

# Transformations probabilistes - possibilistes : modèles graphiques, révision, inférence

## MÉMOIRE

soutenu le 3 juillet 2014

pour l'obtention du

**Master Recherche de l'Université d'Artois**  
**Spécialité Informatique – Option Systèmes Intelligents et Applications**

par

Levray Amélie

*Encadrants :* Karim Tabia      Maître de Conférences  
Salem Benferhat      Professeur des Universités



## **Remerciements**

Je remercie mes tuteurs monsieur Tabia et monsieur Benferhat de m'avoir acceptée sur ce sujet de mémoire, plus particulièrement monsieur Tabia pour son investissement et sa présence qui m'ont permis de mener à bien ce stage. Son aide a été précieuse dans la rédaction de ce mémoire et la préparation de l'exposé.

Ensuite, je remercie monsieur Grégoire, monsieur Marquis de m'avoir permis de faire mon stage de fin d'études au sein du laboratoire. Ainsi que les membres du CRIL, pour ces cinq années d'enseignements de l'informatique.

Enfin, merci à mes petits camarades de master pour la bonne ambiance et l'entraide sur cette dernière année.



# Table des matières

<b>Table des figures</b>	<b>vi</b>
<b>Liste des tableaux</b>	<b>vii</b>
<b>Introduction générale</b>	<b>1</b>
<hr/>	
<b>Chapitre 1 Incertitude et modèles graphiques</b>	
1.1 Introduction . . . . .	3
1.2 Théorie des probabilités . . . . .	4
1.2.1 Distribution de probabilités . . . . .	4
1.2.2 Quelques règles d'inférence . . . . .	5
1.2.3 Réseaux Bayésiens . . . . .	6
1.3 Théorie des possibilités . . . . .	10
1.3.1 Distribution de possibilités . . . . .	10
1.3.2 Conditionnement . . . . .	11
1.3.3 Réseaux possibilistes . . . . .	12
1.3.4 Réseaux possibilistes quantitatifs . . . . .	13
1.3.5 Réseaux possibilistes qualitatifs . . . . .	13
<b>Chapitre 2 Transformations</b>	<b>15</b>
2.1 Principes et objectifs . . . . .	15
2.1.1 Transformation Probabilités $\rightarrow$ Possibilités . . . . .	16
2.1.2 Transformation Possibilités $\rightarrow$ Probabilités . . . . .	16
2.2 Principes des transformations et propriétés . . . . .	16
2.2.1 Principe de consistance de Zadeh . . . . .	16
2.2.2 Principe de consistance de Klir . . . . .	17
2.2.3 Principe de consistance de Dubois et Prade . . . . .	17
2.3 Principales transformations . . . . .	18
2.3.1 Klir transformation ( $KT$ ) . . . . .	18

2.3.2	Optimal transformation ( <i>OT</i> )	18
2.3.3	Symmetric transformation ( <i>ST</i> )	18
2.3.4	Variable transformation ( <i>VT</i> )	19
2.3.5	Exemple	19
2.3.6	Exemples de nouvelles transformations	20
2.3.7	Autres transformations	21
<b>Chapitre 3 Analyse des transformations</b>		<b>23</b>
3.1	Transformations et normalisation	23
3.1.1	Transformation <i>OT</i>	24
3.1.2	Transformation <i>KT</i>	24
3.1.3	Transformation <i>ST</i>	24
3.1.4	Transformation <i>VT</i>	25
3.2	Transformations et requêtes MPE	25
3.3	Préservation de l'ordre des événements	26
3.3.1	<i>OT</i> , une transformation qui ne préserve pas l'ordre des événements	26
3.3.2	Généralisation à toutes les transformations	27
3.4	Transformations et marginalisation	28
3.5	Transformations et conditionnement	30
3.5.1	Cadre quantitatif	30
3.5.2	Cadre qualitatif	31
3.6	Conditionnement avec des informations incertaines	32
3.6.1	Conditionnement de Jeffrey en probabilités	33
3.6.2	Conditionnement de Jeffrey en possibilités	33
3.6.3	Transformations et conditionnement de Jeffrey	34
3.7	Transformations et requêtes MAP	38
<b>Chapitre 4 Analyse des transformations dans les modèles graphiques</b>		<b>40</b>
4.1	Transformations et Indépendances	40
4.1.1	Définitions	41
4.1.2	Transformation <i>OT</i>	41
4.1.3	Transformation <i>KT</i>	43
4.1.4	Transformation <i>ST</i>	44
4.1.5	Transformation <i>VT</i>	45
4.2	Préservation des indépendances	46
4.2.1	Préservation des indépendances d'événements dans le cadre qualitatif	46
4.2.2	Conséquences	48

---

4.3	Transformations et modèles graphiques . . . . .	49
4.3.1	Transformer un réseau Bayésien . . . . .	49
4.3.2	Préservation des distributions jointes . . . . .	50
4.3.3	Préservation des distributions jointes dans le cadre qualitatif . . . . .	51
4.3.4	Préservation des distributions jointes dans le cadre quantitatif . . . . .	53

---

<b>Conclusion et Perspectives</b>	<b>55</b>
-----------------------------------	-----------

**Bibliographie**

# Table des figures

1.1	Exemple de réseau Bayésien . . . . .	8
1.2	Exemple de réseau possibiliste . . . . .	12
3.1	Transformation d'une distribution de probabilités en distribution de possibilités par $OT$ .	26
3.2	Préservation de la marginalisation . . . . .	28
3.3	Préservation du conditionnement . . . . .	31
3.4	Préservation du conditionnement avec des informations incertaines . . . . .	36
4.1	Représentation des événements $\phi$ , $\psi$ et $\alpha$ où $\omega_{max} \in \phi \cap \psi \cap \alpha$ . . . . .	47
4.2	Transformation de modèles graphiques . . . . .	49
4.3	Réseau Bayésien et la transformation des tables locales par $OT$ . . . . .	50
4.4	Représentation de la transformation $TR$ sur un réseau . . . . .	51
4.5	Réseau Bayésien avec des distributions locales et leurs transformées par $ST$ . . . . .	52
4.6	Réseau Bayésien dont les distributions locales sont transformées par $TR$ . . . . .	53
4.7	Réseau Bayésien avec des distributions locales et leurs transformées par $OT$ . . . . .	54
8	Transformations, modèles graphiques et bases de croyances . . . . .	56



# Liste des tableaux

1.1	Exemple de distribution de probabilités . . . . .	4
1.2	Table de probabilités jointe codée par le RB de la Figure 1.1 . . . . .	8
1.3	Table de possibilités jointe codée par le réseau possibiliste de la Figure 1.2 dans le cadre quantitatif . . . . .	13
1.4	Table de possibilités jointe codée par le réseau possibiliste de la Figure 1.2 dans le cadre qualitatif . . . . .	14
2.1	Comparaison de deux distributions ayant le même degré de consistance selon le critère de Zadeh . . . . .	17
2.2	Distribution de probabilités de l'Exemple 12 . . . . .	20
2.3	Distribution de possibilités obtenues par les différentes transformations . . . . .	20
2.4	Probabilités et possibilités des événements sur $2^\Omega$ . . . . .	20
2.5	Résumé des propriétés des transformations . . . . .	21
3.1	Probabilités et possibilités des événements sur $2^\Omega$ . . . . .	26
3.2	Distribution de probabilités représentant la perte de l'ordre des événements . . . . .	27
3.3	Distribution de probabilités où $p(\omega_1) < p(\omega_2) + p(\omega_3)$ et sa transformée . . . . .	28
3.4	Distribution de probabilités et sa transformation par $OT$ . . . . .	29
3.5	Marginalisation de la Table 3.4 sur la variable $C$ . . . . .	29
3.6	Transformation de la distribution $p(A, B)$ de la Table 3.5 par $OT$ . . . . .	29
3.7	Distribution de probabilités pour la preuve de la Proposition 4.6 . . . . .	30
3.8	Distribution de possibilités transformée de la Table 3.7 . . . . .	30
3.9	Distribution de probabilités de l'Exemple 15 . . . . .	34
3.10	Distribution de probabilités révisée de la Table 3.9 . . . . .	35
3.11	Distribution de possibilités résultante de la transformation de la Table 3.9 par $OT$ . . . . .	35
3.12	Distribution de possibilités révisée résultante de la transformation de la Table 3.10 par $OT$ . . . . .	35
3.13	Distribution de possibilités de la Table 3.11 révisée à l'aide de l'Équation 3.17 . . . . .	35
3.14	Distribution de probabilités et sa révision . . . . .	36
3.15	Distribution transformée de la Table 3.14 . . . . .	37
3.16	Distribution de probabilités $p'(A, B)$ de la Table 3.14 transformée par une transformation $TR$ . . . . .	37
3.17	Distribution de possibilités de la Table 3.15 révisée à l'aide de l'Équation 3.17 . . . . .	37
3.18	Distribution de possibilités de la Table 3.15 révisée à l'aide de l'Équation 3.19 . . . . .	37
4.1	Distribution de probabilités et sa transformation par $OT$ . . . . .	42
4.2	Distribution de possibilités de $B$ sachant $AC$ après transformation par $OT$ . . . . .	42
4.3	Distribution de possibilités de $B$ sachant $A$ après transformation par $OT$ . . . . .	42
4.4	Distribution de probabilités et sa transformation par $KT$ . . . . .	43

4.5	Distribution de possibilités de $A$ sachant $C$ après transformation par $KT$ . . . . .	43
4.6	Distribution de possibilités de $A$ après transformation par $KT$ . . . . .	43
4.7	Distribution de probabilités et sa transformation par $ST$ . . . . .	44
4.8	Distribution de possibilités de $B$ sachant $AC$ après transformation par $ST$ . . . . .	44
4.9	Distribution de possibilités de $B$ sachant $A$ après transformation par $ST$ . . . . .	44
4.10	Distribution de probabilités et sa transformation par $VT$ . . . . .	45
4.11	Distribution de possibilités de $B$ sachant $AC$ après transformation par $VT$ . . . . .	46
4.12	Distribution de possibilités de $B$ sachant $A$ après transformation par $VT$ . . . . .	46
4.13	Distribution de probabilités et sa transformée . . . . .	47
4.14	Distribution de probabilités et sa transformation par $KT$ . . . . .	48
4.15	Distribution de possibilités de $A$ sachant $B$ après transformation par $KT$ . . . . .	49
4.16	Distribution de possibilités de $A$ après transformation par $KT$ . . . . .	49
4.17	Comparaison des distributions de possibilités générées par le Réseau 4.3 . . . . .	51
4.18	Distributions de probabilités et possibilités jointes codées par les réseaux de la Figure 4.5	52
4.19	Distribution de possibilités transformée par $ST$ . . . . .	52
4.20	Distributions de probabilités et possibilités jointes codées par le réseau de la Figure 4.7 .	54
4.21	Distribution de possibilités transformée par $OT$ . . . . .	54

# Introduction générale

Le problème de la représentation et du raisonnement dans l'incertain a repris de l'ampleur à l'apparition de l'intelligence artificielle. Le but étant, pour une machine, de modéliser le comportement humain. Plus particulièrement, nous nous intéressons à la manière dont l'information est représentée en présence d'imprécision, d'incertitude et de contradictions. Dans ce domaine, de nombreux cadres existent, le plus connu est la théorie des probabilités. La théorie des probabilités a prouvé son efficacité en présence d'informations empiriques. Mais cette théorie est problématique lorsqu'on fait face à des cas d'ignorance totale ou partielle, ou d'informations incomplètes de manière générale. C'est pour cela qu'il existe des cadres de représentations alternatifs tels que la théorie des possibilités, les probabilités imprécises ou alors la théorie de Dempster-Shafer.

Dans plusieurs problèmes réels, le nombre de variables est important. L'inconvénient de ces cadres est que la taille des représentations augmente de manière exponentielle par rapport au nombre de variables. Nous avons alors besoin de représentations compactes. Les modèles graphiques sont une réponse à ce problème. En effet, les modèles graphiques sont des outils importants pour la représentation et le raisonnement dans l'incertain ; ils factorisent une distribution jointe (de probabilités, de possibilités, etc.) en exploitant les indépendances conditionnelles. Concernant l'utilisation des modèles graphiques, on trouve principalement le raisonnement et l'inférence. En effet, les modèles graphiques disposent d'algorithmes d'inférence efficaces pour répondre à des questions d'intérêt afin de réaliser certaines tâches comme l'explication, le diagnostic, etc. En plus, l'autre grand intérêt des modèles graphiques est leur facilité d'interprétation et d'élicitation, et ils peuvent être construits automatiquement à partir de données.

Étant donné le nombre de cadres utilisés pour la représentation des connaissances et de l'incertain, nous pouvons nous demander comment est-il possible de travailler avec ces différents cadres en même temps, l'utilité donc des transformations d'un cadre à un autre cadre est évidente. L'idée des transformations est d'exprimer les croyances ou les informations incertaines, définies dans un cadre, dans un autre cadre. Le but étant de minimiser la perte d'information, afin que les informations qui nous permettaient d'inférer sur un sujet, nous donne idéalement la même réponse dans le nouveau cadre. Dans ce mémoire, les transformations porteront essentiellement à passer du cadre probabiliste au cadre possibiliste. Dans le cadre possibiliste, nous distinguerons deux échelles différentes dans la manière de représenter les informations. Une échelle quantitative où le degré de possibilités à un sens numérique, alors que l'échelle qualitative exprime juste un ordre sur les croyances.

Dans un premier temps, nous présentons un état de l'art sur la théorie des probabilités et la théorie des possibilités, avec les modèles graphiques associés. Nous étudions ensuite les principales transformations proposées dans la littérature. La plus grande partie de ce mémoire est consacrée à notre contribution ; nous analysons les transformations suivant plusieurs critères, avant de passer à une généralisation à toutes les transformations. Des questions comme la marginalisation, le conditionnement, les requêtes MPE et MAP sont traitées. Nous tentons ensuite de répondre à ces mêmes questions dans les modèles

graphiques.

## Notations

Avant de débiter, il convient de se mettre d'accord sur quelques notations. Dans la suite :

- $\Omega$  est l'univers du discours
- Par convention,  $\omega_i$  est un monde appartenant à l'ensemble  $\Omega$
- $p$  représente une distribution de probabilités et  $\pi$  une distribution de possibilités
- Nous noterons  $TR$  une transformation quelconque
- $V$  représente l'ensemble des variables, que nous notons  $A_i$ , nous utilisons aussi les lettres  $B, C, \dots$
- A chaque variable  $A_i$  est associée un domaine  $D_{A_i}$
- Les lettres grecs  $\alpha, \phi, \psi$  sont utilisées pour les événements. Un événement  $\phi \subseteq \Omega$  est un sous-ensemble de  $\Omega$
- $2^\Omega$  dénote l'ensemble de tous les événements  $\phi \subseteq \Omega$
- La probabilité (resp. la possibilité) d'un monde est notée soit par  $p(\omega_i)$  (resp.  $\pi(\omega_i)$ ) soit par  $p_i$  (resp.  $\pi_i$ )
- $P(a)$  représente la probabilité de l'événement  $\phi = \{A = a\}$ , afin de simplifier la notation
- L'indépendance conditionnelle entre  $X$  et  $Y$  sachant  $Z$  est notée  $X \perp Y | Z$

# Chapitre 1

## Incertitude et modèles graphiques

### Sommaire

---

<b>1.1</b>	<b>Introduction</b>	<b>3</b>
<b>1.2</b>	<b>Théorie des probabilités</b>	<b>4</b>
1.2.1	Distribution de probabilités	4
1.2.2	Quelques règles d'inférence	5
1.2.3	Réseaux Bayésiens	6
<b>1.3</b>	<b>Théorie des possibilités</b>	<b>10</b>
1.3.1	Distribution de possibilités	10
1.3.2	Conditionnement	11
1.3.3	Réseaux possibilistes	12
1.3.4	Réseaux possibilistes quantitatifs	13
1.3.5	Réseaux possibilistes qualitatifs	13

---

### 1.1 Introduction

Avant de dresser une vue d'ensemble des différents cadres de représentation utilisés dans ce mémoire, il convient de préciser un peu le vocabulaire employé. On appelle *information* toute collection de symboles ou de signes produits par l'observation de phénomènes naturels ou artificiels, soit par l'activité cognitive humaine. Une première distinction apparaît entre *informations* dites *objectives*, issue de mesures ou de capteurs ou plus généralement de l'observation directe de phénomènes, et celles dites *subjectives* exprimées par des individus et conçues sans le recours à l'observation directe du réel. L'information peut prendre une forme numérique ou qualitative.

**Définition 1** (Informations incomplètes). Ce sont des informations qui ne permettent pas de répondre de manière satisfaisante et précise aux questions d'intérêt.

**Exemple 1.** Nelson Mandela est né entre 1914 et 1920. Cette information ne permet pas de répondre avec précision à une question du genre *Quel est l'âge de Nelson Mandela ?*

La nature des informations incomplètes n'est pas la même. On distingue *imprécision, flou, incertitude, etc.*

**Définition 2** (Information imprécise). Une information est dite *imprécise* si elle est formulée sous forme d'un ensemble de valeurs contenant la réponse exacte.

**Exemple 2.** Ce morceau de musique est soit de *Bach* soit de *Mozart*.

**Définition 3** (Information floue, vague). Une information est considérée floue si elle est exprimée avec des termes linguistiques qui font intervenir la gradualité et qui ne sont ni tout à fait vraies, ni tout à fait fausses.

**Exemple 3.** Il ne fait pas trop froid dehors.

**Définition 4** (Information incertaine). Une information est dite *incertaine* pour un agent lorsque l'agent ne sait pas si cette information est vraie ou fausse mais l'agent peut privilégier certaines réponses sur d'autres.

**Exemple 4.** Il est très probable qu'il fasse beau en Juillet.

Les deux cadres de représentation utilisés dans ce stage sont les probabilités et les possibilités. Dans cette optique nous allons d'abord définir la théorie des probabilités puis la théorie des possibilités.

## 1.2 Théorie des probabilités

La théorie des probabilités est l'étude *mathématique* des phénomènes caractérisés par l'incertitude. Les débuts de l'étude des probabilités correspondent aux premières observations du hasard dans les jeux par exemple. Les probabilités sont le cadre le plus connu et le plus ancien pour représenter et raisonner avec des informations incertaines. Dans ce qui suit, nous rappelons les principaux concepts de cette théorie, puis nous présenterons les modèles graphiques le plus couramment associés à cette théorie.

### 1.2.1 Distribution de probabilités

**Définition 5** (Mondes possibles). Les mondes possibles sont les états élémentaires dans lesquels le monde considéré peut se trouver. On définit l'univers du discours  $\Omega = \{ \omega_1, \omega_2, \dots, \omega_n \}$  comme étant l'ensemble de tous les mondes possibles ; Dans ce rapport,  $\Omega$  est fini et discret.

**Définition 6** (Distribution de probabilités). Soit  $\Omega$  l'univers du discours, une distribution de probabilités associe à chaque monde  $\omega_i$ , un degré dans l'intervalle  $[0,1]$ .

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

$$\forall \omega_i \in \Omega, \omega_i \rightarrow p(\omega_i) \in [0, 1].$$

**Exemple 5.** Soit  $\Omega = \{ \omega_1, \omega_2, \omega_3, \omega_4 \}$

$\omega_i$	$p(\omega_i)$
$\omega_1$	0.2
$\omega_2$	0.5
$\omega_3$	0.15
$\omega_4$	0.15

TABLE 1.1 – Exemple de distribution de probabilités

La notion de mesure de probabilité est définie sur des événements c'est à dire un ensemble de mondes de  $\Omega$ .

**Définition 7** (Mesure de probabilité). Une mesure de probabilité  $P(\phi)$  associe à chaque événement  $\phi \subseteq \Omega$  un degré (réel) dans  $[0,1]$  qui reflète les chances que  $\phi$  a de se réaliser.  $P(\phi)$  est définie comme suit :

$$P(\phi) = \sum_{\omega_i \in \phi} p(\omega_i) \quad (1.1)$$

La théorie de probabilités a été axiomatisée dans les années 1930 par Andrei Kolmogorov, un mathématicien russe.

### Axiomes de Kolmogorov

**Axiome 1** (Non Négativité).

$$\phi \subseteq \Omega, 0 \leq P(\phi) \leq 1 \quad (1.2)$$

Un événement faisant partie de l'univers du discours peut dans le pire des cas être impossible (probabilité de 0) ou sûr dans le meilleur des cas (probabilité de 1).

**Axiome 2** (Normalisation).

$$P(\Omega) = 1 \Leftrightarrow \sum_{i=1}^n p(\omega_i) = 1 \quad (1.3)$$

La masse de probabilité est entièrement affectée aux événements de l'univers du discours.

**Axiome 3** (Additivité).

$$P(\phi \cup \psi) = P(\phi) + P(\psi) \text{ si } \phi \text{ et } \psi \text{ sont disjoints} \quad (1.4)$$

Les chances d'occurrence de l'un de deux événements disjoints sont la somme des chances d'occurrence des événements individuellement.

## 1.2.2 Quelques règles d'inférence

Étant donné une distribution de probabilités  $p$ , pour inférer les probabilités d'un événement  $\phi$ , nous utilisons certaines règles d'inférence.

**Définition 8** (Règle de marginalisation). Soit  $\{A_1, \dots, A_k\} \subseteq V$ ,

$$P(A_1..A_k) = \sum_{A_{k+1}..A_n} (p(A_1A_2..A_n)) \quad (1.5)$$

*Remarque.* Une distribution marginale est une distribution de probabilités, elle vérifie donc les trois axiomes de la Section 1.2.1.

**Définition 9** (Probabilités conditionnelles). Soient  $\phi$  et  $\psi$  deux événements. La probabilité de  $\phi$  étant donné  $\psi$  est calculée comme suit :

$$P(\phi|\psi) = \frac{P(\phi \cap \psi)}{P(\psi)} \quad (1.6)$$

Pour calculer  $P(\phi|\psi)$ , on ne retient que les modèles appartenant à  $\phi \cap \psi$  puisque  $\psi$  est observé. On normalise en divisant sur  $P(\psi)$ .

On déduit de la Définition 9, la règle du produit qui nous permet de calculer n'importe quelle intersection d'événements.

**Définition 10** (Règle du produit). Soient  $\phi$  et  $\psi$  deux événements.

$$P(\phi \cap \psi) = P(\phi) * P(\psi|\phi) = P(\psi) * P(\phi|\psi) \quad (1.7)$$

La règle de chaînage permet le calcul de n'importe quel membre d'une distribution sur un ensemble de variables aléatoires en utilisant seulement le conditionnement probabiliste.

**Définition 11** (Règle de chaînage). Soit  $V = \{A_1, \dots, A_n\}$

$$P(A_1 A_2 \dots A_n) = P(A_1) * P(A_2|A_1) * \dots * P(A_n|A_1 \dots A_{n-1}) \quad (1.8)$$

**Définition 12** (Règle de Bayes).  $\forall \phi \subseteq \Omega, \forall \psi \subseteq \Omega,$

$$P(\phi|\psi) = \frac{P(\phi) * P(\psi|\phi)}{P(\psi)} \quad (1.9)$$

La règle de Bayes permet d'exprimer  $P(\phi | \psi)$  en fonction de  $P(\psi | \phi)$  (inverser les probabilités). Cela permet, par exemple, de calculer les probabilités des causes si les effets sont connus et les probabilités de ces effets étant données les causes.

Travailler directement sur les distributions n'est pas pratique en terme de complexité (spatiale et temporelle). En effet, la taille de la table peut vite devenir trop volumineuse, nous avons donc besoin de représentations compactes, les modèles graphiques en général, nous fournissent un outil efficace pour la représentation de nos informations. Et parmi-eux, les réseaux Bayésiens, l'une des représentations compactes les plus utilisées.

