

# Extension pondérée des logiques modales dans le cadre des croyances graduelles

Benedicte Legastelois

► **To cite this version:**

Benedicte Legastelois. Extension pondérée des logiques modales dans le cadre des croyances graduelles. Logique en informatique [cs.LO]. Université Pierre et Marie Curie - Paris VI, 2017. Français. <NNT : 2017PA066516>. <tel-01714318v2>

**HAL Id: tel-01714318**

**<https://hal.sorbonne-universite.fr/tel-01714318v2>**

Submitted on 26 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THÈSE DE DOCTORAT  
DE L'UNIVERSITÉ PIERRE ET MARIE CURIE**

Spécialité Informatique

Présentée par

**Bénédicte Legastelois**

**Extension pondérée des logiques modales  
dans le cadre des croyances graduelles**

Soutenance le 30 novembre 2017

devant le jury composé de :

Lluis GODO (IIIA-CSIC, Barcelone)	Rapporteur
Andreas HERZIG (IRIT, Toulouse)	Rapporteur
Philippe CAPET (Ektimo, Cahors)	Examineur
Nicolas MAUDET (LIP6, Paris)	Examineur
Marie-Jeanne LESOT (LIP6, Paris)	Directrice de thèse
Adrien REVAULT D'ALLONNES (LIASD, Paris 8)	Encadrant



# Table des matières

Introduction générale	1
<b>Partie I Logiques modales pondérées</b>	<b>5</b>
<b>Introduction</b>	<b>7</b>
<b>1 État de l'art</b>	<b>9</b>
1.1 Cadre général des logiques modales . . . . .	9
1.2 Axes de comparaison . . . . .	14
1.3 Préservation de la sémantique de Kripke . . . . .	17
1.4 Extensions de la sémantique de Kripke . . . . .	22
1.5 Bilan . . . . .	24
<b>2 Proposition d'une sémantique pondérée</b>	<b>27</b>
2.1 Syntaxe . . . . .	28
2.2 Définition de la sémantique proposée . . . . .	28
2.3 Interprétation en terme de quantification . . . . .	30
2.4 Propriétés . . . . .	32
2.5 Bilan . . . . .	35
<b>3 Proposition d'axiomes pondérés et correspondance</b>	<b>37</b>
3.1 Problématique . . . . .	38
3.2 Principes proposés . . . . .	39
3.3 Typologie des axiomes pondérés établis . . . . .	43
3.4 Type I . . . . .	44
3.5 Type II . . . . .	50
3.6 Type III . . . . .	54
3.7 Type IV . . . . .	59
3.8 Discussion . . . . .	63
3.9 Bilan . . . . .	64

<b>Partie II</b>	<b>Logique doxastique graduelle</b>	<b>65</b>
	<b>Introduction</b>	<b>67</b>
<b>4</b>	<b>État de l'art : formalisation de la croyance</b>	<b>69</b>
4.1	Aperçu épistémologique . . . . .	69
4.2	Modélisation de croyances binaires : formalisations logiques . . . . .	71
4.3	Modélisation de croyances non binaires : formalisations numériques . . . . .	73
4.4	Bilan . . . . .	75
<b>5</b>	<b>Représentation de croyances graduelles : formalisme flou</b>	<b>77</b>
5.1	Motivations et principes . . . . .	78
5.2	Proposition d'un modèle ensembliste flou . . . . .	80
5.3	Lecture doxastique des propriétés floues . . . . .	81
5.4	Bilan . . . . .	83
<b>6</b>	<b>Manipulation des croyances graduelles dans le formalisme flou</b>	<b>85</b>
6.1	Conjonction et disjonction . . . . .	86
6.2	Implication . . . . .	91
6.3	Introspection . . . . .	93
6.4	Bilan . . . . .	94
<b>7</b>	<b>Négation de croyances graduelles</b>	<b>97</b>
7.1	Formalisation de la négation . . . . .	97
7.2	Négation de la formule . . . . .	98
7.3	Négation du degré . . . . .	100
7.4	Négation de la modalité . . . . .	102
7.5	Bilan . . . . .	105
<b>Partie III</b>	<b>Applications</b>	<b>107</b>
	<b>Introduction</b>	<b>109</b>
<b>8</b>	<b>Outil de vérification de formules modales pondérées</b>	<b>111</b>
8.1	Présentation générale . . . . .	111
8.2	Génération et caractérisation de modèles . . . . .	114
8.3	Vérification de formules . . . . .	115
8.4	Bilan . . . . .	118
<b>9</b>	<b>Application à un jeu coopératif</b>	<b>119</b>
9.1	Présentation du jeu Hanabi . . . . .	120
9.2	Formalisation en logique modale et logique modale pondérée . . . . .	121
9.3	Prise de décision en logique modale . . . . .	123
9.4	Expérimentation . . . . .	124
9.5	Bilan . . . . .	127