### 1.2.3 Réseaux Bayésiens

Un réseau Bayésien [Pea89, Dar09] est un système représentant de manière compacte la connaissance et permettant de calculer des probabilités d'intérêt apportant des solutions à différentes sortes de problématiques.

L'intérêt particulier des réseaux Bayésiens est de combiner simultanément des connaissances d'experts et des informations extraites à partir de données. Ils sont utilisés principalement pour le diagnostic (médical ou industriel), l'analyse de risques, la détection de spams, le datamining, modélisation et simulation de systèmes complexes, détection d'intrusions, etc.

Pour résumer, un **réseau Bayésien** est un modèle graphique permettant d'acquérir, de représenter et raisonner avec des **informations complexes et incertaines**.

Dans ce qui suit, un *graphe* correspond à une paire  $G = (V, E)$  où  $V$  constitue l'ensemble des *somets*, aussi connus sous le nom de *nœuds*, et  $E$  est l'ensemble des *liens* entre les nœuds dans  $V$ . Un *arc* est une arête orientée. Un arc de  $A_i$  vers  $A_j$  est noté  $A_i \rightarrow A_j$ .  $A_i$  est appelé *origine* et  $A_j$  l'*extrémité*. Un graphe constitué d'arcs est appelé *graphe orienté*. Pour les graphes orientés, il existe la notion de parents et d'enfants. S'il existe un arc  $A_i \rightarrow A_j$ ,  $A_i$  est appelé *parent* de  $A_j$  et  $A_j$  l'*enfant* de  $A_i$ . L'ensemble des parents de  $A_i$  est noté par  $par(A_i)$ . Une *racine* est un nœud qui n'a pas de parents, alors



qu'une *feuille* est un nœud qui n'a pas d'enfants.

Pearl et ses collègues [GVP90] ont établi que les informations incertaines peuvent être gérées plus efficacement en tirant profit de la notion d'indépendance conditionnelle. Un modèle graphique permet de représenter les relations d'indépendances conditionnelles qui existent entre les variables (ou ensembles de variables).

**Définition 13** (Indépendances conditionnelles). Soient  $X$ ,  $Y$  et  $Z$  trois sous-ensembles de variables. Dire que  $X$  et  $Y$  sont indépendants conditionnellement à  $Z$  revient à dire que :

$$P(X | YZ) = P(X | Z) \quad (1.10)$$

Ces relations d'indépendances conditionnelles se caractérisent par les propriétés suivantes [PP86] : Soient  $X$ ,  $Y$ ,  $W$  et  $Z$  des sous-ensembles de variables dans  $V$ .

- 1 (Symétrie)  $X \perp Y | Z \Leftrightarrow Y \perp X | Z$
- 2 (Décomposition)  $X \perp YW | Z \Rightarrow X \perp Y | Z$  et  $X \perp W | Z$
- 3 (Faible union)  $X \perp YW | Z \Rightarrow X \perp Y | ZW$
- 4 (Contraction)  $X \perp Y | Z$  et  $X \perp W | ZY \Rightarrow X \perp YW | Z$
- 5 (Intersection)  $X \perp W | ZY$  et  $X \perp Y | ZW \Rightarrow X \perp YW | Z$

où  $X \perp Y | Z$  signifie que  $X$  est indépendant de  $Y$  dans le contexte de  $Z$ .

### Structure d'un réseau Bayésien

Un réseau Bayésien est défini par :

- **Un composant graphique** : Les variables et les arcs forment un graphe orienté acyclique (DAG). Chaque sommet du DAG représente une variable du problème modélisé et les arcs représentent les indépendances entre les variables.
- **Un composant numérique** : Chaque lien dans le DAG est quantifié par une distribution conditionnelle  $p(A_i | par(A_i))$  pour chaque nœud  $A_i$  dans le contexte de ses parents  $par(A_i)$ . Pour les variables sans parents, on associe aussi une table locale a priori (sans parents).

Les distributions locales doivent satisfaire les contraintes de **normalisation** suivantes :

- Si  $par(A_i) = \emptyset$  ( $A_i$  est une racine) alors la distribution marginale associée à  $A_i$  doit satisfaire :

$$\sum_{a_i \in D_{A_i}} P(a_i) = 1 \quad (1.11)$$

- Si  $par(A_i) \neq \emptyset$  alors la distribution conditionnelle associée à  $A_i$  doit satisfaire :

$$\forall u_i \in D_{par(A_i)}, \sum_{a_i \in D_{A_i}} P(a_i | u_i) = 1 \quad (1.12)$$

où  $D_{par(A_i)}$  représente le domaine associé aux variables parentes de  $A_i$ .

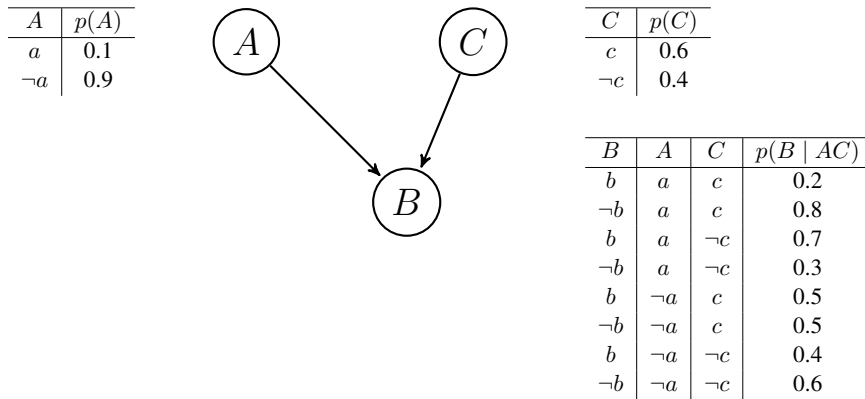


FIGURE 1.1 – Exemple de réseau Bayésien

**Exemple 6.** Soient 3 variables  $A, B, C$  avec pour domaines respectifs  $D_A = \{a, \neg a\}$ ,  $D_B = \{b, \neg b\}$  et  $D_C = \{c, \neg c\}$ .

Nous remarquons que les tables de probabilités locales de la Figure 1.1 sont toutes normalisées.

Un réseau Bayésien code une distribution jointe qui se calcule avec la règle de chaînage.

**Définition 14** (Règle de chaînage). Soient  $A_1, \dots, A_n$  des variables aléatoires. Un réseau Bayésien permet d’écrire la table de probabilités jointe sous une forme factorisée en utilisant la règle de chaînage :

$$P(A_1, \dots, A_n) = \prod_{i=1}^n P(A_i | \text{par}(A_i)) \tag{1.13}$$

*Remarque.* Si toutes les tables locales sont normalisées alors la distribution jointe est également normalisée.

**Exemple 7.** La distribution de probabilité jointe du réseau Bayésien de la Figure 1.1, calculée à partir de la règle de chaînage (1.13) est présentée dans la Table 1.2.

$A$	$B$	$C$	$p(A, B, C)$
$\neg a$	$\neg b$	$\neg c$	0.216
$\neg a$	$\neg b$	$c$	0.27
$\neg a$	$b$	$\neg c$	0.144
$\neg a$	$b$	$c$	0.27
$a$	$\neg b$	$\neg c$	0.012
$a$	$\neg b$	$c$	0.048
$a$	$b$	$\neg c$	0.028
$a$	$b$	$c$	0.012

TABLE 1.2 – Table de probabilités jointe codée par le RB de la Figure 1.1

A titre d’exemple, pour calculer  $p(a \neg b \neg c)$ , selon la structure du graphe :

$$p(a \neg b \neg c) = p(a) \times p(\neg b | a \neg c) \times p(\neg c) = 0.1 \times 0.3 \times 0.4 = 0.012$$

En utilisant la table de probabilités jointe, nous pouvons appliquer les règles d’inférence de la Section 1.2.2 afin de calculer n’importe quelle probabilité d’intérêt. Si nous voulons calculer  $p(a)$  par

exemple, nous utilisons la règle de marginalisation comme ceci :

$$p(a) = \sum_{b_i \in D_B, c_j \in D_C} p(a, b_i, c_j) = 0.012 + 0.048 + 0.028 + 0.012 = 0.1$$

La règle de Bayes nous permet de calculer des probabilités conditionnelles à partir d'autres probabilités conditionnelles comme sur cet exemple :

$$P(ac | b) = \frac{P(b | ac) * P(ac)}{P(b)} = \frac{P(b | ac) * P(a) * P(c)}{P(b)} \approx 0,026$$

L'inconvénient de la table de probabilités jointe est que lorsque le nombre de variables augmente, la table devient très coûteuse à construire car sa taille est exponentielle en fonction du nombre de variables. C'est pour cela que les algorithmes d'inférence travaillent directement sur le réseau Bayésien.

### Inférence

Les réseaux Bayésiens ne servent pas uniquement à la représentation mais aussi au raisonnement (calculer des probabilités d'intérêt). Les modèles graphiques sont dotés de mécanismes d'inférence pour répondre à des questions d'intérêt afin de réaliser certaines tâches comme l'explication, le diagnostic, la classification, etc.

Certaines informations sont disponibles dans les tables de probabilités locales. Lorsque ce n'est pas le cas, nous utilisons l'inférence. L'inférence dans un réseau Bayésien est, en général, la propagation d'une ou plusieurs informations certaines appelées évidences et la mise à jour des autres informations. Cela revient à calculer la probabilité a posteriori de certaines variables non observées étant donné un ensemble de variables observées. Il existe trois principales catégories de requêtes (donc d'inférence) :

- **Probabilité d'un événement** : Dans ces requêtes, nous souhaitons mesurer la probabilité d'un événement donné  $\phi \subseteq \Omega$ . Il s'agit de calculer  $P(\phi)$  depuis le réseau. Les mécanismes utilisés pour ce type de requêtes sont généralement la règle de chaînage et la marginalisation. Dans le cas général, on cherche à calculer  $P(\phi | \psi)$  où  $\psi$  est une observation.
- **Explication la plus probable (MPE)** : L'objectif ici est d'identifier la plus probable instantiation  $v^*$  (tuple de valeurs) de toutes les variables  $V$  du réseau qui soit cohérente avec une instantiation  $e$  d'un sous-ensemble de variables  $E$ . C'est à dire chercher la meilleure explication pour  $E = e$ , calculée comme suit :

$$v^* = \operatorname{argmax}_v (P(v, e)) \quad (1.14)$$

- **Maximum a posteriori** : Dans les requêtes de type maximum a posteriori (MAP), il s'agit de calculer l'hypothèse  $h^*$ , instance de  $H$  (variables d'hypothèses ou variables non observées), la plus probable pour une observation donnée  $e$  (instance des variables observées  $E$ ). Les sous-ensembles  $E$  et  $H$  sont disjoints. Dans les problèmes de classification, par exemple, il s'agit de trouver la classe (variable non observable) la plus probable étant donné les caractéristiques de l'objet à classer (l'évidence ici concerne les valeurs des variables caractérisant les objets étudiés). L'inférence MAP est basée sur la formule suivante :

$$h^* = \operatorname{argmax}_h (P(h | e)) \quad (1.15)$$

De manière générale, les mécanismes utilisés pour l'inférence dans les réseaux Bayésiens sont le conditionnement, la règle de Bayes, la marginalisation et la règle de chaînage. Les réseaux Bayésiens

permettent, grâce à ces mécanismes, de calculer n'importe quelle probabilité d'intérêt qu'on peut utiliser en classification, explication, etc. Notons que les algorithmes d'inférence, de par la nature des résultats qu'ils fournissent, se répartissent en algorithmes d'inférence exacte et en algorithmes d'inférence approchée. Les algorithmes d'inférence dans les réseaux Bayésiens peuvent être répartis selon leurs approches en trois catégories :

1. *Inférence par élimination de variables* : Le principe de ces algorithmes est d'éliminer de proche en proche et par marginalisation certaines variables jusqu'à répondre à la requête. L'ordre d'élimination dépend de la requête et de la structure du réseau. La complexité en temps de ce type d'algorithmes est exponentielle dans la largeur (en terme de nombre de variables) des facteurs (tables relatives à des sous-ensembles de variables) rencontrés lors de l'élimination [ZP96].
2. *Inférence par compilation* : Nous pouvons mettre dans cette catégorie tous les algorithmes qui passent par des structures ou langages intermédiaires pour répondre (efficacement) aux requêtes posées. Parmi ces algorithmes, certains sont basés sur des structures d'arbres. Le plus connu de ces algorithmes est l'arbre de jonction [Jen96] où le réseau Bayésien de départ est compilé en arbre de clusters et séparateurs. Après compilation, certains calculs se font de manière locale aux clusters et séparateurs. La complexité de l'inférence dépend ici de la largeur de l'arbre (consistant en la largeur du plus grand cluster).
3. *Inférence par conditionnement* : L'idée principale dans cette approche est d'exploiter l'évidence sur laquelle porte la requête afin de simplifier le réseau initial en arbre ou polyarbre. À titre d'exemple, les observations sont intégrées en déconnectant les variables observées de leurs descendants directs et en mettant à jour les tables locales de ces derniers. Des algorithmes d'inférence dits par instanciation exploitent également cette idée en instanciant certaines variables et réalisant des inférences sur les réseaux résultants [Pea88].

*Remarque (Complexité).* Bien que le problème d'inférence dépende de la complexité du réseau (plus la topologie du réseau est simple, plus l'inférence est efficace), il a été démontré par G.F Cooper [Coo90] que dans le cas général, l'inférence dans les réseaux Bayésiens est un problème NP-difficile. Des détails récents et une synthèse des principaux résultats de complexité de l'inférence dans les réseaux Bayésiens sont disponibles dans un papier de de Campos et al. [dC11].

La théorie des probabilités souffre de beaucoup de problèmes connus dans les cas d'ignorance totale ou partielle. Dans de telles situations, le cadre probabiliste standard conduit souvent à des inconsistances et des résultats problématiques. C'est la raison pour laquelle des théories alternatives ont été mises au point, comme la théorie des possibilités.

## 1.3 Théorie des possibilités

La théorie des possibilités, introduite par Zadeh [Zad99] et développée par Dubois et Prade [DP88], se situe à la croisée entre les ensembles flous, les probabilités et le raisonnement non-monotone. Cette théorie évalue l'incertitude sur un ensemble totalement ordonné de valeurs, appelé échelle possibiliste, d'une manière qualitative ou quantitative.

### 1.3.1 Distribution de possibilités

Comme la théorie des probabilités, la théorie des possibilités est basée sur la notion de distribution de possibilités. Une distribution de possibilités, notée  $\pi$ , est une application d'un ensemble d'états ou de

mondes possibles  $\Omega$  vers l'échelle  $[0,1]$  traduisant une connaissance partielle sur le monde. Par convention,  $\pi(\omega_i) = 1$  correspond à un état complètement possible,  $\pi(\omega_i) = 0$  correspond à un état impossible.

Il existe également des axiomes pour la théorie des possibilités :

1.  $\Pi(\emptyset) = 0$
2.  $\Pi(\Omega) = 1$  (normalisation)
3. Si  $\phi \subseteq \Omega$  et  $\psi \subseteq \Omega$  alors  $\Pi(\phi \cup \psi) = \max(\Pi(\phi), \Pi(\psi))$  (maxitivité)

Une distribution de possibilités normalisée exprime qu'un de ses états est totalement possible, ce qui se traduit par la condition :

$$\max_{\omega_i \in \Omega}(\pi(\omega_i)) = 1 \quad (1.16)$$

Si  $\max_{\omega_i \in \Omega}(\pi(\omega_i)) < 1$ , ceci indique une contradiction interne (ou inconsistance) dans la représentation.

**Mesures de nécessité et de possibilité :** Dire qu'un événement  $\phi$  est impossible n'implique pas seulement que l'événement complémentaire  $\bar{\phi}$  est possible mais aussi qu'il est certain (nécessaire). Deux mesures duales sont utilisées : la mesure de possibilité, et la mesure de nécessité. La possibilité d'un événement  $\phi$ , noté  $\Pi(\phi)$  est obtenue par la formule  $\Pi(\phi) = \max_{\omega_i \in \phi}(\pi(\omega_i))$  et reflète la situation la plus normale dans laquelle  $\phi$  est vrai. Soit  $\bar{\phi}$  le complémentaire de  $\phi$ . La nécessité, notée  $N(\phi)$ , d'un événement  $\phi$ , définie par la formule  $N(\phi) = 1 - \Pi(\bar{\phi})$ , reflète la situation la plus normale dans laquelle  $\phi$  est faux. La distance entre  $N(\phi)$  et  $\Pi(\phi)$  évalue le niveau d'ignorance sur  $\phi$ .

La particularité de l'échelle possibiliste est qu'elle peut s'interpréter de deux manières :

- **échelle ordinale** : De manière ordinale, si les valeurs affectées ne reflètent que l'ordre entre les différents mondes possibles (l'écart entre les valeurs n'a aucune importance).
- **échelle numérique** : De manière quantitative, si les valeurs affectées prennent un sens sur l'intervalle (l'écart entre deux degrés est pris en compte).

Cette particularité implique des définitions pour le conditionnement différentes, un conditionnement quantitatif (*ratio scale*) et un conditionnement qualitatif (*ordinal*), cependant il existe d'autres types de conditionnement dans le cadre possibiliste mais le plus connu est celui de Hisdal [His78].

### 1.3.2 Conditionnement

Dans le cadre possibiliste, le conditionnement consiste à modifier la distribution de possibilités initiale  $\pi$  à l'arrivée d'une nouvelle information. En fait, nous devons restreindre les états possibles à ceux où la nouvelle information est vraie.

Dans le cadre quantitatif, le conditionnement du degré de l'interprétation  $\omega$  étant donné un événement  $\phi \subseteq \Omega$  est défini comme suit :

$$\pi(\omega_i | \phi) = \begin{cases} \frac{\pi(\omega_i)}{\Pi(\phi)} & \text{si } \omega_i \in \phi \\ 0 & \text{sinon.} \end{cases} \quad (1.17)$$

Notons que c'est exactement la même définition qu'en théorie des probabilités : elle préserve la valeur relative des degrés de possibilités des éléments de  $\phi$ . La seule différence est que  $\Pi(\phi)$  est calculée avec

la règle du maximum et non de la somme.

Dans le cadre qualitatif, le conditionnement est donné par :

$$\pi(\omega_i | \phi) = \begin{cases} 1 & \text{si } \pi(\omega_i) = \Pi(\phi) \text{ et } \omega_i \in \phi \\ \pi(\omega) & \text{si } \pi(\omega_i) < \Pi(\phi) \text{ et } \omega_i \in \phi \\ 0 & \text{sinon.} \end{cases} \quad (1.18)$$

Le conditionnement qualitatif exclut les contre-modèles de  $\phi$  et normalise les modèles de  $\phi$  en mettant le modèle le plus plausible de  $\phi$  à 1.

Tout comme en probabilités, une distribution de possibilités peut vite devenir trop volumineuse, c'est pourquoi un modèle graphique adapté aux possibilités a été introduit : les réseaux possibilistes [BGK02].

### 1.3.3 Réseaux possibilistes

Tout comme les réseaux Bayésiens, les réseaux possibilistes sont composés de deux éléments :

- **une composante graphique** : qui est un DAG  $G = (V, E)$  où  $V$  désigne l'ensemble des nœuds représentant les variables et  $E$  représente l'ensemble des arêtes qui encode les liens entre les nœuds.
- **une composante numérique** : qui quantifie les différents liens. L'incertitude de chaque nœud  $A_i$  est représentée par une table de possibilités locale dans le contexte de ses parents (désigné par  $par(A_i)$ ).

Les distributions locales doivent respecter la contrainte de normalisation pour chaque variable

$$\max_{a_i \in D_{A_i}} (\Pi(a_i | u_i)) = 1, \forall u_i \in D_{par(A_i)} \quad (1.19)$$

**Exemple 8.** Soient 3 variables  $A, B, C$  avec pour domaines respectifs  $D_A = \{a, -a\}$ ,  $D_B = \{b, -b\}$  et  $D_C = \{c, -c\}$ . Un exemple de réseau possibiliste est représenté par la Figure 1.2.

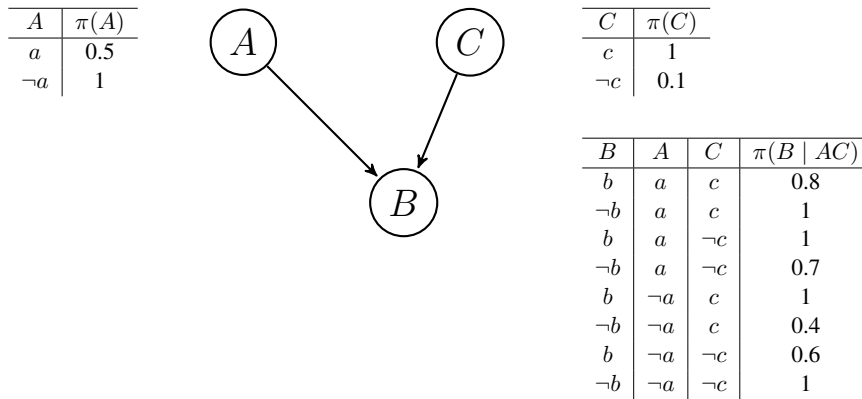


FIGURE 1.2 – Exemple de réseau possibiliste

Nous différencions deux types de réseaux possibilistes : les réseaux possibilistes qualitatifs correspondant à l'interprétation ordinaire de l'échelle possibiliste et les réseaux possibilistes quantitatifs correspondant à l'interprétation numérique. Il est bien connu que les réseaux possibilistes quantitatifs sont proches des réseaux Bayésiens puisque qu'ils partagent les mêmes caractéristiques (essentiellement, l'opérateur du produit) avec des résultats théoriques et pratiques presque identiques.

### 1.3.4 Réseaux possibilistes quantitatifs

Les réseaux possibilistes quantitatifs (basés sur le produit) sont très proches en termes de conditionnement et d'indépendance des réseaux Bayésiens.

La distribution de possibilités  $\pi$  associée à un réseau possibiliste quantitatif est calculée sur un ensemble de variables  $V = \{A_1, \dots, A_n\}$  de la même manière que les réseaux Bayésiens en utilisant la **règle de chaînage** suivante :

$$\pi(A_1, \dots, A_n) = \prod_{i=1}^n \pi(A_i \mid \text{par}(A_i)) \quad (1.20)$$

où  $\pi(A_i \mid \text{par}(A_i))$  correspond à la distribution de possibilités locale associée à la variable  $A_i$  dans le contexte de ses parents.

**Exemple 9.** Soit le réseau possibiliste représenté par le graphe donné par la Figure 1.2. La distribution de possibilités jointe définie par :  $\pi(A, B, C) = \pi(A) \cdot \pi(B \mid AC) \cdot \pi(C)$  est donnée dans la Table 1.3.

$A$	$B$	$C$	$\pi(A, B, C)$
$\neg a$	$\neg b$	$\neg c$	0.1
$\neg a$	$\neg b$	$c$	0.4
$\neg a$	$b$	$\neg c$	0.06
$\neg a$	$b$	$c$	1
$a$	$\neg b$	$\neg c$	0.035
$a$	$\neg b$	$c$	0.5
$a$	$b$	$\neg c$	0.05
$a$	$b$	$c$	0.4

TABLE 1.3 – Table de possibilités jointe codée par le réseau possibiliste de la Figure 1.2 dans le cadre quantitatif

### 1.3.5 Réseaux possibilistes qualitatifs

Un réseau possibiliste qualitatif code de manière compacte une distribution de possibilité jointe sur un ensemble de variables  $V = \{A_1, \dots, A_n\}$  avec la **règle de chaînage** suivante :

$$\pi(A_1, \dots, A_n) = \min_{i=1, \dots, n} \pi(A_i \mid \text{par}(A_i)) \quad (1.21)$$

où  $\pi(A_i \mid \text{par}(A_i))$  correspond à la distribution de possibilités locale associée à la variable  $A_i$  dans le contexte de ses parents.

**Exemple 10.** Soit le réseau possibiliste représenté par le graphe donné par la Figure 1.2. La distribution de possibilité jointe définie par :  $\pi(A, B, C) = \min(\pi(A), \pi(B \mid AC), \pi(C))$  est donnée dans la Table 1.4

## Conclusion

Pour conclure, et avant de présenter les transformations dans le chapitre suivant, nous venons de présenter la théorie des probabilités et la théorie des possibilités qui sont parmi les principaux cadres

$A$	$B$	$C$	$\pi(A, B, C)$
$\neg a$	$\neg b$	$\neg c$	0.1
$\neg a$	$\neg b$	$c$	0.4
$\neg a$	$b$	$\neg c$	0.1
$\neg a$	$b$	$c$	1
$a$	$\neg b$	$\neg c$	0.1
$a$	$\neg b$	$c$	0.5
$a$	$b$	$\neg c$	0.1
$a$	$b$	$c$	0.5

TABLE 1.4 – Table de possibilités jointe codée par le réseau possibiliste de la Figure 1.2 dans le cadre qualitatif

pour représenter et raisonner avec l’incertain. Ainsi que les modèles graphiques, associés à chaque théorie, nous permettant d’avoir des représentations compactes et d’utiliser des algorithmes exploitant les indépendances pour réaliser des tâches d’inférence.

Le chapitre suivant sera consacré aux transformations permettant de passer d’un cadre probabiliste à un cadre possibiliste et vice-versa.



# Chapitre 2

## Transformations

### Sommaire

---

<b>2.1 Principes et objectifs</b> . . . . .	<b>15</b>
2.1.1 Transformation Probabilités $\rightarrow$ Possibilités . . . . .	16
2.1.2 Transformation Possibilités $\rightarrow$ Probabilités . . . . .	16
<b>2.2 Principes des transformations et propriétés</b> . . . . .	<b>16</b>
2.2.1 Principe de consistance de Zadeh . . . . .	16
2.2.2 Principe de consistance de Klir . . . . .	17
2.2.3 Principe de consistance de Dubois et Prade . . . . .	17
<b>2.3 Principales transformations</b> . . . . .	<b>18</b>
2.3.1 Klir transformation ( <i>KT</i> ) . . . . .	18
2.3.2 Optimal transformation ( <i>OT</i> ) . . . . .	18
2.3.3 Symmetric transformation ( <i>ST</i> ) . . . . .	18
2.3.4 Variable transformation ( <i>VT</i> ) . . . . .	19
2.3.5 Exemple . . . . .	19
2.3.6 Exemples de nouvelles transformations . . . . .	20
2.3.7 Autres transformations . . . . .	21

---

Il existe plusieurs représentations de l'incertain et chacune de ces représentations est préférable pour certains types d'incertitude et de problèmes. Il existe également des transformations permettant de passer d'un formalisme à un autre.

Le processus de transformation des probabilités aux possibilités a déjà été étudié dans le passé, par plusieurs chercheurs, mais nous pouvons voir que beaucoup de ces études se limitent aux principes que doivent satisfaire les transformations et proposent les transformations les satisfaisant. Lorsque les informations que nous avons à propos d'un événement, sont à la fois exprimées en probabilités et en possibilités, ces deux descriptions doivent en théorie être cohérentes. Ce qui suggère le principe de consistance. D'autres principes ont été proposés pour garantir d'autres propriétés.

Dans ce chapitre, nous présenterons les objectifs et les principes que doivent satisfaire les transformations. Nous présenterons différentes transformations proposées dans la littérature.

### 2.1 Principes et objectifs

D'une personne à une autre, le cadre choisi pour exprimer ses croyances peut-être différent, c'est pourquoi il y a nécessité de pouvoir passer d'un cadre à un autre sans perte d'informations. L'intérêt est,

par exemple, de pouvoir fusionner les croyances de plusieurs agents.

De plus, transformer des mesures de possibilités en mesure de probabilités ou inversement peut être utile dans n'importe quel problème où on traite des informations incertaines (e.g. des données statistiques).

### 2.1.1 Transformation Probabilités → Possibilités

Transformer une mesure de probabilités en mesure de possibilités peut être utile en présence de sources d'informations plus faibles, ou lorsque calculer avec des possibilités est plus facile que de calculer avec des probabilités. La théorie des possibilités est une théorie mathématique qui traite différents types d'incertitude et est par conséquent une alternative à la théorie des probabilités. La théorie des possibilités est consacrée à manipuler l'information incomplète. De ce fait, en présence d'ignorance, l'analyse est bien plus facilement supportée en possibilité, ou alors si on veut faire du raisonnement qualitatif.

Lorsqu'il y a peu de données, par exemple dans un cadre fréquentiste, où nous avons besoin d'un grand nombre de données, il est difficile d'inférer en probabilité et pour autant en tirer des conclusions fiables. Par exemple, l'article de Masson et Denoeux [MD06] décrit l'utilisation des transformations probabilités-possibilités dans le cas où l'on travaille avec des données empiriques insuffisantes.

### 2.1.2 Transformation Possibilités → Probabilités

La théorie des possibilités a été développée par nécessité de travailler avec des informations incertaines et incomplètes. Ce qui fait qu'il existe beaucoup plus d'outils pour raisonner en probabilités qu'en possibilités. Avoir une transformation qui préserve toutes les informations nous permettrait d'utiliser ces outils, au lieu de devoir les adapter aux possibilités. Par exemple, JavaBayes [Coz01], un puissant outil qui gère les réseaux Bayésiens et est développé pour l'inférence.

Il est aussi intéressant d'avoir de telles transformations dans le cas de prise de décision [Sme89] et d'informations multi-sources.

## 2.2 Principes des transformations et propriétés

Afin de décrire les principales transformations, plusieurs propriétés, appelées *principes de consistance* ont été proposés dans la littérature.

### 2.2.1 Principe de consistance de Zadeh

Zadeh [Zad99] définit le principe de consistance entre probabilités et possibilités comme "*a high degree of possibility does not imply a high degree of probability, and a low degree of probability does not imply a low degree of possibility*". Le degré de consistance entre une distribution de probabilités  $p$  et une distribution de possibilités  $\pi$  est défini par :

$$C_z(\pi, p) = \sum_{i=1..n} \pi_i * p_i \quad (2.1)$$

Cependant, Zadeh signale que  $C_z(\pi, p)$  n'est pas une loi précise ou une relation entre les distributions de probabilités et de possibilités. "*Rather it is an approximate formalization of the heuristic observation that a lessening of the possibility of a event tends to lessen its probability, but not vice-versa*".

En effet, le degré de consistance est d'autant plus proche de 1 que  $\forall \omega_i \in \Omega$ , si  $p(\omega_i) > 0$  alors  $\pi(\omega_i) = 1$ . Nous pouvons avoir plusieurs distributions (différentes) mais ayant le même degré de consistance, est-ce que cela veut dire qu'aucune d'elles n'est meilleure que les autres ? Selon quel critère en choisir une ?

**Exemple 11.** Soit  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ ,  $p$  une distribution de probabilités et  $\pi_1, \pi_2$  deux distributions de possibilités résultantes d'une transformation de  $p$ . La distribution  $\pi_1$  représente l'ignorance totale, alors que  $\pi_2$  représente une ignorance partielle.

	$p$	$\pi_1$	$\pi_2$
$\omega_1$	0.2	1	1
$\omega_2$	0.1	1	1
$\omega_3$	0.3	1	1
$\omega_4$	0.4	1	1
$\omega_5$	0	1	0

TABLE 2.1 – Comparaison de deux distributions ayant le même degré de consistance selon le critère de Zadeh

À travers les distributions de la Table 2.1, nous constatons que  $C_z(\pi_1, p) = 1$  et  $C_z(\pi_2, p) = 1$ , or ici clairement, la distribution  $\pi_2$  est meilleure que  $\pi_1$  qui représente l'ignorance totale.

### 2.2.2 Principe de consistance de Klir

Le concept de consistance a été redéfini par Klir [KG93]. En supposant que les éléments de  $\Omega$  sont ordonnés tel que  $p_i > 0$  et  $p_i \geq p_{i+1}$ ,  $\forall i = \{1..n\}$ . Une transformation devrait se baser sur ces postulats :

- Le *postulat "scaling"* force chaque valeur  $\pi_i$  à être une fonction de  $\frac{p_i}{p_1}$  (où  $p_1 \geq \dots \geq p_n$ ).
- Le *postulat "d'invariance de l'incertain"* selon lequel  $p$  et  $\pi$  doivent avoir la même quantité d'incertain.
- La *condition de consistance* qui établit que ce qui est probable doit être possible, alors  $\pi$  peut être vu comme un majorant de  $p$ .

### 2.2.3 Principe de consistance de Dubois et Prade

Dubois et Prade [PD80] ont donné un exemple pour montrer que le postulat "*scaling*" de Klir peut parfois amener à la violation du principe de consistance. Le deuxième postulat est aussi discutable parce qu'il suppose que les mesures d'information probabiliste et possibiliste sont commensurables.

Ainsi, Dubois et Prade ont défini le principe de consistance, différemment, utilisant ces postulats : [DFMP04]

- La *condition de consistance* :  $P(\phi) \leq \Pi(\phi)$ ,  $\forall \phi \subseteq \Omega$ .
- La *préservation de la préférence* : Supposons que  $\pi$  prend la même forme que  $p$ , alors  $\forall (\omega_1, \omega_2) \in \Omega^2$ ,  $p(\omega_1) > p(\omega_2) \Rightarrow \pi(\omega_1) > \pi(\omega_2)$  et  $p(\omega_1) = p(\omega_2) \Rightarrow \pi(\omega_1) = \pi(\omega_2)$ .
- Le principe du *maximum de spécificité* : Il faut chercher la distribution la plus spécifique qui satisfait les deux postulats précédents.

La notion de spécificité permet de sélectionner une distribution de possibilités en fonction de la quantité d'informations contenue dans cette distribution. La spécificité est définie comme suit :

**Définition 15** (Spécificité). Soit  $\pi_1$  et  $\pi_2$  deux distributions de possibilités sur  $\Omega$ , alors  $\pi_2$  est dite plus spécifique que  $\pi_1$  si et seulement si :  $\forall \omega \in \Omega, \pi_2(\omega) \leq \pi_1(\omega)$ .

## 2.3 Principales transformations

Plusieurs règles de transformations ont été proposées dans la littérature. On présentera les plus connues, *Klir transformation (KT)*, *Optimal transformation (OT)*, *Symmetric transformation (ST)*, et *Variable transformation (VT)* [SAA13].

### 2.3.1 Klir transformation (KT)

Supposons que les éléments de  $\Omega$  sont ordonnés tel que :  $\forall i = \{1..n\}$ ,  $p_i > 0$ ,  $p_i \geq p_{i+1}$  et  $\pi_i > 0$ ,  $\pi_i \geq \pi_{i+1}$  avec  $p_{n+1} = 0$  et  $\pi_{n+1} = 0$ . Klir a examiné le principe de préservation de l'incertain sous deux échelles :

- *Ratio-scale* :  $p \rightarrow \pi$  et  $\pi \rightarrow p$ , appelées transformations normalisées, elles sont définies par :

$$\pi_i = \frac{p_i}{p_1}, \quad p_i = \frac{\pi_i}{n \sum_{i=1}^n \pi_i} \quad (2.2)$$

- *Log-interval scale* :  $p \rightarrow \pi$  et  $\pi \rightarrow p$  sont définies par :

$$\pi_i = \left(\frac{p_i}{p_1}\right)^\alpha, \quad p_i = \frac{\frac{1}{\pi_i^\alpha}}{n \sum_{i=1}^n (\pi_i)^\alpha} \quad (2.3)$$

où  $\alpha$  est un paramètre appartenant à l'intervalle  $]0,1[$ .

### 2.3.2 Optimal transformation (OT)

Proposée par Dubois et Prade [DFMP04] et appelée "Asymetric transformation", cette transformation est définie comme suit :

$$\pi_i = \sum_{j/p_j \leq p_i} p_j, \quad p_i = \sum_{j=1}^n \frac{\pi_j - \pi_{j+1}}{j} \quad (2.4)$$

*OT* est optimale car elle donne la distribution de possibilités la plus spécifique, c'est à dire celle qui perd le moins d'informations [DPS93]. Elle est dite asymétrique puisque les deux formules ne sont pas inverses, elles ne sont pas basées sur le même principe informationnel. Alors que la transformation possibilités-probabilités est régie par le principe de "raison insuffisante" qui a pour but de trouver la distribution de probabilités qui contient autant d'incertain que possible mais qui conserve les caractéristiques de la distribution de possibilités. La transformation probabilités-possibilités est régie par le principe de spécificité maximale, qui cherche à trouver la distribution de possibilités la plus informative possible.

### 2.3.3 Symmetric transformation (ST)

Sandri et al. [DPS93] ont proposé *ST* qui nécessite moins de calculs mais qui est bien loin d'être optimale, car elle ne donne pas la meilleure spécificité possible. *ST* est définie par :

$$\pi_i = \sum_{j=1}^n \min(p_i, p_j) \quad (2.5)$$

### 2.3.4 Variable transformation (VT)

C'est une transformation probabilités-possibilités proposée par Mouchaweb et al. [MMR06] et s'exprime de la façon suivante :

Supposons que les éléments de  $\Omega$  sont ordonnés de telle sorte que :  $\forall i = \{1..n\}$ ,  $p_i > 0$ ,  $p_i \geq p_{i+1}$  avec  $p_{n+1} = 0$ , alors :

$$\pi_i = \left(\frac{p_i}{p_1}\right)^{k \cdot (1-p_i)} \quad (2.6)$$

$k$  est une constante qui doit satisfaire la condition de consistance suivante :

$$\forall \omega \in \Omega : \pi(\omega) \geq p(\omega) \quad (2.7)$$

Cette condition est un cas particulier du principe de consistance de Dubois et Prade défini dans la sous-section 2.2.3. En effet, la condition (2.7) est le cas discret, i.e la distribution ne contient que des singletons, de la condition de Dubois et Prade. Pour garantir le principe de consistance (2.7), la valeur de  $k$  doit appartenir à l'intervalle suivant :

$$0 \leq k \leq \frac{\log(p_n)}{(1-p_n) \cdot \log\left(\frac{p_n}{p_1}\right)} \quad (2.8)$$

Lorsque la valeur de  $k$  est égale à son maximum :  $k_{max} = \frac{\log(p_n)}{(1-p_n) \cdot \log\left(\frac{p_n}{p_1}\right)}$ , la possibilité  $\pi_n$ , calculée selon VT, est égale à  $p_n$ . Si la valeur de  $k$  augmente au dessus de  $k_{max}$  alors la valeur de  $\pi_n$  deviendra plus petite que celle de  $p_n$ , ce qui veut dire que VT ne satisfera plus la condition de consistance définie par 2.7. En effet, VT transforme une distribution de probabilités de manière non-linéaire ; ce qui signifie que VT ajoute plus d'informations aux grandes probabilités qu'aux petites. C'est dû au fait qu'une grande probabilité a une puissance,  $k \cdot (1 - p_i)$ , plus petite que les petites probabilités. La différence entre la transformation de Klir et VT est que la transformation de Klir a une puissance constante  $\alpha$  qui appartient à l'intervalle ]0,1[ afin de préserver l'incertitude, tandis que la puissance,  $k \cdot (1 - p_i)$ , dans VT est variable pour la rendre plus spécifique.

Bouguelid [Bou07] a proposé  $VT_i$ , qui est une amélioration de VT, afin de la rendre aussi spécifique que OT. Alors, un paramètre  $k_i$  est associé à chaque  $\pi_i$ . Formellement,  $\forall i = \{1..n\}$ ,

$$\pi_i = \left(\frac{p_i}{p_1}\right)^{k_i \cdot (1-p_i)} \quad (2.9)$$

où  $k_i$  appartient à l'intervalle :  $0 \leq k_i \leq \frac{\log(p_i + p_{i+1} + \dots + p_n)}{(1-p_n) \cdot \log\left(\frac{p_n}{p_1}\right)}$ ,  $\forall i = \{2..n\}$ .

Nous verrons par la suite, que toutes les transformations définies souffrent tout de même de nombreux problèmes. Voyons avant de passer à la suite, un exemple de ces transformations.

### 2.3.5 Exemple

**Exemple 12.** Considérons 2 variables  $A$  et  $B$  avec pour domaines  $D_A = \{a_1, a_2\}$  et  $D_B = \{b_1, b_2, b_3\}$ , la distribution de probabilités associée est présentée dans la Table 2.2.

Appliquons les transformations OT, KT, ST, VT, et  $VT_i$  sur la distribution de la Table 2.2. On obtient la Table 2.3.

A	B	$p(A, B)$
$a_1$	$b_1$	0.36
$a_1$	$b_2$	0.18
$a_1$	$b_3$	0.06
$a_2$	$b_1$	0.2
$a_2$	$b_2$	0.12
$a_2$	$b_3$	0.08

TABLE 2.2 – Distribution de probabilités de l’Exemple 12

A	B	$\Pi_{KT}$	$\Pi_{OT,VT_i}$	$\Pi_{ST}$	$\Pi_{VT}$
$a_1$	$b_1$	1	1	1	1
$a_1$	$b_2$	0.5	0.44	0.8	0.38
$a_1$	$b_3$	0.16	0.06	0.36	0.06
$a_2$	$b_1$	0.55	0.64	0.84	0.45
$a_2$	$b_2$	0.33	0.26	0.62	0.19
$a_2$	$b_3$	0.22	0.14	0.46	0.09

TABLE 2.3 – Distribution de possibilités obtenues par les différentes transformations

### 2.3.6 Exemples de nouvelles transformations

Il est facile de créer des transformations permettant de passer d’une distribution de probabilités vers une distribution de possibilités tout en respectant le principe de préservation d’ordre, nous en proposons deux assez simples. La première étant définie par :

$$\pi_i = \begin{cases} 1 & \text{si } p_i = \max_j(p_j) \\ p_i & \text{sinon} \end{cases} \quad (2.10)$$

Cette transformation respecte la préservation de l’ordre des interprétations et la normalisation. Cependant elle ne respecte pas la condition de consistance.

*Contre-exemple.* La Table 2.4 nous montre que l’événement  $\phi = \{\omega_2, \omega_3\}$  a une probabilité plus grande que sa possibilité.

$\phi$	$P(\phi)$	$\Pi(\phi)$
$\{\omega_1\}$	0.5	1
$\{\omega_2\}$	0.4	0.4
$\{\omega_3\}$	0.1	0.1
$\{\omega_1, \omega_2\}$	0.9	1
$\{\omega_1, \omega_3\}$	0.6	1
$\{\omega_2, \omega_3\}$	0.5	0.4
$\{\omega_1, \omega_2, \omega_3\}$	1	1

TABLE 2.4 – Probabilités et possibilités des événements sur  $2^\Omega$

Ainsi la condition  $\forall \phi \subseteq \Omega, P(\phi) \leq \Pi(\phi)$  n’est pas satisfaite. □

La deuxième transformation que nous proposons est définie par :

$$\pi_i = \begin{cases} p_i + (1 - \max_j(p_j)) & \text{si } p_i > 0 \\ 0 & \text{sinon} \end{cases} \quad (2.11)$$

Cette seconde transformation quant à elle, respecte deux des trois postulats définis par Dubois et Prade, et présentés plus haut. La condition de consistance et la préservation de l'ordre.

Nous concluons que construire une transformation qui préserve les principes est assez simple, la plus grande difficulté est plutôt de la rendre aussi spécifique que possible.

### 2.3.7 Autres transformations

Les transformations présentées précédemment sont les plus connues et les plus utilisées. D'autres transformations ont été proposées, dans d'autres contextes. Yager a en effet, proposé une transformation d'une distribution de probabilités en une autre distribution de probabilités telle que la distribution de probabilités résultante ait au moins autant d'entropie que la distribution de probabilités originale. Pour développer cette procédure, il a utilisé l'approche de Dubois et Prade lors de leur transformation possibilités-probabilités.

D'autres transformations ont été implémentées dans le cas continu plutôt que dans le cas fini. Dubois et Prade [DPS93] proposent une approche pour la transformation d'une distribution de probabilités continue en une distribution de possibilités spécifique qui permet aux limites supérieures de probabilité d'événements d'être calculées.

Yamada [Yam01], quant à lui, a proposé une transformation basée sur la théorie de l'évidence. Il examine trois nouvelles idées basées sur la théorie de l'évidence : une considère les possibilités sur une échelle ordinale, et les deux autres sur une échelle numérique. Et en ce qui concerne l'échelle ordinale, son travail montre que le principe de préservation d'ordre donne une unique structure ordinale des possibilités. En ce qui concerne l'échelle numérique, il examine deux méthodes et montre que l'une des deux basée sur l'équidistribution mène à une unique transformation satisfaisant les principes de consistance et de préservation d'ordre, et il se trouve que cette transformation est celle proposée dans un papier de Dubois et Prade [DP83].

## Conclusions

La Table 2.5 résume les caractéristiques de  $KT$ ,  $OT$ ,  $ST$ ,  $VT$  et  $VT_i$ . Pour chaque transformation, il est précisé si elle traite le cas discret (D) et le cas continu (C) et si elle respecte le principe de consistance (Cs), la préservation de la préférence (PP) et le maximum de spécificité (MS). Clairement,  $OT$  et  $VT_i$  sont les règles les plus intéressantes dans le cas discret pour la transformation  $p \rightarrow \pi$ .

TR	$p \rightarrow \pi$	$\pi \rightarrow p$	Propriétés				
			D	C	Cs	PP	MS
$KT$	×	×	×	×		×	
$OT$	×	×	×	×	×	×	×
$ST$	×	×	×	×	×	×	
$VT$	×		×			×	
$VT_i$	×		×		×	×	×

TABLE 2.5 – Résumé des propriétés des transformations

La représentation des probabilités et la représentation des possibilités ne sont pas deux représentations équivalentes de l'incertain. Par conséquent, il ne devrait pas y avoir de symétrie entre les deux transformations. La représentation possibiliste est plus faible car elle est conçue pour traiter des informations incomplètes (ignorance totale ou partielle). Et aussi car les mesures de possibilités sont basées sur une structure maxitive plutôt qu'une additive.

Le but des chapitres suivants est d'analyser les transformations connues et vues dans ce chapitre d'abord à travers les distributions, puis dans les modèles graphiques. En plus, nous étudierons les transformations en générale sur des questions comme le conditionnement, les indépendances, les requêtes type MPE, MAP, etc.