---

<b>Conclusions et perspectives</b>	<b>129</b>
<b>Bibliographie</b>	<b>135</b>
<b>Annexes</b>	<b>141</b>
<b>A Implémentation du jeu Hanabi</b>	<b>141</b>
<b>B Rappels de manipulation conjonctive et disjonctive</b>	<b>145</b>
B.1 Principes : factorisation et distribution . . . . .	145
B.2 Logique classique . . . . .	146
B.3 Logique modale . . . . .	146
B.4 Logique floue . . . . .	147
B.5 Logique possibiliste . . . . .	148



# Introduction générale

## Contexte et motivation

Plusieurs cryptozoologues se retrouvent lors d'un congrès et discutent de leurs travaux respectifs, leurs enquêtes sur l'existence de créatures mythologiques, imaginaires ou simplement disparues, basées sur le croisement de diverses sources, orales et écrites. Trois d'entre eux étudient les espèces animales connues par tradition ou mémoire collective, notamment sous forme de mythes et légendes. Le premier déclare qu'il croit dur comme fer à l'existence, passée ou présente, des licornes. Le second affirme qu'il ne croit que très peu à leur existence et le dernier qu'il n'y croit pas du tout.

Deux points sont à souligner, discutés tour à tour ci-dessous : la part subjective de leur conviction et la nécessité de moduler leurs croyances par des degrés.

**Subjectivité des croyances** La composante subjective de la croyance est cruciale car les croyances constituent une notion essentiellement non factuelle. Il n'existe aucune preuve irréfutable de l'existence des licornes, la principale part d'influence sur les convictions des cryptozoologues provient de leur perception des choses. Les biais internes qui contribuent à modeler une description très personnelle de la réalité sont fortement liés aux convictions profondes de chacun : dans le cas des cryptozoologues, s'ils trouvent plusieurs témoignages écrits de l'apparition d'un ongulé muni d'une corne sur la tête, ils peuvent, selon leur interprétation, considérer ces observations comme preuves supplémentaires de l'existence des licornes ou douter de la crédibilité des témoins.

La croyance est un concept qui fait l'objet d'études dans plusieurs domaines, tant en sciences cognitives, qu'en philosophie de l'esprit et épistémologie et en logique. D'un point de vue cognitif, la croyance est un comportement essentiel pour l'être humain : le fait de croire permet de se faire une opinion et d'améliorer son propre comportement social (Schwitzgebel, 2015).

D'un point de vue philosophique : le domaine de l'épistémologie étudie notamment tous les aspects de la croyance, de son fonctionnement aux mécanismes de raisonnement qui lui sont rattachés (Huber, 2016). Ce type de travaux est à mettre en relation avec les sciences

cognitives, qui permettent l'étude empirique de la croyance. En adoptant ce point de vue expérimental, il est possible d'observer les comportements liés aux croyances et de mettre en lumière les liens entre l'aspect théorique et l'aspect pratique des capacités humaines concernant la croyance.

D'autre part, la modélisation théorique de la croyance passe par une formalisation logique : elle permet de poser un cadre de définition et de raisonnement, respectant des propriétés choisies et mettant en œuvre les arguments philosophiques proposés en amont. Dans ce cadre, la logique modale apporte un ensemble d'outils formels pour représenter et raisonner sur des notions non factuelles comme la croyance (Hintikka, 1962).