# Chapitre 3

## Analyse des transformations

### Sommaire

---

<b>3.1 Transformations et normalisation</b> . . . . .	<b>23</b>
3.1.1 Transformation $OT$ . . . . .	24
3.1.2 Transformation $KT$ . . . . .	24
3.1.3 Transformation $ST$ . . . . .	24
3.1.4 Transformation $VT$ . . . . .	25
<b>3.2 Transformations et requêtes MPE</b> . . . . .	<b>25</b>
<b>3.3 Préservation de l'ordre des événements</b> . . . . .	<b>26</b>
3.3.1 $OT$ , une transformation qui ne préserve pas l'ordre des événements . . . . .	26
3.3.2 Généralisation à toutes les transformations . . . . .	27
<b>3.4 Transformations et marginalisation</b> . . . . .	<b>28</b>
<b>3.5 Transformations et conditionnement</b> . . . . .	<b>30</b>
3.5.1 Cadre quantitatif . . . . .	30
3.5.2 Cadre qualitatif . . . . .	31
<b>3.6 Conditionnement avec des informations incertaines</b> . . . . .	<b>32</b>
3.6.1 Conditionnement de Jeffrey en probabilités . . . . .	33
3.6.2 Conditionnement de Jeffrey en possibilités . . . . .	33
3.6.3 Transformations et conditionnement de Jeffrey . . . . .	34
<b>3.7 Transformations et requêtes MAP</b> . . . . .	<b>38</b>

---

Nous venons de présenter l'ensemble des cadres de représentations de l'incertain utilisé dans ce mémoire, nous avons aussi défini quelques transformations existantes et couramment utilisées. Le but de ce chapitre est d'étudier ces transformations à travers plusieurs questions. Nous nous intéressons en effet à la normalisation, les requêtes MPE, la marginalisation, le conditionnement, les requêtes MAP, etc.

### 3.1 Transformations et normalisation

Dans les principes énoncés dans le chapitre précédent, il n'est pas explicitement dit qu'une transformation doit satisfaire la normalisation. Pourtant, il est important que chaque transformation donne une distribution de possibilités normalisée, le but de cette section est de vérifier que la normalisation est bien préservée lors de la transformation.

### 3.1.1 Transformation $OT$

$OT$  est considérée comme la transformation la plus optimale, satisfaisant tous les principes définis dans la Section 2.2. Il en est de même pour la normalisation, en effet à partir de n'importe quelles distributions de probabilités normalisées,  $OT$  transforme en distributions de possibilités normalisées.

**Proposition 3.1.** Soit  $p$  une distribution de probabilités sur  $\Omega$ . Soit  $\pi_{OT} = OT(p)$  la transformation de  $p$  avec  $OT$ . Si  $p$  est normalisée, alors  $\pi_{OT}$  est normalisée.

*Démonstration.* Soit  $p$  une distribution de probabilités normalisée, on a :

$$\sum_{\omega_i \in \Omega} p(\omega_i) = 1$$

Soit  $\pi_{OT}$  la transformation définie par :

$$\pi_{OT}(\omega_i) = \sum_{\omega_j / p(\omega_j) \leq p(\omega_i)} p(\omega_j)$$

Considérons  $\omega_{max}$  le monde ayant la plus grande probabilité selon  $p$  (i.e  $\forall \omega_i \in \Omega, p(\omega_i) \leq p(\omega_{max})$ ). Ainsi,

$$\pi_{OT}(\omega_{max}) = \sum_{\omega_j / p(\omega_j) \leq p(\omega_{max})} p(\omega_j) = \sum_{\omega_i \in \Omega} p(\omega_i) = 1$$

La distribution de possibilités  $\pi_{OT}$  est bien normalisée. □

### 3.1.2 Transformation $KT$

**Proposition 3.2.** Soit  $p$  une distribution de probabilités sur  $\Omega$ . Soit  $\pi_{KT} = KT(p)$  la transformation de  $p$  avec  $KT$ . Si  $p$  est normalisée, alors  $\pi_{KT}$  est normalisée.

*Démonstration.* Soit  $p$  une distribution de probabilités sur  $\Omega$  ordonnée tel que  $p_i > 0, p_i \geq p_{i+1}$  avec  $p_{n+1} = 0$ .

$$\pi_1 = \frac{p_1}{p_1} = 1$$

La distribution transformée avec  $KT$  est bien normalisée. □

### 3.1.3 Transformation $ST$

**Proposition 3.3.** Soit  $p$  une distribution de probabilités sur  $\Omega$ . Soit  $\pi_{ST} = ST(p)$  la transformation de  $p$  avec  $ST$ . Si  $p$  est normalisée, alors  $\pi_{ST}$  est normalisée.

*Démonstration.* Rappelons que la transformation  $ST$  est définie par :  $\pi_i = \sum_{j=1}^n \min(p_i, p_j)$ . Notons la plus grande probabilité  $p_{max}$ , nous obtenons :

$$\pi_{max} = \sum_{j=1}^n \min(p_{max}, p_j) = \sum_{j=1}^n p_j = 1$$

La distribution  $\pi_{ST}$  est donc normalisée. □

### 3.1.4 Transformation VT

**Proposition 3.4.** Soit  $p$  une distribution de probabilités sur  $\Omega$ . Soit  $\pi_{VT} = VT(p)$  la transformation de  $p$  avec VT. Si  $p$  est normalisée, alors  $\pi_{VT}$  est normalisée.

*Démonstration.* Supposons que les éléments de  $\Omega$  sont ordonnés de telle sorte que :  $\forall i = \{1..n\}$ ,  $p_i > 0$ ,  $p_i \geq p_{i+1}$  avec  $p_{n+1} = 0$ , et :

$$\pi_i = \left(\frac{p_i}{p_1}\right)^{k \cdot (1-p_i)}$$

où  $k$  est une constante appartenant à l'intervalle :  $0 \leq k \leq \frac{\log(p_n)}{(1-p_n) \cdot \log(\frac{p_n}{p_1})}$ .

La transformation nous donne bien une distribution de possibilités normalisée, en effet

$$\pi_1 = \left(\frac{p_1}{p_1}\right)^{k \cdot (1-p_1)} = 1^{k \cdot (1-p_1)} = 1$$

□

## 3.2 Transformations et requêtes MPE

La question que nous nous posons ici est la suivante : *Étant donnée une requête de type MPE, avons nous la même réponse si nous raisonnons dans un cadre probabiliste ou dans un cadre possibiliste.*

**Proposition 3.5.** Soit  $p$  une distribution de probabilités et  $\pi_{TR}$  une transformation préservant l'ordre des interprétations alors

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega} (p(\omega_i, \phi)) = \operatorname{argmax}_{\omega_i \in \Omega} (\pi(\omega_i, \phi)) \quad (3.1)$$

*Démonstration.* La démonstration est assez triviale. Considérons l'ordre sur  $\Omega = \{\omega_1, \dots, \omega_n\}$  suivant :  $p(\omega_1) > p(\omega_2) > \dots > p(\omega_n)$ . Si la transformation  $TR$  préserve l'ordre des interprétations, alors :  $\pi(\omega_1) > \pi(\omega_2) > \dots > \pi(\omega_n)$ .

Calculer  $\operatorname{argmax}_{\omega_i \in \Omega} (p(\omega_i, \phi))$ , c'est calculer l'interprétation ayant la probabilité maximale parmi les interprétations appartenant à l'événement  $\phi$ . Soit  $\omega_j = \operatorname{argmax}_{\omega_i \in \Omega} (p(\omega_i, \phi))$  alors

$$\forall \omega_i \in \phi, \text{ tel que } \omega_j \neq \omega_i, p(\omega_j) > p(\omega_i)$$

Comme l'ordre des interprétations est préservé alors :

$$\forall \omega_i \in \phi, \text{ tel que } \omega_j \neq \omega_i, \pi_{TR}(\omega_j) > \pi_{TR}(\omega_i)$$

Ainsi,  $\omega_j = \operatorname{argmax}_{\omega_i \in \Omega} (\pi_{TR}(\omega_i, \phi))$ . La réponse à la requête MPE est bien la même. □

Comme les transformations  $OT$ ,  $KT$ ,  $ST$  et  $VT$  préservent l'ordre des interprétations, nous en déduisons le Corollaire 3.1.

**Corollaire 3.1.** Soit  $p$  une distribution de probabilités et  $\pi_{TR}$  la distribution de possibilités résultant de la transformation  $TR \in \{OT, KT, ST, VT\}$ .

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega} (p(\omega_i, \phi)) = \operatorname{argmax}_{\omega_i \in \Omega} (\pi_{OT}(\omega_i, \phi)) \quad (3.2)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega} (p(\omega_i, \phi)) = \operatorname{argmax}_{\omega_i \in \Omega} (\pi_{KT}(\omega_i, \phi)) \quad (3.3)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(p(\omega_i, \phi)) = \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{ST}(\omega_i, \phi)) \quad (3.4)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(p(\omega_i, \phi)) = \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{VT}(\omega_i, \phi)) \quad (3.5)$$

Comme conséquence directe, si dans un problème donné, nous nous intéressons uniquement aux requêtes MPE, alors transformer en possibiliste et raisonner dans ce cadre garantit les mêmes résultats.

### 3.3 Préservation de l'ordre des événements

Les transformations probabilistes-possibilistes essaient de préserver l'ordre des interprétations ( $\omega_i \in \Omega$ ), mais qu'en est-il des événements ( $\phi \subseteq \Omega$ ) ?

#### 3.3.1 OT, une transformation qui ne préserve pas l'ordre des événements

Nous avons pu constater et il est facilement démontrable que l'ordre des événements après n'importe quelle transformation existante n'est pas forcément préservé. Illustrons cela avec l'exemple suivant :

**Exemple 13.** Prenons une distribution de probabilités sur 3 mondes  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ . La distribution et la transformation par OT sont représentées sur la Figure 3.1.

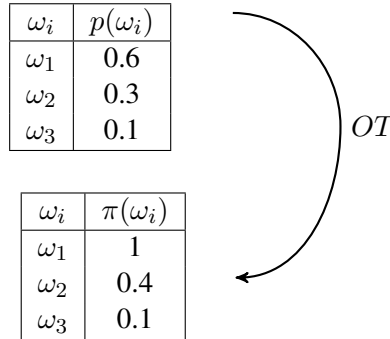


FIGURE 3.1 – Transformation d'une distribution de probabilités en distribution de possibilités par OT

Nous pouvons écrire la Table 3.1, décrivant la probabilité et possibilité de l'événement  $\phi \subseteq \Omega$ .

$\phi$	$P(\phi)$	$\Pi(\phi)$
$\{\omega_1\}$	0.6	1
$\{\omega_2\}$	0.3	0.4
$\{\omega_3\}$	0.1	0.1
$\{\omega_1, \omega_2\}$	0.9	1
$\{\omega_1, \omega_3\}$	0.7	1
$\{\omega_2, \omega_3\}$	0.4	0.4
$\{\omega_1, \omega_2, \omega_3\}$	1	1

TABLE 3.1 – Probabilités et possibilités des événements sur  $2^\Omega$

Nous remarquons bien que l'ordre sur les événements n'est pas préservé puisque par exemple :

$$P(\{\omega_1, \omega_2\}) > P(\{\omega_1, \omega_3\}) \text{ alors que } \Pi(\{\omega_1, \omega_2\}) = \Pi(\{\omega_1, \omega_3\})$$

L'Exemple 13 nous montre que l'ordre des événements n'est pas préservé pour la transformation  $OT$ . L'étude faite pour les autres transformations  $KT$ ,  $ST$  et  $VT$  conclut au même résultat.

### 3.3.2 Généralisation à toutes les transformations

La question qui se pose donc est de savoir si nous pouvons construire une nouvelle transformation, respectant les principes de consistance et préservation d'ordre et permettant la préservation de l'ordre des événements. La Proposition 3.6 montre qu'il est impossible d'obtenir telle transformation.

**Proposition 3.6.** *Soit  $TR$  une transformation probabilités-possibilités,  $p$  une distribution de probabilités et  $\pi_{TR}$  la distribution de possibilités obtenue par la transformation  $TR$ .*

$$\forall \phi, \psi \subseteq \Omega, \text{ tel que } \phi \neq \psi, P(\phi) < P(\psi) \not\Rightarrow \Pi_{TR}(\phi) < \Pi_{TR}(\psi) \quad (3.6)$$

*Démonstration.* Soit  $p$  la distribution de la Table 3.2 sur  $\Omega = \{\omega_1, \omega_2, \omega_3\}$  avec  $\begin{cases} \alpha > \beta > \gamma \\ \alpha + \beta + \gamma = 1 \end{cases}$ .

$\omega_i$	$p(\omega_i)$
$\omega_1$	$\alpha$
$\omega_2$	$\beta$
$\omega_3$	$\gamma$

TABLE 3.2 – Distribution de probabilités représentant la perte de l'ordre des événements

Si  $TR$  préserve l'ordre des interprétations, alors  $\pi(\omega_1) > \pi(\omega_2) > \pi(\omega_3)$  et  $\pi$  satisfait la normalisation  $\pi(\omega_1) = 1$ .

Nous avons  $P(\{\omega_1, \omega_2\}) = \alpha + \beta > P(\{\omega_1\})$  mais  $\Pi(\{\omega_1, \omega_2\}) = 1$  et  $\Pi(\{\omega_1\}) = 1$ . Alors que  $P(\{\omega_1, \omega_2\}) > P(\{\omega_1\})$ , en possibilités,  $\Pi(\{\omega_1, \omega_2\}) = \Pi(\{\omega_1\})$ . L'ordre des événements n'est donc pas préservé.  $\square$

Cette démonstration nous montre juste que si nous avons un ordre strict sur deux interprétations en probabilités, nous pouvons avoir l'égalité de ces deux interprétations en possibilités. Mais est-ce vraiment le cas ou pouvons-nous avoir  $\forall \phi, \psi \subseteq \Omega$ , avec  $\phi \neq \psi$ ,  $\Pi(\phi) < \Pi(\psi) \Rightarrow P(\phi) < P(\psi)$ ? Ou alors avoir si  $P(\phi) < P(\psi)$  alors  $\Pi(\phi) \leq \Pi(\psi)$ ? La réponse est que nous ne pouvons pas affirmer que ces deux cas sont possibles. En effet, nous avons dans certaines situations  $P(\phi) < P(\psi)$  alors que  $\Pi(\phi) > \Pi(\psi)$ .

**Proposition 3.7.** *Soit  $TR$  une transformation probabilités-possibilités,  $p$  une distribution de probabilités telle que  $p(\omega_{max}) < \sum_{\omega_i \neq \omega_{max}} p(\omega_i)$ . Soit  $\pi_{TR}$  la distribution de possibilités obtenue par la transformation  $TR$  et  $\phi = \Omega \setminus \{\omega_{max}\}$ .*

$$P(\phi) > P(\{\omega_{max}\}) \Rightarrow \Pi_{TR}(\phi) < \Pi_{TR}(\{\omega_{max}\}) \quad (3.7)$$

$\omega_i$	$p(\omega_i)$	$\pi_{TR}(\omega_i)$
$\omega_1$	0.4	1
$\omega_2$	0.35	$\alpha$
$\omega_3$	0.25	$\beta$

TABLE 3.3 – Distribution de probabilités où  $p(\omega_1) < p(\omega_2) + p(\omega_3)$  et sa transformée

*Démonstration.* Soit  $p$  une distribution de probabilités sur  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ . On suppose que cette distribution est telle que  $\begin{cases} p(\omega_1) > p(\omega_2) > p(\omega_3) \\ p(\omega_1) < p(\omega_2) + p(\omega_3) \end{cases}$ . La Table 3.3 représente  $p$  et sa transformée  $\pi_{TR}$ .

La transformation  $TR$  transforme  $p$  en  $\pi_{TR}$  en préservant l'ordre des interprétations ainsi  $1 > \alpha > \beta$ . Calculons la probabilité et la possibilité des événements  $\phi = \{\omega_2, \omega_3\}$  et  $\psi = \{\omega_1\}$ , nous avons

- $P(\phi) = 0.6$  et  $P(\psi) = 0.4$ ,
- $\Pi(\phi) = \max(\beta, \gamma) = \beta$  et  $\Pi(\psi) = \alpha$

donc  $\Pi(\phi) < \Pi(\psi)$  alors que  $P(\phi) > P(\psi)$ . □

En résumé, nous distinguons deux cas, dans le premier cas où  $\omega_{max} > \sum_{\omega_i \in \Omega, \omega_i \neq \omega_{max}} p(\omega_i)$  (ie.  $\omega_{max} > 0.5$ ), nous relâchons l'égalité c'est à dire qu'au lieu d'avoir strictement inférieur, nous pouvons avoir juste inférieur. Dans le second cas, où  $\omega_{max} < 0.5$  alors il existe au moins un événement  $\phi$  tel que  $P(\phi) > P(\{\omega_{max}\})$  mais  $\Pi(\phi) < \Pi(\{\omega_{max}\})$ .

Ce qu'il convient de vérifier ensuite est la marginalisation. Dans les cadres probabiliste et possibiliste, la marginalisation se fait différemment. Alors qu'en probabilités, elle est basée sur le principe d'additivité, en possibilités elle l'est sur le principe de maxitivité.

### 3.4 Transformations et marginalisation

La question se pose alors de savoir quelles répercussions sur le raisonnement cette différence a lorsque nous transformons du cadre probabiliste vers le cadre possibiliste. Dans cette section, nous nous intéressons à l'effet des transformations sur la marginalisation.

Soient  $V = \{A_1, \dots, A_n\}$  et  $\{A_i, \dots, A_j\} \subseteq V$ . La Figure 3.2 illustre cette question.

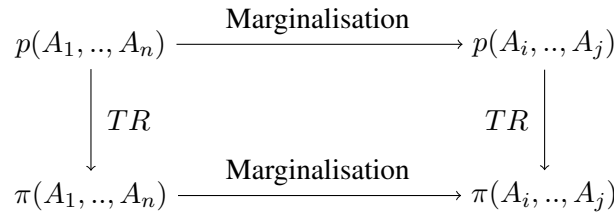


FIGURE 3.2 – Préservation de la marginalisation

Voyons maintenant un exemple avec la transformation  $OT$ .

**Exemple 14.** Soient  $p$  une distribution de probabilités sur trois variables binaires  $A, B$  et  $C$ . La Table 3.4 représente la distribution de probabilités et sa transformation par  $OT$ , et la Table 3.5 représente la marginalisation de la Table 3.4 sur la variable  $C$ .

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{OT}(A, B, C)$
$a$	$b$	$c$	0.09	0.292
$a$	$b$	$\neg c$	0.09	0.292
$a$	$\neg b$	$c$	0.21	0.712
$a$	$\neg b$	$\neg c$	0.21	0.712
$\neg a$	$b$	$c$	0.288	1
$\neg a$	$b$	$\neg c$	0.032	0.04
$\neg a$	$\neg b$	$c$	0.072	0.112
$\neg a$	$\neg b$	$\neg c$	0.008	0.008

 TABLE 3.4 – Distribution de probabilités et sa transformation par  $OT$ 

$A$	$B$	$p(A, B)$	$\pi_{OT}(A, B)$
$a$	$b$	0.18	0.292
$a$	$\neg b$	0.42	0.712
$\neg a$	$b$	0.32	1
$\neg a$	$\neg b$	0.08	0.112

 TABLE 3.5 – Marginalisation de la Table 3.4 sur la variable  $C$ 

$A$	$B$	$\pi_{OT}(A, B)$
$a$	$b$	0.26
$a$	$\neg b$	1
$\neg a$	$b$	0.58
$\neg a$	$\neg b$	0.08

 TABLE 3.6 – Transformation de la distribution  $p(A, B)$  de la Table 3.5 par  $OT$ 

La transformation par  $OT$  de la distribution  $p(A, B)$  de la Table 3.5 est décrite par la Table 3.6. Nous constatons que l'ordre des interprétations n'est pas préservé. En fait, marginaliser sur  $C$ , revient à calculer la probabilité des événements :  $\{abc, ab\neg c\}$ ,  $\{a\neg bc, a\neg b\neg c\}$ ,  $\{\neg abc, \neg ab\neg c\}$ ,  $\{\neg a\neg bc, \neg a\neg b\neg c\}$ . Or nous avons vu dans la Proposition 3.6, que l'ordre des événements n'était pas préservé, donc nous pouvons conclure que la marginalisation ne va pas préserver l'ordre des interprétations. Il en est de même pour les transformations  $KT$ ,  $ST$  et  $VT$ . Ajoutons aussi que l'ordre peut même s'inverser comme dans notre exemple.

**Proposition 3.8.** Soit  $TR$  une transformation probabilités-possibilités,  $p$  une distribution de probabilités sur l'ensemble de variables  $\{A_1, \dots, A_n\}$ . Soit  $A' \subseteq \{A_1, \dots, A_n\}$  et  $D_{A'}$  le domaine associé à  $A'$ . Soit  $\pi_{TR-mg}$  la distribution obtenue avec  $TR$  puis par marginalisation possibiliste sur  $A'$  (avec l'opérateur  $\max$ ). Soit  $\pi_{mg-TR}$  la distribution obtenue par marginalisation probabiliste puis par  $TR$  vers le possibiliste.

$$\forall \phi, \psi \subseteq D_{A'}, \Pi_{mg-TR}(\phi) < \Pi_{mg-TR}(\psi) \not\Rightarrow \Pi_{TR-mg}(\phi) < \Pi_{TR-mg}(\psi) \quad (3.8)$$

*Démonstration.* Soient deux variables  $A$  avec  $D_A = \{a, \neg a\}$  et  $B$  avec  $D_B = \{b, \neg b\}$  et la distribution de probabilités représentée par la Table 3.7.

Nous pouvons marginaliser sur  $A$  ou  $B$ , ainsi  $p(a) = 0.5$  et  $p(\neg a) = 0.5$  et pour  $B$ ,  $p(b) = 0.7$  et  $p(\neg b) = 0.3$ . En transformant les marginales, nous obtenons  $\pi(a) = 1$  et  $\pi(\neg a) = 1$  et pour  $B$ ,  $\pi(b) = 1$

$A$	$B$	$p(A, B)$
$a$	$b$	0.4
$a$	$\neg b$	0.1
$\neg a$	$b$	0.3
$\neg a$	$\neg b$	0.2

TABLE 3.7 – Distribution de probabilités pour la preuve de la Proposition 4.6

et  $\pi(\neg b) = y$ . Notons  $y$  l'inconnue qui change en fonction des transformations.

Maintenant, considérons la construction des distributions de possibilités marginales à partir de la transformation de la Table 3.7, que l'on trouve dans la Table 3.8. Les  $x_i$  sont tous différents et inférieurs à 1. En

$A$	$B$	$\pi_{TR}(A, B)$
$a$	$b$	1
$a$	$\neg b$	$x_1$
$\neg a$	$b$	$x_2$
$\neg a$	$\neg b$	$x_3$

TABLE 3.8 – Distribution de possibilités transformée de la Table 3.7

marginalisant sur  $B$ ,  $\pi(b) = 1$  et  $\pi(\neg b) = \max(x_1, x_3)$ . En marginalisant sur  $A$ ,  $\pi(a) = 1$  et  $\pi(\neg a) = \max(x_2, x_3)$ . Par conséquent, en transformant les marginales, on a  $\pi_{mg-TR}(a) = \pi_{mg-TR}(\neg a) = 1$  alors qu'en marginalisant la transformée, on a  $\pi_{TR-mg}(a) > \pi_{TR-mg}(\neg a)$ . Donc aucune transformation  $TR$  ne peut préserver la marginalisation si elle préserve l'ordre des interprétations. □

Lorsque nous sommes face à une évidence nouvelle (information sûre ou observation), nous avons besoin de la prendre en compte. Le conditionnement nous permet de répondre à cette question. Qu'en est-il lorsque que nous transformons les distributions ? Les deux prochaines sections répondrons à cette question, d'abord quand l'évidence traitée est certaine (ie. la probabilité de l'événement est de 1) puis, quand un événement est incertain, alors nous conditionnerons sur une nouvelle distribution avec le conditionnement de Jeffrey.

### 3.5 Transformations et conditionnement

Le conditionnement et les requêtes MAP peuvent être vus sur le même plan (mis à part que les requêtes MAP utilisent la marginalisation en plus). Ici, nous essayons de montrer que les transformations dès le moment qu'elles vérifient les principes de Dubois et Prade (Section 2.2), alors elles répondent à la question suivante : *Étant donné une observation, trouvons nous les mêmes informations après conditionnement en probabilités et en possibilités ?* Illustrons cette question par le Schéma 3.3, et essayons de répondre à cette question dans les deux cadres possibilistes étudiés, dans le cadre quantitatif et le cadre qualitatif. Pour information, dans un de ses papiers [Sud92], Sudkamp a montré que nous n'obtenons pas la même distribution. Cependant, nous nous intéressons ici à l'ordre des interprétations.

#### 3.5.1 Cadre quantitatif

Le conditionnement en probabilités et le conditionnement en possibilités dans le cadre quantitatif se ressemblent dans le sens où ils sont tous les deux basés sur le rapport pour la renormalisation. Rappelons que le conditionnement (dans tous les cadres) avec une information certaine  $\phi \subseteq \Omega$ , exclut les



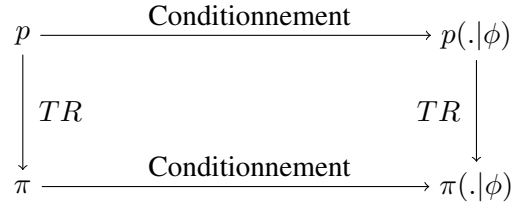


FIGURE 3.3 – Préservation du conditionnement

interprétations  $\omega_i \notin \phi$  et renormalise les interprétations  $\omega_j \in \phi$  pour obtenir une distribution a posteriori normalisée. Comme les transformations préservent aussi l'ordre des interprétations, nous pouvons supposer que l'ordre dans le conditionnement des transformations et des distributions de probabilités sera le même.

**Proposition 3.9.** *Soit  $\phi \subseteq \Omega$  une observation. Soit  $\pi_{TR-cd}$  la distribution obtenue avec  $TR$  puis par conditionnement possibiliste quantitatif avec  $\phi$ . Soit  $\pi_{cd-TR}$  la distribution obtenue par conditionnement probabiliste puis par  $TR$  vers le possibiliste.*

$$\forall \omega_i, \omega_j \in \Omega, \pi_{TR-cd}(\omega_i | \phi) < \pi_{TR-cd}(\omega_j | \phi) \Rightarrow \pi_{cd-TR}(\omega_i | p \phi) < \pi_{cd-TR}(\omega_j | p \phi) \quad (3.9)$$

L'interprétation de cette proposition est que si nous transformons puis conditionnons, l'ordre des interprétations  $\omega_i \in \Omega$  sera le même que si nous conditionnons d'abord et transformons après.

*Démonstration.* Soit l'observation  $\phi \subseteq \Omega$ . Soient  $\omega_i, \omega_j \in \Omega^2$  et  $p(\omega_i) > p(\omega_j)$ . Si  $\omega_i \in \phi$  et  $\omega_j \in \Omega$ , alors

$$\frac{p(\omega_i)}{P(\phi)} > \frac{p(\omega_j)}{P(\phi)} \Rightarrow p(\omega_i | \phi) > p(\omega_j | \phi)$$

Puisque  $TR$  est une transformation qui préserve l'ordre, nous pouvons affirmer que

$$\pi_{cd-TR}(\omega_i | \phi) > \pi_{cd-TR}(\omega_j | \phi)$$

De l'autre côté, si  $TR$  est une transformation qui préserve l'ordre alors

$$p(\omega_i) > p(\omega_j) \Rightarrow \pi_{TR}(\omega_i) > \pi_{TR}(\omega_j)$$

En conditionnant quantitativement par  $\phi$ , nous avons

$$\pi_{TR}(\omega_i) > \pi_{TR}(\omega_j) \Rightarrow \frac{\pi(\omega_i)}{\Pi(\phi)} > \frac{\pi(\omega_j)}{\Pi(\phi)} \Rightarrow \pi_{TR-cd}(\omega_i | \phi) > \pi_{TR-cd}(\omega_j | \phi)$$

Les autres cas,  $\omega_i \in \phi, \omega_j \notin \phi$  et  $\omega_i \notin \phi, \omega_j \notin \phi$  sont triviaux. □

### 3.5.2 Cadre qualitatif

Après avoir vérifié dans le cadre quantitatif, nous nous intéressons au cadre qualitatif. Le conditionnement dans ce cadre est différent du cadre quantitatif, et est pour rappel défini comme suit :

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

Le conditionnement qualitatif peut se voir comme le calcul suivant :  $\min(\pi(\omega_i), \Pi(\phi))$  quand l'interprétation  $\omega_i$  appartient à l'événement  $\phi$ . Ce conditionnement garantit aussi que l'ordre est préservé. Ainsi, la Proposition 3.10 montre que comme dans le cadre quantitatif, l'ordre des interprétations est conservé.

**Proposition 3.10.** Soit  $\phi \subseteq \Omega$  une observation. Soit  $\pi_{TR-cd}$  la distribution obtenue avec  $TR$  puis par conditionnement possibiliste qualitatif avec  $\phi$ . Soit  $\pi_{cd-TR}$  la distribution obtenue par conditionnement probabiliste puis par  $TR$  vers le possibiliste.

$$\forall \omega_i, \omega_j \in \Omega, \pi_{TR-cd}(\omega_i | \phi) < \pi_{TR-cd}(\omega_j | \phi) \Rightarrow \pi_{cd-TR}(\omega_i | \phi) < \pi_{cd-TR}(\omega_j | \phi) \quad (3.10)$$

*Démonstration.* Soit l'observation  $\phi \subseteq \Omega$ . Comme pour le cadre quantitatif, nous éliminons toutes les interprétations qui n'appartiennent pas à l'événement  $\phi$ . Soient  $\omega_i, \omega_j \in \phi$  avec  $p(\omega_i) > p(\omega_j)$  alors

$$\frac{p(\omega_i)}{P(\phi)} > \frac{p(\omega_j)}{P(\phi)} \Rightarrow p(\omega_i | \phi) > p(\omega_j | \phi)$$

Puisque  $TR$  est une transformation qui préserve l'ordre, nous affirmons que

$$\pi_{cd-TR}(\omega_i | \phi) > \pi_{cd-TR}(\omega_j | \phi)$$

De l'autre côté, si  $TR$  est une transformation qui préserve l'ordre alors

$$p(\omega_i) > p(\omega_j) \Rightarrow \pi_{TR}(\omega_i) > \pi_{TR}(\omega_j)$$

En conditionnant qualitativement par  $\phi$ ,

$$\pi_{TR}(\omega_i) > \pi_{TR}(\omega_j) \Rightarrow \min(\pi(\omega_i), \Pi(\phi)) > \min(\pi(\omega_j), \Pi(\phi)) \Rightarrow \pi_{TR-cd}(\omega_i | \phi) > \pi_{TR-cd}(\omega_j | \phi)$$

Nous précisons que si  $\omega_i \in \phi$  alors  $\pi(\omega_i) \leq \Pi(\phi)$ .  $\square$

Comme le conditionnement dans le cadre quantitatif et dans le cadre qualitatif préservent l'ordre des interprétations, et comme les transformations vues dans l'état de l'art respectent le principe de Dubois et Prade, nous pouvons écrire le corollaire suivant valable dans les deux cadres :

**Corollaire 3.2.** Soit  $\pi_{TR-cd}$  la distribution obtenue avec  $TR$  puis par conditionnement possibiliste de  $\phi$ . Soit  $\pi_{cd-TR}$  la distribution obtenue par conditionnement probabiliste puis par  $TR$  vers le possibiliste. La transformation  $TR \in \{OT, KT, ST, VT\}$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{OT-cd}(\omega_i | \phi)) = \operatorname{argmax}_{\omega_i}(\pi_{cd-OT}(\omega_i | \phi)) \quad (3.11)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{KT-cd}(\omega_i | \phi)) = \operatorname{argmax}_{\omega_i}(\pi_{cd-KT}(\omega_i | \phi)) \quad (3.12)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{ST-cd}(\omega_i | \phi)) = \operatorname{argmax}_{\omega_i}(\pi_{cd-ST}(\omega_i | \phi)) \quad (3.13)$$

$$\forall \phi \subseteq \Omega, \operatorname{argmax}_{\omega_i \in \Omega}(\pi_{VT-cd}(\omega_i | \phi)) = \operatorname{argmax}_{\omega_i}(\pi_{cd-VT}(\omega_i | \phi)) \quad (3.14)$$

### 3.6 Conditionnement avec des informations incertaines

Conditionner avec des informations incertaines diffère du conditionnement vu précédemment. C'est prendre en compte l'incertitude de la nouvelle information. Par exemple, un médecin reçoit un patient, son premier symptôme est une toux, la probabilité de deux affections vient alors dans l'esprit du médecin. Alors que le patient énonce un deuxième symptôme, la probabilité d'une des affections se voit augmenter, il doit alors revoir son jugement. Dans cette section, chaque transformation étudiée qui respecte le principe de consistance de Dubois et Prade, préserve l'ordre des interprétations.

### 3.6.1 Conditionnement de Jeffrey en probabilités

La règle de conditionnement de Jeffrey [Jef65] permet de réviser une distribution de probabilités initiale  $p$  en  $p'$  étant donnée une nouvelle distribution représentant la nouvelle information incertaine. Cette méthode implique :

1. **Un moyen de spécifier l'évidence incertaine** : L'incertitude est de la forme  $(\phi_i, \alpha_i)$  avec  $\alpha_i = P'(\phi_i)$  qui signifie qu'après la opération de révision, la probabilité a posteriori de chaque événement  $\phi_i$  doit être égale à  $\alpha_i$ . Ici, la primauté est donnée à la nouvelle information.
2. **Une méthode pour calculer la distribution de probabilités révisée** : La méthode de Jeffrey suppose que bien qu'il y ait un désaccord à propos des événements  $\phi_i$  dans l'ancienne et la nouvelle distribution, le conditionnement probabiliste de n'importe quel événement  $\psi \subseteq \Omega$  étant donné un événement incertain  $\phi_i$  reste le même dans la distribution avant et après la révision. A savoir,

$$\forall \phi_i \subseteq \Omega, \forall \psi \subseteq \Omega, P(\psi | \phi_i) = P'(\psi | \phi_i) \quad (3.15)$$

L'interprétation sous-jacente de la révision impliquée par l'Équation 3.15 est que la distribution de probabilités révisée  $p'$  ne doit pas changer le degré de probabilité conditionnelle de n'importe quel événement  $\psi$  étant donné un événement incertain  $\phi_i$ .

Dans le cadre probabiliste, appliquer la règle de Bayes puis la marginalisation nous permet de réviser le degré de probabilité de n'importe quel événement  $\psi \subseteq \Omega$  de la manière suivante :

$$P'(\psi) = \sum_{\phi_i} \alpha_i * \frac{P(\psi, \phi_i)}{P(\phi_i)} \quad (3.16)$$

*Remarque.* La distribution de probabilités révisée  $p'$  obtenue en utilisant la règle de Jeffrey est unique et existe toujours [CD05]. Nous retrouvons la même chose pour la règle de Jeffrey dans le cadre possibiliste.

### 3.6.2 Conditionnement de Jeffrey en possibilités

En théorie des possibilités, la connaissance est encodée par une distribution de possibilités. Étant donné  $\pi$  la distribution de possibilités encodant nos croyances initiales, et  $\Pi'(\phi_i)$  l'incertitude planant sur un événement  $\phi_i$ , incertitude notée sous la forme  $(\phi_i, \alpha_i)$  tel que  $\Pi'(\phi_i) = \alpha_i$ . Alors la distribution de possibilités  $\pi'$  révisée selon la règle de Jeffrey [DP97] doit satisfaire les conditions suivantes :

1.  $\forall \phi_i, \Pi'(\phi_i) = \alpha_i$
2.  $\forall \phi_i \subseteq \Omega, \forall \psi \subseteq \Omega, \Pi'(\psi | \phi_i) = \Pi(\psi | \phi_i)$

Définissons le conditionnement de Jeffrey dans le cadre quantitatif puis qualitatif. La révision basée sur la règle de Jeffrey dans le cadre quantitatif a été formalisée comme suit par Dubois et Prade [DP97] :

**Définition 16.** Soient  $\pi$  une distribution de possibilités et  $\phi_1, \dots, \phi_n$  un ensemble exhaustif et mutuellement exclusif d'événements où l'incertitude est de la forme  $\Pi'(\phi_i) = \alpha_i$ . Le degré de possibilités révisé d'un événement  $\psi \subseteq \Omega$  est calculé à l'aide de la formule :

$$\forall \psi \subseteq \Omega, \Pi'(\psi) = \max_{\phi_i} (\alpha_i * \frac{\Pi(\psi, \phi_i)}{\Pi(\phi_i)}) \quad (3.17)$$

D'après la Définition 16, le degré de possibilités révisé d'une interprétation  $\omega_j \in \Omega$  est calculé suivant le Lemme 3.1.

**Lemme 3.1.**

$$\forall \omega_j \in \phi_i, \pi'(\omega_j) = \alpha_i * \frac{\pi(\omega_j)}{\Pi(\phi_i)}. \quad (3.18)$$

Il est important de noter que la Définition 16 satisfait les axiomes de la théorie des possibilités (la distribution révisée est normalisée).

La révision de croyances en possibilités dans le cadre qualitatif, peut être réalisée par la règle de Jeffrey de la Définition 17 proposée également par Dubois et Prade [DP97].

**Définition 17.** Soient  $\pi$  une distribution de possibilités et  $\phi_1, \dots, \phi_n$  un ensemble exhaustif et mutuellement exclusif d'événements où l'incertitude est de la forme  $\Pi(\phi_i) = \alpha_i$ . Le degré de possibilités révisé d'un événement  $\psi \subseteq \Omega$  est calculé à l'aide de la formule :

$$\forall \psi \subseteq \Omega, \Pi'(\psi) = \max_{\phi_i} (\min(\Pi(\psi | \phi_i), \alpha_i)) \quad (3.19)$$

Tout comme pour le cadre quantitatif, nous vérifions facilement que  $\Pi'$  de la Définition 17 satisfait les 3 axiomes définissant une mesure de possibilités (cf. 1.3). D'après la Définition 17, le degré de possibilités révisé d'une interprétation  $\omega_j \in \Omega$  est calculé suivant le Lemme 3.2.

**Lemme 3.2.**

$$\forall \omega_j \in \phi_i, \pi'(\omega_j) = \begin{cases} \alpha_i & \text{si } \omega_j \geq \alpha_i \text{ ou } \pi(\omega_j) = \Pi(\phi_i); \\ \pi(\omega_j) & \text{sinon.} \end{cases} \quad (3.20)$$

### 3.6.3 Transformations et conditionnement de Jeffrey

Regardons les transformations quand nous conditionnons avec des informations incertaines. Il s'agit de comparer la table construite en révisant la distribution de possibilités transformée, et la table transformée à partir de la distribution de probabilités révisée. Nous nous intéressons à deux problèmes, en premier savoir si les deux transformations sont équivalentes, puis de savoir si l'ordre des interprétations est le même. Commençons par un petit exemple.

**Exemple 15.** Étant donné deux variables  $A$  et  $B$ , avec pour domaine  $D_A = \{a, \neg a\}$  et  $D_B = \{b, \neg b\}$ , la distribution de probabilités de la Table 3.9 exprime le degré pour chaque monde.

$A$	$B$	$p(A, B)$
$a$	$b$	0.4
$a$	$\neg b$	0.1
$\neg a$	$b$	0.2
$\neg a$	$\neg b$	0.3

TABLE 3.9 – Distribution de probabilités de l'Exemple 15

A travers la Table 3.9, on peut lire la probabilité de l'événement  $P(\{A = a\}) = 0.5$ . Une nouvelle information nous fait revoir la probabilité de cet événement, et  $P'(\{A = a\}) = 0.7$  donc  $P'(\{A = \neg a\}) = 0.3$ . Selon la règle de Jeffrey, la nouvelle distribution révisée  $p'$  donne lieu par exemple à ce calcul,  $p'(ab) = p(ab|a) * p'(a) + p(ab|\neg a) * p'(\neg a) = 0.8 * 0.7 + 0 * 0.3 = 0.56$ . Le reste des calculs remplit la nouvelle table comme suit :

$A$	$B$	$p'(A, B)$
$a$	$b$	0.56
$a$	$\neg b$	0.14
$\neg a$	$b$	0.12
$\neg a$	$\neg b$	0.18

TABLE 3.10 – Distribution de probabilités révisée de la Table 3.9

Transformons maintenant ces tables par  $OT$ , nous obtenons les Tables 3.11 et 3.12. Afin de comparer les tables révisées, il nous faut réviser dans le cadre possibiliste (ici, quantitatif).

$A$	$B$	$\pi_{OT}(A, B)$
$a$	$b$	1
$a$	$\neg b$	0.1
$\neg a$	$b$	0.3
$\neg a$	$\neg b$	0.6

TABLE 3.11 – Distribution de possibilités résultante de la transformation de la Table 3.9 par  $OT$

$A$	$B$	$\pi_{r-OT}(A, B)$
$a$	$b$	1
$a$	$\neg b$	0.26
$\neg a$	$b$	0.12
$\neg a$	$\neg b$	0.44

TABLE 3.12 – Distribution de possibilités révisée résultante de la transformation de la Table 3.10 par  $OT$

Dans la Table révisée 3.13, la révision est faite avec la distribution de possibilités  $\Pi'(\{A = a\}) = 1$  et  $\Pi'(\{A = \neg a\}) = 0.3$  (ie. la transformée par  $OT$  de la distribution de probabilités utilisée pour la révision dans le cadre probabiliste).

Il s'agit maintenant de comparer les Tables 3.12 et 3.13 :

$A$	$B$	$\pi_{OT-r}(A, B)$
$a$	$b$	1
$a$	$\neg b$	0.1
$\neg a$	$b$	0.15
$\neg a$	$\neg b$	0.3

TABLE 3.13 – Distribution de possibilités de la Table 3.11 révisée à l'aide de l'Équation 3.17

Dans la Table 3.12,  $\pi'(a\neg b) > \pi'(\neg ab)$

Alors que dans la Table 3.13,  $\pi'(a\neg b) < \pi'(\neg ab)$

Ceci nous permet de conclure, que lorsque la nouvelle information n'est pas certaine, alors la transformation  $OT$  ne préserve pas l'ordre des interprétations. Ce contre-exemple prouve la Proposition 3.11.

**Proposition 3.11.** Soient  $\pi_{r-OT}$  la distribution de probabilités révisée puis transformée par  $OT$ , et  $\pi_{OT-r}$  la distribution de possibilités révisée avec la règle de Jeffrey de l'Équation (3.1).

$$\forall \omega_i, \omega_j \in \Omega, \pi_{r-OT}(\omega_i) < \pi_{r-OT}(\omega_j) \not\Rightarrow \pi_{OT-r}(\omega_i) < \pi_{OT-r}(\omega_j) \quad (3.21)$$

La Proposition 3.11 a été prouvée dans le cas de la transformation  $OT$ , mais est-ce le cas pour les autres transformations, voyons la généralisation à n'importe quelle transformation de la Proposition (3.11).

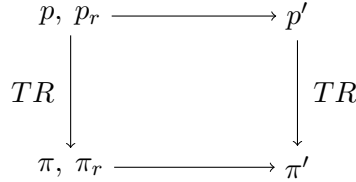


FIGURE 3.4 – Préservation du conditionnement avec des informations incertaines

**Proposition 3.12.** Soient  $\pi_{r-TR}$  la distribution de probabilités révisée puis transformée par  $TR$ , et  $\pi_{TR-r}$  la distribution de possibilités transformée puis révisée.

$$\forall \omega_i, \omega_j \in \Omega, \pi_{r-TR}(\omega_i) < \pi_{r-TR}(\omega_j) \not\Rightarrow \pi_{TR-r}(\omega_i) < \pi_{TR-r}(\omega_j) \quad (3.22)$$

*Démonstration.* Le but, ici, est de prouver que réviser avec des informations incertaines puis transformer, ne donne ni la même distribution, ni le même ordre des interprétations que de transformer d'abord puis de réviser. La proposition ci-dessus tient dans les deux cadres possibilistes, il nous faut donc le vérifier dans le cadre quantitatif et qualitatif. Attaquons nous d'abord au cadre quantitatif.

Considérons la distribution de probabilités de la Table 3.14, avec  $P(a) = \alpha_1 + \alpha_2$  et  $P(\neg a) = \alpha_2 + \alpha_2$  et  $\alpha_1 > \alpha_2$ , nous voulons maintenant réviser notre distribution  $p(A, B)$  en prenant en compte que  $p(A)$  représente maintenant l'ignorance totale. La distribution  $p'(A, B)$  de la Table 3.14 établit la révision de  $p(A, B)$  par la distribution équiprobable  $P'(a) = 0.5$  et  $P'(\neg a) = 0.5$ .

$A$	$B$	$p(A, B)$	$p'(A, B)$
$a$	$b$	$\alpha_1$	$\frac{\alpha_1}{\alpha_1 + \alpha_2} \times 0.5$
$a$	$\neg b$	$\alpha_2$	$\frac{\alpha_2}{\alpha_1 + \alpha_2} \times 0.5$
$\neg a$	$b$	$\alpha_2$	0.25
$\neg a$	$\neg b$	$\alpha_2$	0.25

TABLE 3.14 – Distribution de probabilités et sa révision

Nous souhaitons comparer la distribution de possibilités révisée de la Table 3.16 avec la Table 3.17. La première représentant la distribution transformée de la Table 3.14 et la deuxième, la distribution révisée de  $\pi(A, B)$  de la Table 3.15. La révision de la distribution de possibilités se fait avec la transformée de la distribution qui nous a servi dans le cadre probabiliste. Cette distribution de probabilités décrivait l'ignorance totale, ainsi, sa transformée le doit aussi, la révision est donc faite à partir de la distribution de possibilités  $\Pi'(a) = 1$  et  $\Pi'(\neg a) = 1$ .

Pour construire la distribution de possibilités  $\pi'(A, B)$  qui s'obtient par la transformation de  $p'(A, B)$  de la Table 3.14, il nous faut connaître l'ordre des interprétations. Ici,  $p'(ab) > p'(\neg ab) = p'(\neg a \neg b) > p'(a \neg b)$ , et puisque la transformation préserve l'ordre, nous avons avec la Table 3.16, où  $\gamma_1 < \gamma_2$ .

Dans le but maintenant, de comparer nos tables, il nous faut réviser avec la règle de Jeffrey dans le cadre possibiliste la Table 3.15, la Table 3.17 résulte de cette révision.

$A$	$B$	$\pi_{TR}(A, B)$
$a$	$b$	1
$a$	$\neg b$	$\beta$
$\neg a$	$b$	$\beta$
$\neg a$	$\neg b$	$\beta$

TABLE 3.15 – Distribution transformée de la Table 3.14

$A$	$B$	$\pi_{r-TR}(A, B)$
$a$	$b$	1
$a$	$\neg b$	$\gamma_1$
$\neg a$	$b$	$\gamma_2$
$\neg a$	$\neg b$	$\gamma_2$

TABLE 3.16 – Distribution de probabilités  $p'(A, B)$  de la Table 3.14 transformée par une transformation  $TR$ 

$A$	$B$	$\pi_{TR-r}(A, B)$
$a$	$b$	1
$a$	$\neg b$	$\beta$
$\neg a$	$b$	1
$\neg a$	$\neg b$	1

TABLE 3.17 – Distribution de possibilités de la Table 3.15 révisée à l'aide de l'Équation 3.17

Sans forcément avoir besoin de connaître l'ordre des interprétations représentées par des inconnues, nous pouvons tout de même constater que l'ordre n'est pas préservé étant donné que dans notre distribution de la Table 3.17, trois interprétations sont au même niveau, en effet nous avons  $\pi'(ab) = \pi'(\neg ab) = \pi'(\neg a\neg b) = 1$ . Dans la Table 3.16, une seule interprétation nous donne une possibilité de 1.