**Gradualité des croyances** Lorsque les cryptozoologues expriment chacun leur tour leur croyance en l'existence des licornes, ils attribuent explicitement à cette croyance un indicateur visant à en moduler la valeur. Ainsi, les expressions linguistiques « dur comme fer » et « très peu » tempèrent le niveau de croyance. Toutes les nuances possibles entre une croyance très faible et une croyance extrêmement haute peuvent être exprimées verbalement. Nous parlons alors de pondération ou gradualisation de la croyance, formellement représentée par exemple par une valeur numérique, le *degré de croyance*. La manipulation de ces croyances pondérées nécessite d'étendre la logique modale afin d'en accroître l'expressivité et améliorer le réalisme du raisonnement ainsi modélisé.

Le pouvoir de représentation gagné par cet ajout de pondération est contre-balancé par une nécessité de formaliser ces notions graduelles de façon appropriée. En effet, s'il est aisé d'exprimer des croyances partielles dans le langage naturel, le passage au cas numérique n'est pas évident : les termes linguistiques qui permettent de moduler la croyance, comme « croire dur comme fer » ou « croire très peu », sont multiples et ne se combinent pas naturellement. Il est donc nécessaire de proposer une structure arithmétique pour représenter et combiner ces degrés, ainsi que d'examiner les conséquences qu'elle implique et les propriétés qu'elle induit sur leur interprétation. Plusieurs types de structures ont été proposées, basées par exemple sur les formalismes flou (Bou et al., 2009), possibiliste (Fariñas del Cerro & Herzig, 1991; Dubois et al., 2012), ou encore probabiliste (Bacchus et al., 1996).

## Travaux de thèse

Afin de traiter la problématique complexe qu'est la formalisation du raisonnement sur des croyances pondérées, nous proposons de travailler sur deux plans : d'une part, dans le cadre des logiques modales pondérées pour donner un point de vue général sur la façon dont il est possible d'introduire formellement un degré dans la notion de croyance ; d'autre

---

part, dans le cadre particulier de la logique des croyances, dont nous voulons définir une extension graduelle.

Ainsi, les travaux présentés dans cette thèse sont organisés autour de cet axe. D'abord, dans le cadre des logiques modales pondérées, nous proposons une étude comparative des modèles existants, selon une classification croisée de leurs formalismes. Nous proposons également une nouvelle sémantique pour l'interprétation de formules modales pondérées basée sur la sémantique des mondes possibles, qui repose sur un décompte relatif de mondes validant une formule. Enfin, nous réalisons une étude approfondie des axiomes de la logique modale pour les reconsidérer dans le cadre pondéré. Nous établissons un moyen de tenir compte de la modalité pondérée et en étudions les conséquences en terme de correspondance entre axiomes et modèles. De cette étude, nous dérivons une typologie des axiomes, distinguant quatre catégories selon les relations entre l'axiome pondéré, sa variante classique et la correspondance avec les contraintes sur les cadres de Kripke associées.

D'un autre côté, pour le cas particulier des croyances graduelles, nous proposons un modèle qui repose sur une interprétation ensembliste des croyances graduelles et qui se différencie de l'idée d'incertitude qui accompagne souvent l'interprétation des degrés de croyance : dans le cadre formel de la théorie des sous-ensembles flous (Zadeh, 1965), nous proposons de considérer l'ensemble de croyances comme un sous-ensemble flou de l'ensemble des formules bien formées. Nous étudions les conséquences de ce choix d'interprétation, notamment en analysant le comportement des propriétés du formalisme appliquées aux degrés de croyance. Nous dégageons ainsi les propriétés propres au modèle proposé pour représenter les croyances graduelles.

De plus, nous proposons un système axiomatique basé sur l'interprétation floue choisie. D'une part, nous analysons le comportement des degrés vis-à-vis des connecteurs logiques, ce qui permet de considérer les différentes manipulations et combinaisons de croyances graduelles, autour des opérateurs de conjonction et de disjonction, ainsi que les conséquences formelles de l'application d'une négation à une croyance pondérée. D'autre part, nous étudions l'extension des axiomes classiques de la logique doxastique dans le formalisme flou : l'intégration d'un degré pour moduler les croyances modifie l'expressivité et l'interprétabilité de ces axiomes et doit être adaptée au choix d'interprétation.