Montrons maintenant dans le cadre qualitatif, c'est à dire que cette fois-ci au lieu de réviser avec l'Équation 3.15, nous révisons avec l'Équation 3.19, ce qui nous donne la Table 3.18. Remarquons que

$A$	$B$	$\pi_{TR-r}(A, B)$
$a$	$b$	1
$a$	$\neg b$	$\beta$
$\neg a$	$b$	1
$\neg a$	$\neg b$	1

TABLE 3.18 – Distribution de possibilités de la Table 3.15 révisée à l'aide de l'Équation 3.19

c'est exactement la même table que lorsque nous avons révisé dans le cadre quantitatif (Table 3.17), ce qui nous permet de tirer les mêmes conclusions que dans le cadre quantitatif.

Nous concluons donc, et ceci dans les deux cadres, que la révision avec des informations incertaines ne garde pas l'ordre des interprétations, et par conséquent, ne donne pas du tout les mêmes distributions.  $\square$

### 3.7 Transformations et requêtes MAP

L'étude faite pour la révision nous amène à nous intéresser aux requêtes MAP, qui utilisent le conditionnement et la marginalisation. Il s'agit de chercher la valeur la plus plausible d'un sous ensemble de variables étant donnée une observation.

Soit la structure suivante  $\Omega$  avec des variables  $A_1, \dots, A_n$ . Soit  $\Omega = D_{A_1} * \dots * D_{A_n}$  (avec  $A_1, \dots, A_n$  des variables). Soit  $A' \subseteq \Omega$  un sous-ensemble de variables et  $\phi \subseteq \Omega$  un événement, nous souhaitons marginaliser sur  $\Omega \setminus A'$  pour trouver la distribution a posteriori sur  $A'$ . La question est de savoir si l'égalité suivante est vérifiée :

$$\operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{TR}(a' \mid \phi)) \quad (3.23)$$

**Proposition 3.13.** *Soit  $p$  une distribution de probabilités et  $\pi_{TR}$  la distribution de possibilités résultant de la transformation  $TR$ , et soit  $\phi$  une observation certaine.*

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{TR}(a' \mid \phi)) \quad (3.24)$$

*Démonstration.* Trivial. A l'aide de la Proposition 4.6, nous pouvons conclure que les requêtes MAP étant basées sur la marginalisation, ne préservent pas l'ordre.  $\square$

Nous avons vu dans le cas où l'observation est certaine que l'ordre des interprétations est préservé, cependant la marginalisation ne préserve pas l'ordre, ainsi nous pouvons écrire le corollaire suivant :

**Corollaire 3.3.** *Soit  $p$  une distribution de probabilités et  $\pi_{TR}$  la distribution de possibilités résultant de la transformation  $TR \in \{OT, KT, ST, VT\}$ , et soit  $\phi$  une observation certaine.*

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{OT}(a' \mid \phi)) \quad (3.25)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{KT}(a' \mid \phi)) \quad (3.26)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{ST}(a' \mid \phi)) \quad (3.27)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{VT}(a' \mid \phi)) \quad (3.28)$$

Dans le cas où l'observation est incertaine, le conditionnement de Jeffrey ne préserve pas l'ordre des interprétations, et vu que la marginalisation non plus, nous avons :

**Corollaire 3.4.** *Soit  $p$  une distribution de probabilités et  $\pi_{TR}$  la distribution de possibilités résultant de la transformation  $TR \in \{OT, KT, ST, VT\}$ , et soit  $\phi$  une observation incertaine.*

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{OT}(a' \mid \phi)) \quad (3.29)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{KT}(a' \mid \phi)) \quad (3.30)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{ST}(a' \mid \phi)) \quad (3.31)$$

$$a = \operatorname{argmax}_{a' \in D_{A'}}(p(a' \mid \phi)) \not\Rightarrow a = \operatorname{argmax}_{a' \in D_{A'}}(\pi_{VT}(a' \mid \phi)) \quad (3.32)$$



## Conclusion

Dans ce chapitre, nous avons voulu nous intéresser aux transformations au niveau des distributions, en étudiant leurs effets sur la normalisation, les requêtes MPE et MAP, ainsi que sur la marginalisation et le conditionnement. La conclusion générale est qu'aucune transformation ne préserve l'ordre des interprétations lors de la marginalisation et du conditionnement avec des informations incertaines. Il est intéressant de noter que quand nous ne nous intéressons pas explicitement à la probabilité ou possibilité d'une interprétation ou d'un événement comme pour les requêtes MPE, les deux cadres (probabilités et possibilités) nous donnent les mêmes résultats pour MPE.

L'auteur du papier [Sud92], s'est intéressé aux sujets des transformations aussi, en se posant d'autres questions. Travaillant sur les distributions aussi, son travail a visé plus particulièrement les propriétés qu'il appelle de second ordre : l'indépendance, la marginalisation et le conditionnement. Il a cherché à savoir si les distributions lorsque nous les transformons puis appliquons une propriété étaient les mêmes que si nous appliquions en premier une propriété et ensuite nous les transformions. Il a donc prouvé qu'aucune transformation ne donnerait les mêmes distributions. Alors qu'il s'est intéressé aux distributions et que dans le cadre qualitatif, nous avons voulu nous intéresser à l'ordre des interprétations dans les deux cadres possibilistes.

Dans le chapitre suivant, nous nous posons les questions de ce chapitre, mais en travaillant avec les modèles graphiques.

## Chapitre 4

# Analyse des transformations dans les modèles graphiques

### Sommaire

---

<b>4.1 Transformations et Indépendances</b> . . . . .	<b>40</b>
4.1.1 Définitions . . . . .	41
4.1.2 Transformation $OT$ . . . . .	41
4.1.3 Transformation $KT$ . . . . .	43
4.1.4 Transformation $ST$ . . . . .	44
4.1.5 Transformation $VT$ . . . . .	45
<b>4.2 Préservation des indépendances</b> . . . . .	<b>46</b>
4.2.1 Préservation des indépendances d'événements dans le cadre qualitatif . . . . .	46
4.2.2 Conséquences . . . . .	48
<b>4.3 Transformations et modèles graphiques</b> . . . . .	<b>49</b>
4.3.1 Transformer un réseau Bayésien . . . . .	49
4.3.2 Préservation des distributions jointes . . . . .	50
4.3.3 Préservation des distributions jointes dans le cadre qualitatif . . . . .	51
4.3.4 Préservation des distributions jointes dans le cadre quantitatif . . . . .	53

---

Ce chapitre présente les transformations dans les modèles graphiques. Les modèles graphiques se basent sur les relations d'indépendances pour factoriser une distribution globale. Nous étudions dans un premier temps, la préservation d'indépendances dans les distributions, puis pour palier ce problème, nous introduisons les transformations de modèles graphiques. Nous basons le reste du chapitre sur l'étude des distributions jointes résultantes de la transformation du réseau.

### 4.1 Transformations et Indépendances

Le chapitre précédent au sujet des transformations nous montre clairement que nous perdons de l'information en passant d'une distribution de probabilités à une distribution de possibilités. Nous nous intéressons au cas des indépendances, en partant d'une distribution de probabilités, nous montrons que certaines indépendances ne sont pas forcément préservées. Rappelons d'abord certaines définitions.

### 4.1.1 Définitions

#### Indépendances marginales

- *Indépendance d'événements* : Soient deux événements  $\phi \subseteq \Omega$  et  $\psi \subseteq \Omega$  et  $p$  une distribution de probabilités.

*Si  $P(\phi | \psi) = P(\phi)$  alors  $\phi$  et  $\psi$  sont dits indépendants dans  $p$ .*

- *Indépendance de variables* : Soient  $A_1$  et  $A_2$  deux variables avec leurs domaines respectifs  $D_1$  et  $D_2$ .

*$A_1$  est indépendant de  $A_2$  si  $\forall a_1 \in D_1, \forall a_2 \in D_2, P(a_1 | a_2) = P(a_1)$*

Nous noterons  $P(A_1 | A_2) = P(A_1)$  ou  $A_1 \perp A_2$  pour simplifier les notations.

De manière similaire, sur une distribution de possibilités  $\pi$  :

- $\forall \phi \subseteq \Omega, \forall \psi \subseteq \Omega$ , si  $\Pi(\phi | \psi) = \Pi(\phi)$  alors  $\phi$  et  $\psi$  sont dits indépendants.
- Si  $\Pi(A_1 | A_2) = \Pi(A_1)$  alors  $A_1$  et  $A_2$  sont dits indépendants.

#### Indépendances conditionnelles

- *Indépendance d'événements* : Soient trois événements  $\phi \subseteq \Omega, \alpha \subseteq \Omega$  et  $\psi \subseteq \Omega$  et  $p$  une distribution de probabilités.

*Si  $P(\phi | \psi\alpha) = P(\phi | \alpha)$  alors  $\phi$  et  $\psi$  sont indépendants sachant  $\alpha$*

- *Indépendance de variables* : Soient  $A_1, A_2$  et  $A_3$  trois variables avec leurs domaines respectifs  $D_1, D_2$  et  $D_3$ .

*$A_1$  est indépendante de  $A_2$  sachant  $A_3$  si  $P(A_1 | A_2A_3) = P(A_1 | A_3)$*

Nous noterons  $A_1 \perp A_2 | A_3$  cette relation d'indépendance conditionnelle pour simplifier des notations.

De manière similaire, sur une distribution de possibilités  $\pi$  :

- Si  $\Pi(\phi | \psi\alpha) = \Pi(\phi | \alpha)$  alors  $\phi$  et  $\psi$  sont dits indépendants conditionnellement à  $\alpha$
- Si  $\Pi(A_1 | A_2A_3) = \Pi(A_1 | A_3)$  alors  $A_1$  et  $A_2$  sont dits indépendants conditionnellement  $A_3$

Nous allons, dans un premier temps, voir si les transformations  $OT, KT, ST$  et  $VT$  préservent les indépendances puis dans un second temps poser la question de préservation des indépendances dans le cas général.

### 4.1.2 Transformation $OT$

**Proposition 4.1.** *Soit  $p$  une distribution de probabilités,  $\pi = OT(p)$  la distribution de possibilités résultant de la transformation  $OT$ . Soit  $I_p$  (resp.  $I_\pi$ ) l'ensemble des indépendances dans  $p$  (resp.  $\pi$ ).*

$$I \in I_p \not\Rightarrow I \in I_\pi \quad (4.1)$$

Le contre-exemple suivant nous permet de dire que  $OT$  ne préserve pas toutes les indépendances.

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{OT}(A, B, C)$
$a$	$b$	$c$	0.09	0.292
$a$	$b$	$\neg c$	0.09	0.292
$a$	$\neg b$	$c$	0.21	0.712
$a$	$\neg b$	$\neg c$	0.21	0.712
$\neg a$	$b$	$c$	0.288	1
$\neg a$	$b$	$\neg c$	0.032	0.04
$\neg a$	$\neg b$	$c$	0.072	0.112
$\neg a$	$\neg b$	$\neg c$	0.008	0.008

TABLE 4.1 – Distribution de probabilités et sa transformation par  $OT$

$A$	$B$	$C$	$\pi_{OT}(B _pAC)$	$\pi_{OT}(B _mAC)$
$a$	$b$	$c$	$\approx 0.41$	0.292
$a$	$b$	$\neg c$	$\approx 0.41$	0.292
$a$	$\neg b$	$c$	1	1
$a$	$\neg b$	$\neg c$	1	1
$\neg a$	$b$	$c$	1	1
$\neg a$	$b$	$\neg c$	1	1
$\neg a$	$\neg b$	$c$	0.112	0.112
$\neg a$	$\neg b$	$\neg c$	0.2	0.008

TABLE 4.2 – Distribution de possibilités de  $B$  sachant  $AC$  après transformation par  $OT$

$A$	$B$	$\pi_{OT}(B _pA)$	$\pi_{OT}(B _mA)$
$a$	$b$	$\approx 0.41$	0.292
$a$	$\neg b$	1	1
$\neg a$	$b$	1	1
$\neg a$	$\neg b$	0.112	0.112

TABLE 4.3 – Distribution de possibilités de  $B$  sachant  $A$  après transformation par  $OT$

*Contre-exemple.* La distribution de probabilités de la Table 4.1 nous donne l'indépendance :  $B \perp C | A$ , les calculs de  $\Pi_{OT}(B|AC)$  et  $\Pi_{OT}(B|A)$  sont présentés dans les Tables 4.2 et 4.3. Vérifions si en transformant par  $OT$ , l'indépendance  $B \perp C | A$  est conservée.

Les Tables 4.2 et 4.3 nous montre que  $B$  et  $C$  qui étaient indépendants sachant  $A$ , ne le sont plus (ni dans le cadre qualitatif, ni dans le cadre quantitatif). En effet, dans le cadre quantitatif :

$$\left. \begin{array}{l} \pi(\neg b|_p\neg a\neg c) = 0.2 \\ \pi(\neg b|_p\neg a) = 0.112 \end{array} \right\} \text{donc } \pi(\neg b|_p\neg a\neg c) \neq \pi(\neg b|_p\neg a)$$

Et dans le cadre qualitatif :

$$\left. \begin{array}{l} \pi(\neg b|_m\neg a\neg c) = 0.008 \\ \pi(\neg b|_m\neg a) = 0.112 \end{array} \right\} \text{donc } \pi(\neg b|_m\neg a\neg c) \neq \pi(\neg b|_m\neg a)$$

Comme  $\pi(\neg b|_p\neg a\neg c) \neq \pi(\neg b|_p\neg a)$ , alors  $B \not\perp C | A$ . □

### 4.1.3 Transformation $KT$

**Proposition 4.2.** Soit  $p$  une distribution de probabilités,  $\pi = KT(p)$  la distribution de possibilités résultant de la transformation  $KT$ . Soit  $I_p$  (resp.  $I_\pi$ ) l'ensemble des indépendances dans  $p$  (resp.  $\pi$ ).

$$I \in I_p \not\Rightarrow I \in I_\pi \quad (4.2)$$

Comme  $OT$ , le contre-exemple suivant montre que la transformation  $KT$  ne préserve pas forcément toutes les indépendances.

*Contre-exemple.* La Table 4.4 représente la distribution de probabilités et sa transformation par  $KT$ . L'indépendance décrite ci dessus est :  $A \perp C$ . Le calcul  $\Pi_{KT}(A|C)$  et  $\Pi_{KT}(A)$  est présenté dans les

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{KT}(A, B, C)$
$a$	$b$	$c$	0.054	0.2
$a$	$b$	$\neg c$	0.032	$\approx 0.12$
$a$	$\neg b$	$c$	0.006	$\approx 0.02$
$a$	$\neg b$	$\neg c$	0.008	$\approx 0.03$
$\neg a$	$b$	$c$	0.27	1
$\neg a$	$b$	$\neg c$	0.144	$\approx 0.53$
$\neg a$	$\neg b$	$c$	0.27	1
$\neg a$	$\neg b$	$\neg c$	0.216	0.8

TABLE 4.4 – Distribution de probabilités et sa transformation par  $KT$

Tables 4.5 et 4.6.

$A$	$C$	$\pi_{KT}(A _p C)$	$\pi_{KT}(A _m C)$
$a$	$c$	0.2	0.2
$a$	$\neg c$	$\approx 0.148$	$\approx 0.12$
$\neg a$	$c$	1	1
$\neg a$	$\neg c$	1	1

TABLE 4.5 – Distribution de possibilités de  $A$  sachant  $C$  après transformation par  $KT$

$A$	$\pi_{KT}(A)$
$a$	0.2
$\neg a$	1

TABLE 4.6 – Distribution de possibilités de  $A$  après transformation par  $KT$

Nous remarquons dans ces tables que  $A$  et  $C$  qui étaient indépendants, ne le sont plus (ni dans le cadre qualitatif, ni dans le cadre quantitatif). En effet :

$$\left. \begin{array}{l} \pi(a|_p c) \approx 0.148 \\ \pi(a) = 0.2 \end{array} \right\} \text{ donc } \pi(a|_p c) \neq \pi(a)$$

Et dans le cadre qualitatif :

$$\left. \begin{array}{l} \pi(a|_m c) \approx 0.12 \\ \pi(a) = 0.2 \end{array} \right\} \text{ donc } \pi(a|_m c) \neq \pi(a)$$

Comme  $\pi(a|_m c) \neq \pi(a)$  alors  $A \not\perp C$ . □

#### 4.1.4 Transformation $ST$

**Proposition 4.3.** Soit  $p$  une distribution de probabilités,  $\pi = ST(p)$  la distribution de possibilités résultant de la transformation  $ST$ . Soit  $I_p$  (resp.  $I_\pi$ ) l'ensemble des indépendances dans  $p$  (resp.  $\pi$ ).

$$I \in I_p \not\Rightarrow I \in I_\pi \quad (4.3)$$

Dans le contre-exemple suivant, nous montrons que  $ST$  ne préserve pas forcément toutes les indépendances.

*Contre-exemple.* Dans la distribution de probabilités de la Table 4.7,  $B \perp C | A$ . Vérifions si en transformant par  $ST$ , nous gardons cette indépendance. Les calculs de  $\pi_{ST}(B|AC)$  et  $\pi_{ST}(B|A)$  sont présentés dans les Tables 4.8 et 4.9.

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{OT}(A, B, C)$
$a$	$b$	$c$	0.09	0.562
$a$	$b$	$\neg c$	0.09	0.562
$a$	$\neg b$	$c$	0.21	0.922
$a$	$\neg b$	$\neg c$	0.21	0.922
$\neg a$	$b$	$c$	0.288	1
$\neg a$	$b$	$\neg c$	0.032	0.232
$\neg a$	$\neg b$	$c$	0.072	0.472
$\neg a$	$\neg b$	$\neg c$	0.008	0.064

TABLE 4.7 – Distribution de probabilités et sa transformation par  $ST$

$A$	$B$	$C$	$\pi_{ST}(B _p AC)$	$\pi_{ST}(B _m AC)$
$a$	$b$	$c$	$\approx 0.61$	0.562
$a$	$b$	$\neg c$	$\approx 0.61$	0.562
$a$	$\neg b$	$c$	1	1
$a$	$\neg b$	$\neg c$	1	1
$\neg a$	$b$	$c$	1	1
$\neg a$	$b$	$\neg c$	1	1
$\neg a$	$\neg b$	$c$	0.472	0.472
$\neg a$	$\neg b$	$\neg c$	$\approx 0.28$	0.064

TABLE 4.8 – Distribution de possibilités de  $B$  sachant  $AC$  après transformation par  $ST$

$A$	$B$	$\pi_{ST}(B _p A)$	$\pi_{ST}(B _m A)$
$a$	$b$	$\approx 0.61$	0.562
$a$	$\neg b$	1	1
$\neg a$	$b$	1	1
$\neg a$	$\neg b$	0.472	0.472

TABLE 4.9 – Distribution de possibilités de  $B$  sachant  $A$  après transformation par  $ST$

Dans ces tables,  $B$  et  $C$  ne sont indépendants ni dans le cadre qualitatif, ni dans le cadre quantitatif. En effet, le calcul dans le cadre quantitatif montre que :

$$\left. \begin{array}{l} \pi(\neg b|_p \neg a \neg c) \approx 0.28 \\ \pi(\neg b|_p \neg a) = 0.472 \end{array} \right\} \text{ donc } \pi(\neg b|_p \neg a \neg c) \neq \pi(\neg b|_p \neg a)$$

Et dans le cadre qualitatif :

$$\left. \begin{array}{l} \pi(\neg b|_m \neg a \neg c) = 0.064 \\ \pi(\neg b|_m \neg a) = 0.472 \end{array} \right\} \text{ donc } \pi(\neg b|_m \neg a \neg c) \neq \pi(\neg b|_m \neg a)$$

Comme  $\pi(\neg b|\neg a \neg c) \neq \pi(\neg b|\neg a)$ , alors  $B \not\perp C | A$ . □

#### 4.1.5 Transformation VT

**Proposition 4.4.** Soit  $p$  une distribution de probabilités,  $\pi = VT(p)$  la distribution de possibilités résultant de la transformation VT. Soit  $I_p$  (resp.  $I_\pi$ ) l'ensemble des indépendances dans  $p$  (resp.  $\pi$ ).

$$I \in I_p \Leftrightarrow I \in I_\pi \tag{4.4}$$

Le contre-exemple suivant montre que les indépendances ne sont pas toujours conservées.

*Contre-exemple.* Soient  $A, B$  et  $C$ , 3 variables ayant pour domaine  $D_A = \{a, \neg a\}$ ,  $D_B = \{b, \neg b\}$  et  $D_C = \{c, \neg c\}$ . L'ensemble  $\Omega$  des mondes est le produit cartésien des domaines respectifs des variables  $A, B$  et  $C$ .

La constante  $k$  utilisée dans la transformation VT est définie, pour rappel, comme suit :

$$0 \leq k \leq \frac{\log(p_n)}{(1-p_n) \cdot \log(\frac{p_n}{p_1})}$$

A	B	C	$p(A, B, C)$	$\pi_{VT}(A, B, C)$
a	b	c	0.09	≈0.35
a	b	¬c	0.09	≈0.35
a	¬b	c	0.21	≈0.78
a	¬b	¬c	0.21	≈0.78
¬a	b	c	0.288	1
¬a	b	¬c	0.032	≈0.12
¬a	¬b	c	0.072	≈0.28
¬a	¬b	¬c	0.008	≈0.03

TABLE 4.10 – Distribution de probabilités et sa transformation par VT

La constante  $k$  devait être choisie entre 0 et  $\frac{\log(p_n)}{(1-p_n) \cdot \log(\frac{p_n}{p_1})} \approx 1.35$ , ici nous avons choisi pour le contre exemple  $k = 1$  pour simplifier les calculs. Cependant, nous avons constaté, et il est important de noter, qu'il n'existe pas de  $k$  qui nous permettra de garder les indépendances (nous verrons par la suite pourquoi). La Table 4.10 engendre l'indépendance (dans le cadre probabiliste) suivante :  $B \perp C | A$ . Vérifions si après transformation, on a toujours la même chose, dans le cadre qualitatif et quantitatif. Calculons donc  $\Pi_{VT}(B|AC)$  et  $\Pi_{VT}(B|A)$ . Les résultats sont présentés dans les Tables 4.11 et 4.12.

Nous concluons de ces tables que  $B$  et  $C$  qui étaient indépendants, ne le sont plus (ni dans le cadre qualitatif, ni dans le cadre quantitatif). En effet nous avons calculé dans le cadre quantitatif :

$$\left. \begin{array}{l} \pi(\neg b|_p \neg a \neg c) \approx 0.24 \\ \pi(\neg b|_p \neg a) \approx 0.28 \end{array} \right\} \text{ donc } \pi(\neg b|_p \neg a \neg c) \neq \pi(\neg b|_p \neg a)$$

Et dans le cadre qualitatif :

$$\left. \begin{array}{l} \pi(\neg b|_m \neg a \neg c) \approx 0.03 \\ \pi(\neg b|_m \neg a) \approx 0.28 \end{array} \right\} \text{ donc } \pi(\neg b|_m \neg a \neg c) \neq \pi(\neg b|_m \neg a)$$

Comme  $\pi(\neg b|\neg a \neg c) \neq \pi(\neg b|\neg a)$ , alors  $B \not\perp C | A$ . □

A	B	C	$\pi_{VT}(B _pAC)$	$\pi_{VT}(B _mAC)$
a	b	c	$\approx 0.45$	$\approx 0.35$
a	b	$\neg c$	$\approx 0.45$	$\approx 0.35$
a	$\neg b$	c	1	1
a	$\neg b$	$\neg c$	1	1
$\neg a$	b	c	1	1
$\neg a$	b	$\neg c$	1	1
$\neg a$	$\neg b$	c	$\approx 0.28$	$\approx 0.28$
$\neg a$	$\neg b$	$\neg c$	$\approx 0.24$	$\approx 0.03$

TABLE 4.11 – Distribution de possibilités de B sachant AC après transformation par VT

A	B	$\pi_{VT}(B _pA)$	$\pi_{VT}(B _mA)$
a	b	$\approx 0.45$	$\approx 0.35$
a	$\neg b$	1	1
$\neg a$	b	1	1
$\neg a$	$\neg b$	$\approx 0.28$	$\approx 0.28$

TABLE 4.12 – Distribution de possibilités de B sachant A après transformation par VT

Jusqu'ici, nous avons montré que les transformations OT, KT, ST et VT ne préservent pas toutes les indépendances, *qu'en est-il dans le cas général ? Peut-il exister une transformation qui préserve toutes les indépendances en passant de p à  $\pi$  ?* La section suivante répond à cette question.

## 4.2 Préservation des indépendances

Nous avons montré dans la Section 3.3 du chapitre précédent que l'ordre des événements n'était préservé pour aucune transformation, la question que nous nous posons ici est : *y a-t'il une répercussion sur la préservation des indépendances ?*

### 4.2.1 Préservation des indépendances d'événements dans le cadre qualitatif

La question immédiate est bien sûr de savoir si l'indépendance d'événements n'est pas préservée en général. Nous répondons à cette question d'abord dans le cadre qualitatif.

**Proposition 4.5.** *Soient 3 événements  $\phi, \psi$  et  $\alpha \subseteq \Omega$ , p une distribution de probabilités,  $\pi$  la transformée respectant les critères de préservation d'ordre.*

$$P(\phi | \psi\alpha) = P(\phi | \alpha) \Leftrightarrow \Pi(\phi |_{min} \psi\alpha) = \Pi(\phi |_{min} \alpha) \quad (4.5)$$

Pour rappel, avec 2 événements seulement, ce qui est un cas particulier de l'équation ci-dessus, la nouvelle formule s'écrit :

$$P(\phi | \psi) = P(\phi) \Leftrightarrow \Pi(\phi | \psi) = \Pi(\phi) \quad (4.6)$$

*Démonstration.* Soit la distribution de probabilités p et sa transformation  $\pi$  de la Table 4.13, où A et B sont indépendants en probabilités. Et soit l'ordre  $1 > \alpha_1 > \alpha_2 > \alpha_3 > \alpha_4 > 0$  avec  $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$ .

Comme la transformation préserve l'ordre des interprétations, nous obtenons  $1 > \beta_2 > \beta_3 > \beta_4$ . Dans le cadre qualitatif, nous devons avoir par exemple,  $\Pi(a|_{min}b) = \Pi(a)$  ie.  $\beta_4 = \beta_3$  or  $\beta_4 < \beta_3$  donc impossible d'avoir la préservation de l'indépendance entre A et B.  $\square$



$A$	$B$	$p(A, B)$	$\pi(A, B)$
$a$	$b$	$\alpha_4$	$TR(\alpha_4) = \beta_4$
$a$	$\neg b$	$\alpha_3$	$TR(\alpha_3) = \beta_3$
$\neg a$	$b$	$\alpha_2$	$TR(\alpha_2) = \beta_2$
$\neg a$	$\neg b$	$\alpha_1$	$TR(\alpha_1) = 1$

TABLE 4.13 – Distribution de probabilités et sa transformée

En clair, dans le cadre qualitatif, aucune transformation ne préservera, en général, toutes les indépendances d'événements. Cependant après quelques recherches, nous avons remarqué que dans certains cas, l'indépendance peut être préservée.

Nous a aussi concerné la question dans le cadre quantitatif, les résultants n'étant pas concluant, ils ne sont pas intégrés. Ce que que nous avons pu constater, c'est que dans certains cas, l'indépendance n'est pas préservé, et dans d'autres elle l'est. Mais rien ne nous a permis de généraliser cela.

Par contre, nous avons remarqué que lorsque nous avons l'interprétation la plus probable parmi les événements concernés dans l'intersection des événements, l'indépendance était préservée. Ce qui nous a permis d'écrire la Proposition 4.6.

**Proposition 4.6.** Soient 3 événements  $\phi, \psi$  et  $\alpha \subseteq \Omega$ , tels que l'indépendance  $\phi \perp \psi \mid \alpha$  est vérifiée dans le cadre probabiliste. Soit  $\omega_{max}$  le monde le plus probable parmi les modèles  $\omega_j \in \phi \cup \psi \cup \alpha$ .

$$\text{Si } \omega_{max} \in \alpha \cap \psi \cap \phi \text{ alors } \Pi(\phi \mid \psi\alpha) = \Pi(\phi \mid \alpha) \quad (4.7)$$

Ainsi,  $\phi$  et  $\psi$  sont indépendants conditionnellement à  $\alpha$  dans  $\pi$  obtenue par transformation de  $p$ .

*Démonstration.* Soient  $\Omega$  l'univers du discours et 3 événements  $\phi, \psi$  et  $\alpha \subseteq \Omega$ . Le Schéma 4.1 modélise ces 3 événements.

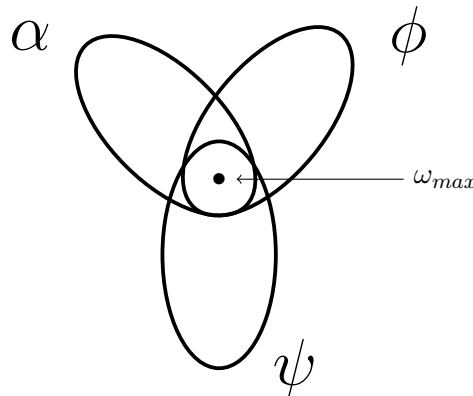


FIGURE 4.1 – Représentation des événements  $\phi, \psi$  et  $\alpha$  où  $\omega_{max} \in \phi \cap \psi \cap \alpha$

Partant du fait que  $\phi \perp \psi \mid \alpha$  dans  $p$ , et que  $\pi$  est obtenue par transformation de  $p$ , par conséquent, préserve l'ordre des interprétations, rien ne nous assure d'avoir

$$\Pi(\phi \mid \psi\alpha) = \Pi(\phi \mid \alpha)$$

Par contre, si le monde  $\omega_{max}$  est le monde le plus probable (le plus possible par conséquent) dans  $\phi \cup \psi \cup \alpha$ , alors :

$$\Pi(\phi \mid \psi\alpha) = \Pi(\phi \mid \alpha) \text{ puisque } \frac{\pi(\omega_{max})}{\pi(\omega_{max})} = \frac{\pi(\omega_{max})}{\pi(\omega_{max})}$$

Dans cette configuration et ce aussi bien dans le cadre qualitatif que dans le cadre quantitatif, nous pouvons donc affirmer que l'indépendance est préservée quelle que soit la transformation utilisée.  $\square$

### 4.2.2 Conséquences

Nous venons de montrer, dans le cadre qualitatif en particulier, que l'ordre des événements n'était pas préservé en général. Nous pouvons donc déduire le Corollaire 4.1.

**Corollaire 4.1.** *Puisque l'indépendance des événements n'est pas forcément préservée alors nous pouvons dire que les indépendances entre variables ne seront pas forcément préservées.*

*Idée de la preuve :* Calculer l'indépendance entre des variables revient à calculer l'indépendance d'événements où un événement est l'ensemble des mondes correspondant à l'affectation d'une variable à une valeur de son domaine. Comme l'indépendance d'événements n'est pas préservée, l'indépendance de variables ne l'est pas non plus.

Les recherches sur la préservation des indépendances de variables montre qu'en passant du cadre des probabilités au cadre des possibilités, nous n'obtenons pas souvent les indépendances trouvées en probabilités. Mais dans l'autre sens, *pouvons-nous avoir des indépendances en possibilités qui n'étaient pas présentes en probabilités ?*

Lors d'une transformation de  $p$  vers  $\pi$ , non seulement on a  $I \in I_p \not\Rightarrow I \in I_\pi$ , mais les transformations peuvent aussi créer de nouvelles indépendances. Voyons l'exemple avec la transformation  $KT$  dans le cadre quantitatif.

**Proposition 4.7.** *Soit  $p$  une distribution de probabilités,  $\pi_{KT}$  la distribution de possibilités résultant de la transformation  $KT$ . Soit  $I_p$  (resp.  $I_\pi$ ) l'ensemble des indépendances dans  $p$  (resp.  $\pi_{KT}$ ).*

$$I' \in I_\pi \not\Rightarrow I' \in I_p \tag{4.8}$$

Le contre-exemple suivant montre les deux sens de l'équation.

*Contre-exemple.* La Table 4.14 représente une distribution de probabilités et sa transformation par  $KT$ .

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{KT}(A, B, C)$
$a$	$b$	$c$	0.012	$\approx 0.036$
$a$	$b$	$\neg c$	0.096	$\approx 0.286$
$a$	$\neg b$	$c$	0.048	$\approx 0.143$
$a$	$\neg b$	$\neg c$	0.144	$\approx 0.429$
$\neg a$	$b$	$c$	0.07	$\approx 0.208$
$\neg a$	$b$	$\neg c$	0.224	$\approx 0.667$
$\neg a$	$\neg b$	$c$	0.07	$\approx 0.208$
$\neg a$	$\neg b$	$\neg c$	0.336	1

TABLE 4.14 – Distribution de probabilités et sa transformation par  $KT$

Dans  $p$ , l'unique indépendance trouvée est  $A \perp C$  alors qu'après la transformation dans  $\pi_{KT}$ , on a  $A \not\perp C$ . Cependant ce n'est pas la seule remarque que nous constatons, les Tables 4.15 et 4.16 montre

$A$	$B$	$\pi_{KT}(A _pB)$
$a$	$b$	$\approx 0.429$
$a$	$\neg b$	$\approx 0.429$
$\neg a$	$b$	1
$\neg a$	$\neg b$	1

TABLE 4.15 – Distribution de possibilités de  $A$  sachant  $B$  après transformation par  $KT$

$A$	$\pi_{KT}(A)$
$a$	$\approx 0.429$
$\neg a$	1

TABLE 4.16 – Distribution de possibilités de  $A$  après transformation par  $KT$

qu'en plus de perdre l'indépendance entre  $A$  et  $C$ , une nouvelle indépendance s'est créée, maintenant dans  $\pi_{KT}$ , nous avons l'indépendance  $A \perp B$ . Et en calculant,  $\pi_{KT}(A|_pC)$  et  $\pi_{KT}(A)$  nous n'obtenons l'indépendance entre  $A$  et  $C$ .  $\square$

### 4.3 Transformations et modèles graphiques

Un moyen de préserver les indépendances de variables est de passer par les modèles graphiques, en effet les modèles graphiques sont l'outil représentant les relations d'indépendances. Ainsi en transformant le modèle graphique probabiliste en modèle graphique possibiliste, les indépendances seront préservées.

Illustrons par la Figure 4.2, la transformation d'un réseau Bayésien en un réseau possibiliste et la transformation des distributions jointes.

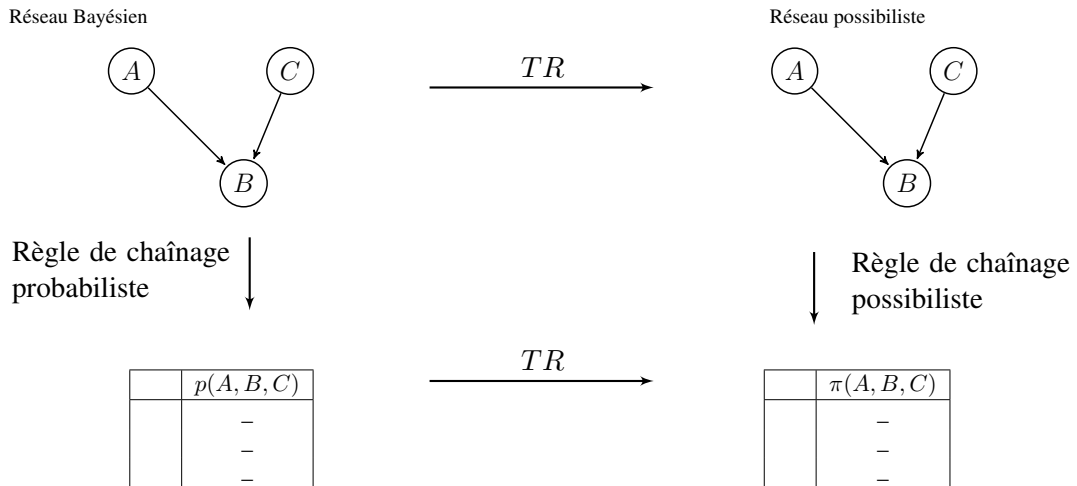


FIGURE 4.2 – Transformation de modèles graphiques

#### 4.3.1 Transformer un réseau Bayésien

Transformer un modèle graphique, ici un réseau Bayésien, consiste à transformer les tables locales et conserver la structure qui encode les relations d'indépendances conditionnelles.

**Définition 18.** Soit un réseau Bayésien  $G$ , un réseau possibiliste  $G_{TR}$  est un réseau composé de :

- Une composante graphique qui est la même que le réseau  $G$
- Une composante numérique, c'est à dire l'ensemble des tables locales de  $G$  transformées par  $TR$

Notons  $I_G$  l'ensemble des indépendances dans un réseau Bayésien  $G$  et  $I_{G_{TR}}$  l'ensemble des indépendances dans le réseau possibiliste transformé par  $TR$ .

**Proposition 4.8.** Soient  $G$  un réseau Bayésien et  $G_{TR}$  le réseau possibiliste transformé par  $TR$ .

$$\forall I \in I_G \Rightarrow I \in I_{G_{TR}} \quad (4.9)$$

La preuve est triviale puisque c'est le même graphe qui encode les relations d'indépendances dans les deux réseaux alors il est clair que toutes les relations d'indépendances sont préservées. Illustrons cette proposition à l'aide d'un exemple.

**Exemple 16.** Considérons le réseau Bayésien de la Figure 4.3 et la transformation  $OT$ . Ce réseau représente 3 variables  $A$ ,  $B$  et  $C$  et l'indépendance entre  $A$  et  $C$ .

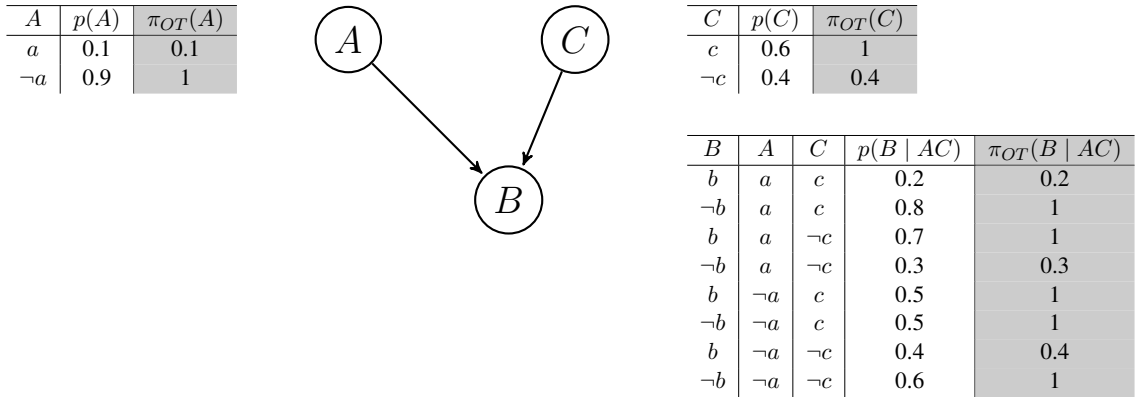


FIGURE 4.3 – Réseau Bayésien et la transformation des tables locales par  $OT$

### 4.3.2 Préservation des distributions jointes

Nous noterons la distribution jointe transformée par  $TR$ ,  $\pi_{RB-TR}$  et la distribution jointe obtenue par la règle de chaînage dans le cadre quantitatif (resp. dans le cadre qualitatif),  $\pi_{RP-*}$  (resp.  $\pi_{RP-min}$ ).

A l'aide des règles de chaînage vues dans les deux cadres, nous pouvons calculer les tables jointes, en reprenant le réseau de la Figure 4.3, calculons dans la Table 4.17, les distributions jointes. Si nous voulons calculer les indépendances, nous trouverons bien que  $A$  et  $C$  sont indépendants.

Le problème est que l'ordre des interprétations est modifié. Par exemple, nous avons  $\pi_{RB-OT}(a\neg b\neg c) = \pi_{RB-OT}(abc)$  alors que dans le distribution  $\pi_{RP-*}$  calculée à partir de la règle de chaînage 1.20,  $\pi_{RP-*}(a\neg b\neg c) < \pi_{RP-*}(abc)$ . Dans le cadre qualitatif, nous avons  $\pi_{RB-OT}(a\neg b\neg c) > \pi_{RB-OT}(abc)$  alors  $\pi_{RP-min}(a\neg b\neg c) = \pi_{RP-min}(abc)$ .

Nous nous intéressons donc à la différence entre l'ordre induit par la règle de chaînage probabiliste et l'ordre induit par la transformation puis le chaînage possibiliste. Et à partir de maintenant, nous séparons les deux cadres possibilistes, nous nous attaquons au cadre qualitatif en premier.

$A$	$B$	$C$	$p(A, B, C)$	$\pi_{RB-OT}(A, B, C)$	$\pi_{RP-*}(A, B, C)$	$\pi_{RP-min}(A, B, C)$
$\neg a$	$\neg b$	$\neg c$	0.216	0.46	0.4	0.4
$\neg a$	$\neg b$	$c$	0.27	1	1	1
$\neg a$	$b$	$\neg c$	0.144	0.244	0.16	0.4
$\neg a$	$b$	$c$	0.27	1	1	1
$a$	$\neg b$	$\neg c$	0.012	0.024	0.012	0.1
$a$	$\neg b$	$c$	0.048	0.1	0.1	0.1
$a$	$b$	$\neg c$	0.028	0.052	0.04	0.1
$a$	$b$	$c$	0.012	0.024	0.02	0.1

TABLE 4.17 – Comparaison des distributions de possibilités générées par le Réseau 4.3

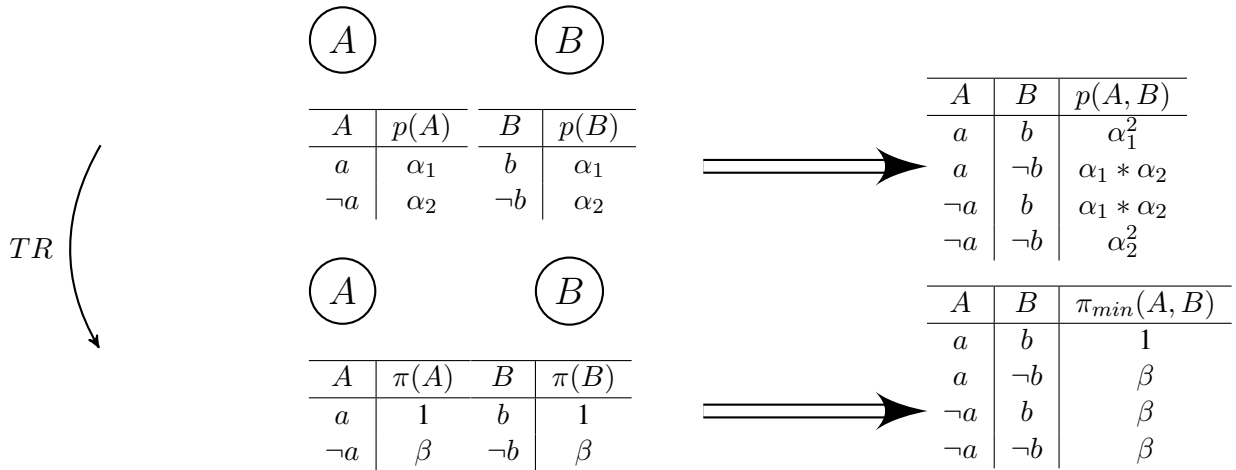
### 4.3.3 Préservation des distributions jointes dans le cadre qualitatif

Dans un premier temps, nous construisons la distribution de probabilités jointe à l'aide de la règle de chaînage, pour la transformer par  $TR$ , nous obtenons donc un ordre, puis après avoir transformé les tables locales du réseau, et calculer avec la règle de chaînage possibiliste dans le cadre qualitatif, la distribution jointe, nous obtenons un second ordre.

**Proposition 4.9.** Soient  $\pi_{RB-TR}$  la distribution de possibilités transformée par  $TR$  et  $\pi_{RP-min}$  la distribution jointe calculée par la règle de chaînage qualitative :

$$\forall \omega_1, \omega_2 \in \Omega, \pi_{RB-TR}(\omega_1) < \pi_{RB-TR}(\omega_2) \Rightarrow \pi_{RP-min}(\omega_1) < \pi_{RP-min}(\omega_2) \quad (4.10)$$

*Démonstration.* Soit un réseau Bayésien avec 2 variables  $A$  et  $B$ . Ici,  $A \perp B$  et  $A$  et  $B$  ont la même distribution (ici,  $\alpha_1 > \alpha_2$ ).


 FIGURE 4.4 – Représentation de la transformation  $TR$  sur un réseau

Clairement,  $p(ab) > p(a\neg b) = p(\neg ab) > p(\neg a\neg b)$ . Et en possibilités,  $\pi(ab) > \pi(a\neg b) = \pi(\neg ab) = \pi(\neg a\neg b)$ .

Dans cet exemple, quelle que soit la transformation, si elle préserve l'ordre, nous perdons l'ordre après transformation du réseau.  $\square$

La preuve démontre que l'ordre n'est pas préservé strictement, la question maintenant est de savoir

si l'ordre des interprétations peut s'inverser. Étant donné un ordre sur la distribution  $\pi_{RB-TR}$  tel que  $\pi_{RB-TR}(\omega_i) > \pi_{RB-TR}(\omega_j)$ , pouvons-nous avoir  $\pi_{RB-TR}(\omega_i) < \pi_{RB-TR}(\omega_j)$  ?

Pour répondre à cette question, nous allons d'abord nous intéresser aux transformations existantes. Par exemple, prenons la transformation  $ST$ , le Graphe 4.5 représente un réseau Bayésien et sa transformation par  $ST$ , puis les deux distributions jointes générées depuis le réseau (Table 4.18).