Enfin, nous complétons ces travaux par des mises en œuvre du modèle sémantique proposé dans un contexte pratique, pour le cas général et pour le cas doxastique. Le cas général est implémenté dans un outil d'interprétation de formules. Ce dernier est muni d'une interface graphique permettant de générer et d'afficher des modèles de Kripke contraints par des paramètres définis par l'utilisateur et d'entrer des formules modales pondérées pour la vérification. D'autre part, nous mettons en œuvre la lecture doxastique de cette sémantique dans le cadre d'un jeu de cartes coopératif à annonces publiques, dans lequel les joueurs

ignorent les cartes composant leur main. Il s'agit, pour chaque joueur, de tenir compte des annonces de ses coéquipiers pour deviner le contenu de sa main. L'utilisation de la sémantique pondérée intervient dans la prise de décision quant aux cartes à jouer.

### **Organisation du manuscrit**

Le manuscrit est organisé en trois parties, les deux premières présentent les travaux théoriques et la dernière en présente les applications pratiques.

La première partie détaille les travaux sur les logiques modales pondérées du point de vue général. Elle présente dans le chapitre 1 une étude comparative des modèles existants. Elle décrit dans le chapitre 2 notre proposition de sémantique pour la modalité pondérée. Elle en analyse les conséquences en terme d'axiomes modaux et ouvre la voie d'une théorie de la correspondance dans le cadre pondéré, dans le chapitre 3.

La deuxième partie du manuscrit se concentre sur le cas de la logique doxastique. Elle propose un état de l'art sur la représentation des croyances classiques et graduelles, dans le chapitre 4, et présente, dans le chapitre 5, notre proposition de modélisation floue pour des croyances graduelles, ainsi qu'une étude approfondie des propriétés de ce choix d'interprétation. Le chapitre 6 décrit les possibilités de manipulation des croyances graduelles autour des opérateurs logiques, qui découlent de l'interprétation choisie, avec une analyse des axiomes doxastiques dans le contexte du formalisme flou pour les croyances graduelles. Enfin, cette partie présente, dans le chapitre 7, une étude de l'application de la négation à une croyance graduelle.

La troisième partie présente les applications pratiques qui mettent en œuvre les travaux théoriques décrits dans le manuscrit. D'une part, elle détaille l'implémentation de l'outil de vérification de formules modales pondérées et de génération de modèles de Kripke. D'autre part, elle présente l'implémentation du jeu coopératif à annonces publiques dont la résolution se base sur la sémantique que nous proposons pour la modalité pondérée.

Enfin, le manuscrit se termine par une conclusion générale et les perspectives de travail ouvertes par thèse.

Première partie

Logiques modales pondérées



# Introduction

Les logiques modales constituent une extension de la logique classique, permettant de manipuler des éléments non factuels. En particulier, elles sont utilisées pour modéliser le raisonnement dit épistémique, sur les croyances et le savoir (Hintikka, 1962).

Dans cette première partie, nous nous intéressons au cadre général dans lequel il est possible d'étendre la logique modale au cas non binaire. En d'autres termes, nous proposons d'étudier les logiques modales pondérées, sans nous restreindre à un cas particulier, donc, par exemple, hors du cadre épistémique. Nous abordons ainsi tous les aspects de la logique modale classique et en proposons une déclinaison qui correspond au cas de non-binarité des notions représentées par la modalité.

Le premier chapitre de cette partie présente une étude complète et approfondie des modèles de logique modale pondérée existants. Nous proposons de les organiser selon deux axes croisés, selon l'objet sur lequel porte la pondération dans la sémantique des mondes possibles et selon la nature de cette pondération, qui dépend de la structure arithmétique dans laquelle elle est interprétée. D'autres modèles, hors de ces catégories, sont également présentés. Ils correspondent soit à des cas particuliers de pondération de la logique modale, soit à des contextes d'application spécifiques.

Le deuxième chapitre présente la syntaxe et la sémantique que nous proposons pour interpréter des formules comportant une modalité pondérée. Elle est basée sur un principe de comptage relatif qui préserve la structure des mondes possibles classique. Nous en étudions les conséquences en caractérisant la sémantique proposée pour en déterminer les principales propriétés.

Dans le troisième chapitre, nous proposons une étude des axiomes modaux avec intégration de modalités pondérées : si l'on remplace chaque modalité classique par une modalité pondérée, il est nécessaire de discuter et déterminer la valeur de chacun des poids, conformément à ce que nous considérons comme une valeur pertinente, également définie dans ce chapitre. Les conséquences de la pondération des axiomes sur leur correspondance sémantique sont aussi examinées. Nous établissons une typologie qui permet d'organiser les axiomes pondérés selon s'ils permettent de conserver la correspondance classique ou non.



# Chapitre 1

## État de l'art

L'enrichissement des logiques classiques par l'introduction de modalités permet de manipuler des notions non factuelles au-delà de la véracité des formules : par exemple, il est possible d'exprimer « je crois que  $\varphi$  » ou « il est nécessaire que  $\varphi$  », indépendamment du fait que  $\varphi$  soit vraie.

Les logiques modales augmentent ainsi l'expressivité de la logique classique en ajoutant deux opérateurs :  $\Box$  et  $\Diamond$ . D'un point de vue sémantique, ceux-ci permettent de construire des formules dont la valeur de vérité ne dépend donc pas seulement de leurs sous-formules.

Les logiques modales pondérées augmentent encore le pouvoir d'expression, en enrichissant les deux opérateurs modaux avec des degrés, entiers ou réels, de type  $\Box_\alpha$  et  $\Diamond_\alpha$ . Étant donné les multiples possibilités qu'offre l'ajout d'un tel poids, il existe une grande diversité de logiques modales pondérées, dont ce chapitre propose de donner une vue d'ensemble.

Après avoir brièvement rappelé le cadre général des logiques modales et introduit les notations utilisées dans tout le manuscrit, il présente dans la section 1.2 les différents axes que nous proposons et qui permettent de caractériser et comparer les différentes logiques modales pondérées. Il les décrit ensuite successivement, en les distinguant d'abord selon qu'elles préservent la sémantique des mondes possibles de Kripke, en section 1.3, et ou qu'elles en proposent une extension, en section 1.4.

### 1.1 Cadre général des logiques modales : quelques rappels et notations

Cette section contient des rappels de logique modale, suivant les trois éléments d'un système logique : syntaxe, sémantique et axiomes. Pour plus de détails, le lecteur peut par exemple se rapporter aux ouvrages de Blackburn et al. (2001), Chellas (1980) ou la page de l'Encyclopédie de Philosophie de Stanford (Garson, 2016).

### 1.1.1 Syntaxe

Les logiques modales sont une extension de la logique propositionnelle. Leur syntaxe est composée des quatre connecteurs classiques : « et » noté  $\wedge$ , « ou » noté  $\vee$ , « non » noté  $\neg$ , « implique » noté  $\rightarrow$ , ainsi que d'un opérateur modal  $\Box$  et son opérateur dual noté  $\Diamond$  défini comme  $\Diamond = \neg\Box\neg$ .

Le langage de la logique modale est défini ainsi, en utilisant une grammaire formelle dont l'alphabet  $\mathcal{P}$  est un ensemble de variables propositionnelles :

$$\begin{aligned}
 \mathcal{P} &::= p, q, \dots \\
 \mathcal{F} &::= \mathcal{P} \\
 &| \mathcal{F} \wedge \mathcal{F} \\
 &| \mathcal{F} \vee \mathcal{F} \\
 &| \neg\mathcal{F} \\
 &| \mathcal{F} \rightarrow \mathcal{F} \\
 &| \Box\mathcal{F} \\
 &| \Diamond\mathcal{F}
 \end{aligned}$$

On appelle formule de logique modale toute formule appartenant à ce langage.

### 1.1.2 Sémantique des mondes possibles

Il existe plusieurs sémantiques pour interpréter les formules de logique modale. Nous présentons ici la sémantique de Kripke (Kripke, 1963), ou sémantique des mondes possibles.