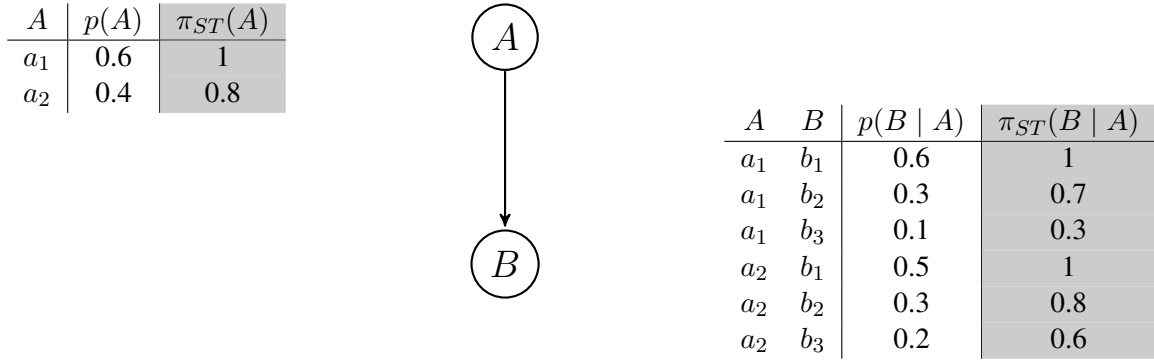


FIGURE 4.5 – Réseau Bayésien avec des distributions locales et leurs transformées par  $ST$

A	B	$p(A, B)$	$\pi_{RP-min}(A, B)$
$a_1$	$b_1$	0.36	1
$a_1$	$b_2$	0.18	0.7
$a_1$	$b_3$	0.06	0.3
$a_2$	$b_1$	0.2	0.8
$a_2$	$b_2$	0.12	0.8
$a_2$	$b_3$	0.08	0.6

TABLE 4.18 – Distributions de probabilités et possibilités jointes codées par les réseaux de la Figure 4.5

Afin de comparer nos deux tables de possibilités, nous transformons la distribution de probabilités jointe de la Table 4.18 par  $ST$ , on obtient la Table 4.19.

A	B	$\pi_{RB-ST}(A, B)$
$a_1$	$b_1$	1
$a_1$	$b_2$	0.8
$a_1$	$b_3$	0.36
$a_2$	$b_1$	0.84
$a_2$	$b_2$	0.62
$a_2$	$b_3$	0.46

TABLE 4.19 – Distribution de possibilités transformée par  $ST$

De prime abord, nous remarquons que les deux distributions ne sont pas identiques. Et non seulement, sont elles différentes mais en plus l'ordre des interprétations est totalement faussé. Si nous regardons de plus près aux interprétations  $a_1b_2$  et  $a_2b_2$ , dans la Table 4.18, nous avons  $\pi_{RP-min}(a_1b_2) < \pi_{RP-min}(a_2b_2)$  alors que dans la Table 4.19, nous avons  $\pi_{RB-ST}(a_1b_2) > \pi_{RB-ST}(a_2b_2)$ . L'ordre de ces deux interprétations n'est pas identique. Nous observons ce phénomène avec  $VT$  aussi pour le même exemple.

Maintenant, la question est de savoir si oui ou non, il existe une transformation dans laquelle nous serions sûr que l'ordre est respecté strictement.

**Proposition 4.10.** Soient  $\pi_{RB-TR}$  la distribution de possibilités transformée par  $TR$  et  $\pi_{RP-min}$  la distribution jointe calculée par la règle de chaînage qualitative :

$$\forall \omega_1, \omega_2 \in \Omega, \pi_{RP-min}(\omega_1) < \pi_{RP-min}(\omega_2) \not\Rightarrow \pi_{RB-TR}(\omega_1) < \pi_{RB-TR}(\omega_2) \quad (4.11)$$

*Démonstration.* Soit le réseau Bayésien représenté par la Figure 4.6, ici la table  $p(A)$  est une permutation de la table  $p(B)$ . Comme le rappelle la proposition, la transformation  $TR$  par principe donne donc la même possibilité aux mondes. Étant donné que la transformation  $TR$  préserve l'ordre des interprétations,

<div style="display: flex; justify-content: center; align-items: center; gap: 20px;"> <div style="text-align: center;"> <div style="border: 1px solid black; border-radius: 50%; width: 30px; height: 30px; margin: 0 auto; display: flex; align-items: center; justify-content: center;">A</div> <table border="1" style="margin-top: 10px; border-collapse: collapse; text-align: left;"> <thead> <tr> <th style="width: 10%;">A</th> <th style="width: 15%;">p(A)</th> <th style="width: 15%;">π<sub>TR</sub>(A)</th> </tr> </thead> <tbody> <tr> <td>a<sub>1</sub></td> <td>0.4</td> <td>1</td> </tr> <tr> <td>a<sub>2</sub></td> <td>0.2</td> <td>α<sub>2</sub></td> </tr> <tr> <td>a<sub>3</sub></td> <td>0.25</td> <td>α<sub>1</sub></td> </tr> <tr> <td>a<sub>4</sub></td> <td>0.15</td> <td>α<sub>3</sub></td> </tr> </tbody> </table> </div> <div style="text-align: center;"> <div style="border: 1px solid black; border-radius: 50%; width: 30px; height: 30px; margin: 0 auto; display: flex; align-items: center; justify-content: center;">B</div> <table border="1" style="margin-top: 10px; border-collapse: collapse; text-align: left;"> <thead> <tr> <th style="width: 10%;">B</th> <th style="width: 15%;">p(B)</th> <th style="width: 15%;">π<sub>TR</sub>(B)</th> </tr> </thead> <tbody> <tr> <td>b<sub>1</sub></td> <td>0.15</td> <td>α<sub>3</sub></td> </tr> <tr> <td>b<sub>2</sub></td> <td>0.2</td> <td>α<sub>2</sub></td> </tr> <tr> <td>b<sub>3</sub></td> <td>0.25</td> <td>α<sub>1</sub></td> </tr> <tr> <td>b<sub>4</sub></td> <td>0.4</td> <td>1</td> </tr> </tbody> </table> </div> </div>	A	p(A)	π <sub>TR</sub> (A)	a <sub>1</sub>	0.4	1	a <sub>2</sub>	0.2	α <sub>2</sub>	a <sub>3</sub>	0.25	α <sub>1</sub>	a <sub>4</sub>	0.15	α <sub>3</sub>	B	p(B)	π <sub>TR</sub> (B)	b <sub>1</sub>	0.15	α <sub>3</sub>	b <sub>2</sub>	0.2	α <sub>2</sub>	b <sub>3</sub>	0.25	α <sub>1</sub>	b <sub>4</sub>	0.4	1	
A	p(A)	π <sub>TR</sub> (A)																													
a <sub>1</sub>	0.4	1																													
a <sub>2</sub>	0.2	α <sub>2</sub>																													
a <sub>3</sub>	0.25	α <sub>1</sub>																													
a <sub>4</sub>	0.15	α <sub>3</sub>																													
B	p(B)	π <sub>TR</sub> (B)																													
b <sub>1</sub>	0.15	α <sub>3</sub>																													
b <sub>2</sub>	0.2	α <sub>2</sub>																													
b <sub>3</sub>	0.25	α <sub>1</sub>																													
b <sub>4</sub>	0.4	1																													

FIGURE 4.6 – Réseau Bayésien dont les distributions locales sont transformées par  $TR$

nous avons  $\alpha_1 > \alpha_2 > \alpha_3$ .

Observons quelques interprétations de la distribution jointe en probabilités et en possibilités.  
En probabilités :

- $p(a_1b_1) = 0.4 * 0.15 = 0.06$
- $p(a_2b_2) = 0.2 * 0.2 = 0.04$

$$\Rightarrow p(a_1b_1) > p(a_2b_2) \quad (*_1)$$

En possibilités :

- $\pi_{RP-min}(a_1b_1) = \alpha_3$
- $\pi_{RP-min}(a_2b_2) = \alpha_2$

$$\Rightarrow \pi_{RP-min}(a_1b_1) < \pi_{RP-min}(a_2b_2) \quad (*_2)$$

De  $(*_1)$  et  $(*_2)$  nous pouvons conclure que l'ordre s'inversera dans le cadre qualitatif quelle que soit la transformation  $TR$ . □

*Avons-nous le même constat dans le cadre quantitatif ?*

#### 4.3.4 Préservation des distributions jointes dans le cadre quantitatif

Dans le cadre quantitatif, nous montrons que les transformations définies dans le chapitre 2, peuvent comme dans le cadre qualitatif, inverser les interprétations lors de la transformations du réseau Bayésien en réseau possibiliste. Par exemple avec la transformation  $OT$ . Reprenons le réseau 4.5, et adaptons le avec la transformation  $OT$ .

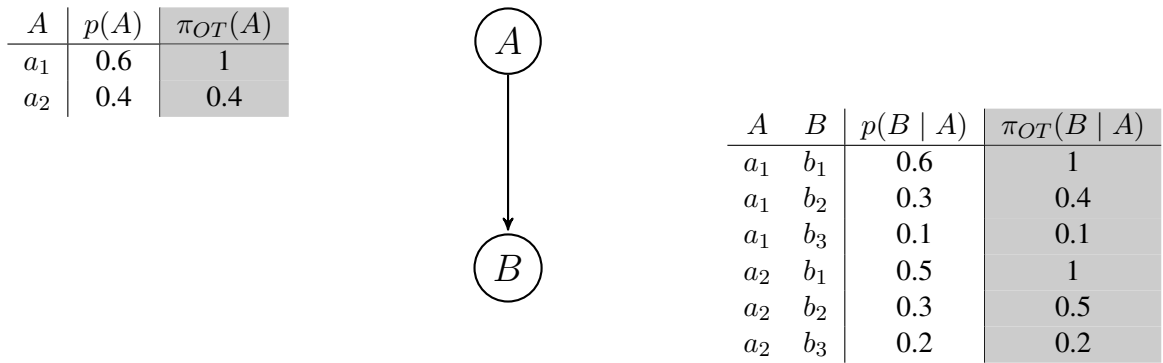


FIGURE 4.7 – Réseau Bayésien avec des distributions locales et leurs transformées par  $OT$

$A$	$B$	$p(A, B)$	$\pi_{RP-*}(A, B)$
$a_1$	$b_1$	0.36	1
$a_1$	$b_2$	0.18	0.4
$a_1$	$b_3$	0.06	0.1
$a_2$	$b_1$	0.2	0.4
$a_2$	$b_2$	0.12	0.2
$a_2$	$b_3$	0.08	0.08

TABLE 4.20 – Distributions de probabilités et possibilités jointes codées par le réseau de la Figure 4.7

$A$	$B$	$\pi_{RB-OT}(A, B)$
$a_1$	$b_1$	1
$a_1$	$b_2$	0.44
$a_1$	$b_3$	0.06
$a_2$	$b_1$	0.64
$a_2$	$b_2$	0.26
$a_2$	$b_3$	0.14

TABLE 4.21 – Distribution de possibilités transformée par  $OT$

Afin de comparer nos deux tables de possibilités, nous transformons la distribution de probabilités jointe de la Table 4.18 par  $OT$ , on obtient la Table 4.19.

Comme dans le cadre qualitatif, nous remarquons que l'ordre des interprétations peut s'inverser. Si nous regardons de plus près aux interprétations  $a_2b_3$  et  $a_1b_3$ , dans la Table 4.20, nous avons  $\pi_{RP-*}(a_2b_3) < \pi_{RP-*}(a_1b_3)$  alors que dans la Table 4.21, nous avons  $\pi_{RB-OT}(a_2b_3) > \pi_{RB-OT}(a_1b_3)$ .

### Conclusion

Dans ce chapitre, nous avons voulu regarder les transformations dans les modèles graphiques. Nous nous sommes rendu compte qu'avec les transformations existantes, nous perdions l'ordre des interprétations dans la distribution jointe. Nous avons aussi prouvé que peu importe la transformation, dans le cadre qualitatif, elle ne garantit pas la préservation de l'ordre des interprétations. Cela en plus de nous donner des distributions différentes.



# Conclusion et Perspectives

## Conclusion

Ce stage avait pour but d'étudier les transformations probabilistes-possibilistes. Dans cette optique, deux cadres de représentation de l'incertain ont été étudiés, la théorie des probabilités et la théorie des possibilités, afin de mieux appréhender le sujet et comprendre l'intérêt des transformations. De nombreux auteurs ont étudié cette question dans le passé, nous donnant une base solide d'étude. Notre état de l'art comprend ainsi l'étude des transformations existantes, la plus connue et la plus utilisée étant *OT* (*Optimal Transformation*) car jusqu'à présent elle donne la meilleure spécificité. Dans la première partie de la contribution, nous avons traité les distributions en appliquant une transformation puis une règle d'inférence et l'inverse afin de comparer les distributions de possibilités obtenues. Dans la seconde partie, nous nous sommes intéressés aux transformations dans les modèles graphiques.

Concernant les résultats, nous avons remarqué que les transformations souffrent de beaucoup de problèmes liés à l'inférence. La première question que nous nous sommes posée était sur l'inférence MPE (*Most Probable Explanation*), et nous avons remarqué que les requêtes MPE en probabilités et en possibilités après transformation donnent les mêmes résultats. En s'intéressant à la règle de marginalisation, passer d'un formalisme à un autre, aucune transformation ne nous permet de conserver l'ordre des interprétations. Ce phénomène est lié à la préservation de l'ordre des événements, nous avons montré que lors d'une transformation, l'ordre des événements n'était pas préservé. Quand nous révisons une distribution, nous distinguons deux cas, dans le premier cas lorsqu'on révisé avec des informations certaines, les distributions obtenues sont différentes mais on constate que l'ordre des interprétations est préservé. Dans le deuxième cas, nous révisons avec des informations incertaines, les interprétations dans ce cas ne sont plus ordonnées de la même manière. Ces deux premiers résultats nous ont permis de déduire que l'inférence MAP (*Maximum A Posteriori*) ne donnera pas les mêmes résultats.

Nous avons étudié les indépendances de variables dans les distributions, les résultats montrent que les indépendances, dans le cas général, ne sont pas préservées. Ce qui fait des modèles graphiques l'outil idéal quant à la préservation des indépendances, en plus de leurs formes compactes.

Dans les modèles graphiques, nous avons analysé la différence entre la distribution jointe obtenue par transformation et la distribution jointe obtenue à partir du réseau transformé. Nous avons constaté que l'ordre des interprétations n'était pas toujours le même. On distingue deux cas, dans le cadre qualitatif, nous avons prouvé que pour n'importe quelle transformation, non seulement l'ordre strict est perdu, mais il existe des cas où l'ordre peut s'inverser. Dans le cadre quantitatif, nous avons montré que les transformations existantes ne préservaient pas l'ordre, et pouvaient aussi l'inverser.

Nous avons étudié les transformations dans deux cas, le premier face aux distributions et le second dans les modèles graphiques. Il ressort de ce travail que transformer les distributions nous permet de garantir l'ordre des interprétations alors que dans les modèles graphiques nous le perdons. Cependant,

l'avantage des modèles graphiques c'est qu'ils nous permettent de conserver les relations d'indépendances ce qui n'est pas le cas des distributions.

Nous avons aussi étudié les transformations dans les deux cadres possibilistes, c'est à dire, le cadre qualitatif et quantitatif. On constate qu'aucune transformation lorsqu'on est dans le cadre qualitatif ne préserve l'ordre des interprétations quand on utilise avec les règles d'inférence, ou même dans les modèles graphiques. Dans le cadre quantitatif, les transformations existantes nous donnent les mêmes résultats que dans le cadre qualitatif, à savoir, quand on applique une règle d'inférence puis on transforme et l'inverse, l'ordre des interprétations n'est pas garanti. Nous avons laissé le cas général comme perspective.

## Perspectives

Pour la suite, une analyse approfondie dans le cadre quantitatif est nécessaire. À savoir, trouver si c'est possible qu'une transformation préserve l'ordre des interprétations dans les modèles graphiques.

Il faut noter que toute notre analyse et contribution s'est cantonnée aux transformations des probabilités vers les possibilités. Avec le temps, notre attention se serait portée sur les transformations dans l'autre sens, c'est à dire des possibilités vers les probabilités. Comme nous l'avons fait remarquer dans ce rapport, les outils d'inférence probabiliste sont très nombreux, alors une transformation qui garantirait des résultats similaires dans les deux cadres nous permettrait d'utiliser ces outils.

Nous pouvons bien sûr nous attaquer à d'autres formalismes, par exemple trouver des transformations entre les probabilités imprécises et les possibilités. En effet, on peut voir une distribution de probabilités imprécises comme une famille de distributions de probabilités. Nous pouvons aussi analyser d'autres types d'informations (ex. préférences) exprimées dans différents cadres.

Une autre perspective concerne le traitement d'informations multi-sources exprimées dans des cadres différents. Il est intéressant de voir, par exemple, comment fusionner plusieurs distributions différentes.

Nous pourrions nous intéresser aussi aux transformations vers des bases de croyances comme le montre la Figure 8.

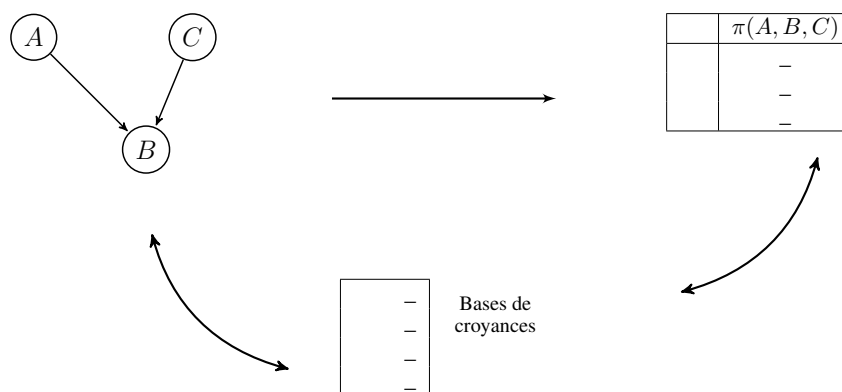


FIGURE 8 – Transformations, modèles graphiques et bases de croyances

Il serait intéressant de voir s'il existe des transformations pouvant passer du cadre probabiliste au du cadre possibiliste quantitatif ou qualitatif, ou inversement en passant par des bases de croyances qui sont une autre forme de représentation compacte. En effet, si une transformation se trouve avoir de bonnes

---

propriétés (ex. efficacité de l'inférence) dans un cadre, il serait intéressant de pouvoir transformer dans l'autre cadre.

# Bibliographie

- [BGK02] C. BORGELT, J. GEBHARDT, et R. KRUSE. « Graphical Models ». Dans *In Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, pages 51–68. Wiley, 2002. [1.3.2](#)
- [Bou07] M.S BOUGUELID. Contribution à l'application de la reconnaissance des formes et la théorie des possibilités au diagnostic adaptatif et prédictif des systèmes dynamiques. Université de Champagne-Ardenne, Decembre 2007. [2.3.4](#)
- [CD05] H. CHAN et A. DARWICHE. « On the revision of probabilistic beliefs using uncertain evidence. ». *Artif. Intell.*, 163(1) :67–90, 2005. [3.6.1](#)
- [Coo90] G. F. COOPER. « The computational complexity of probabilistic inference using bayesian belief networks ». *Artificial Intelligence*, 42(2-3) :393–405, mar 1990. [1.2.3](#)
- [Coz01] F G. COZMAN. « JavaBayes, Bayesian Networks in Java ». <http://www.cs.cmu.edu/~javabayes/Home/>, 1998 - 2001. [2.1.2](#)
- [Dar09] A. DARWICHE. *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press, 2009. [1.2.3](#)
- [dC11] C. P. de CAMPOS. « New Complexity Results for MAP in Bayesian Networks. ». Dans Toby WALSH, éditeur, *IJCAI*, pages 2100–2106. IJCAI/AAAI, 2011. [1.2.3](#)
- [DFMP04] D. DUBOIS, L. FOULLOY, G. MAURIS, et H. PRADE. « Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities ». *Reliable Computing*, 10(4) :273–297, 2004. [2.2.3](#), [2.3.2](#)
- [DP83] D. DUBOIS et H. PRADE. « Unfair coins and necessity measures : a possibilistic interpretation of histograms ». *Fuzzy Sets and Systems*, 10(1) :15–20, 1983. DP034. [2.3.7](#)
- [DP88] D. DUBOIS et H. PRADE. *Possibility Theory : An Approach to Computerized Processing of Uncertainty (traduction revue et augmentée de Théorie des Possibilités)*. Plenum Press, New York, 1988. [1.3](#)
- [DP97] D. DUBOIS et H. PRADE. « A synthetic view of belief revision with uncertain inputs in the framework of possibility theory. ». *Int. J. Approx. Reasoning*, 17(2-3) :295–324, 1997. [3.6.2](#), [3.6.2](#), [3.6.2](#)
- [DPS93] D. DUBOIS, H. PRADE, et S. SANDRI. On Possibility/Probability Transformations. Dans R. LOWEN et M. ROUBENS, éditeurs, *Fuzzy Logic*, pages 103–112. Kluwer Academic Publishers, Dordrecht, 1993. [2.3.2](#), [2.3.3](#), [2.3.7](#)

- 
- [GVP90] D. GEIGER, T. VERMA, et J. PEARL. « Identifying independence in bayesian networks. ». *Networks*, 20(5) :507–534, 1990. [1.2.3](#)
- [His78] E. HISDAL. « Conditional possibilities independence and non interaction ». *Fuzzy Sets and Systems*, pages 283–297, 1978. [1.3.1](#)
- [Jef65] R. JEFFREY. « The Logic of Decision ». 1965. [3.6.1](#)
- [Jen96] F. V. JENSEN. *Introduction to Bayesian Networks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st édition, 1996. [2](#)
- [KG93] G. J. KLIR et J. F. GEER. Information-Preserving Probability- Possibility Transformations : Recent Developments. Dans R. LOWEN et M. ROUBENS, éditeurs, *Fuzzy Logic*, pages 417–428. Kluwer Academic Publishers, Dordrecht, 1993. [2.2.2](#)
- [MD06] M.-H. MASSON et T. DENOEU. « Inferring a possibility distribution from empirical data. ». *Fuzzy Sets and Systems*, 157(3) :319–340, 2006. [2.1.1](#)
- [MMR06] P. Billaudel M.S MOUCHAWEH, M.S Bouguelid et B. RIERA. Variable probability-possibility transformation. pages 417–428. September 2006. [2.3.4](#)
- [PD80] H. PRADE et D. DUBOIS. *Fuzzy Sets and Systems*, pages 258–259, 1980. [2.2.3](#)
- [Pea88] J. PEARL. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988. [3](#)
- [Pea89] J. PEARL. *Probabilistic reasoning in intelligent systems - networks of plausible inference*. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann, 1989. [1.2.3](#)
- [PP86] J. PEARL et A. PAZ. « Graphoids : Graph-Based Logic for Reasoning about Relevance Relations or When would x tell you more about y if you already know z ? ». Dans *ECAI*, pages 357–363, 1986. [1.2.3](#)
- [SAA13] Y. B. SLIMEN, R. AYACHI, et N. B. AMOR. « Probability-Possibility Transformation : - Application to Bayesian and Possibilistic Networks. ». Dans Francesco MASULLI, Gabriella PASI, et Ronald R. YAGER, éditeurs, *WILF*, volume 8256 de *Lecture Notes in Computer Science*, pages 122–130. Springer, 2013. [2.3](#)
- [Sme89] P. SMETS. « Constructing the Pignistic Probability Function in a Context of Uncertainty ». Dans *UAI*, pages 29–40, 1989. [2.1.2](#)
- [Sud92] T. SUDKAMP. « On probability-possibility transformations ». *Fuzzy Sets and Systems*, pages 73–81, 1992. [3.5](#), [3.7](#)
- [Yam01] K. YAMADA. « Probability-Possibility Transformation Based On Evidence Theory ». pages 70–75, 2001. [2.3.7](#)
- [Zad99] L. A. ZADEH. « Fuzzy sets as a basis for a theory of possibility ». *Fuzzy Sets and Systems*, 100 :9–34, 1999. [1.3](#), [2.2.1](#)
- [ZP96] N. L. ZHANG et D. L. POOLE. « Exploiting Causal Independence in Bayesian Network Inference. ». *J. Artif. Intell. Res. (JAIR)*, 5 :301–328, 1996. [1](#)