**Modèles** La sémantique des mondes possibles de la logique modale définit un modèle  $\mathcal{M}$  comme un triplet  $\langle W, R, s \rangle$  :  $W$ , appelé l'univers, est un ensemble dont les éléments sont appelés mondes possibles ;  $R$  est une relation binaire d'accessibilité entre ces mondes (on note  $wRw'$  si  $w'$  est accessible depuis  $w$ ) ;  $s : W \times \mathcal{P} \mapsto \{0, 1\}$  est une fonction de valuation qui associe à une variable propositionnelle et un monde une valeur de vérité. Un tel modèle est illustré sur la figure 1.1, représentant trois mondes par des rectangles, la relation  $R$  par des flèches et donnant la valeur de vérité de deux variables  $p$  et  $q$  pour chaque monde.

Le couple  $\langle W, R \rangle$  est appelé *cadre*. Un modèle  $\mathcal{M}$ , triplet  $\langle W, R, s \rangle$  peut être noté comme un couple  $\langle F, s \rangle$  où  $F$  est le cadre  $\langle W, R \rangle$  et  $s$  la fonction de valuation.

Étant donné un cadre  $\langle W, R \rangle$  et un monde  $w \in W$ , on note  $R_w$  l'ensemble des mondes accessibles depuis  $w$  :

$$R_w = \{w' \in W \mid wRw'\} \tag{1.1}$$

**Calcul de valeur de vérité** La conséquence sémantique  $\models$  est utilisée de manière usuelle dans ce document : «  $\mathcal{M}, w \models \varphi$  » est une extension de  $s$  qui signifie « dans le modèle  $\mathcal{M}$ ,

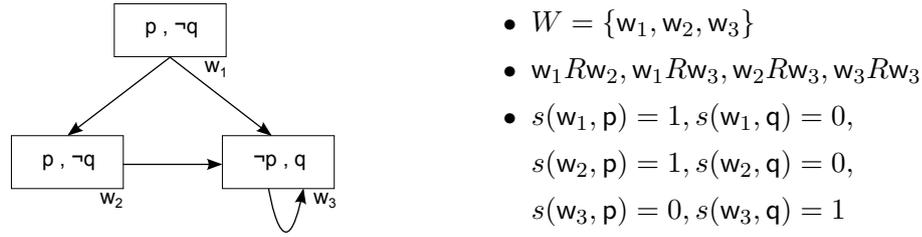


FIGURE 1.1 – Exemple de modèle dans la sémantique de Kripke

le monde  $w$  valide  $\varphi$  », ou «  $\varphi$  est vraie dans le monde  $w$  », selon le calcul de valeur de vérité rappelé ci-dessous. On note  $\mathcal{M}$  lorsqu'une formule est valide dans tous les mondes de  $\mathcal{M}$ . Nous utilisons également  $\vDash \varphi$  pour dire que  $\varphi$  est une tautologie : elle est vraie quel que soit le modèle  $\mathcal{M}$ . Enfin, dans la notation,  $\mathcal{M}$  est omis s'il n'y a pas d'ambiguïté et on s'autorise à écrire  $w \vDash \varphi$ .

La valeur de vérité d'une formule modale  $\varphi$  dépend du monde dans lequel on se place pour son interprétation. Si  $\varphi$  ne contient pas d'opérateur modal, on a  $\mathcal{M}, w \vDash \varphi$  si la formule est vraie dans  $w$  d'après les règles de la logique classique. Pour l'interprétation de la formule  $\Box\varphi$ , on a  $w \vDash \Box\varphi$  si et seulement si  $\varphi$  est vraie dans tous les mondes accessibles depuis  $w$ . Formellement :

$$\mathcal{M}, w \vDash \Box\varphi \Leftrightarrow \forall w' \in W, (w' \in R_w \Rightarrow \mathcal{M}, w' \vDash \varphi) \quad (1.2)$$

Par dualité, l'interprétation de la modalité  $\Diamond$  fait appel au quantificateur existentiel :

$$\mathcal{M}, w \vDash \Diamond\varphi \Leftrightarrow \exists w' \in W, (w' \in R_w \wedge w' \vDash \varphi) \quad (1.3)$$

**Mondes validant** On appelle « monde validant » (respectivement « monde invalidant ») un monde qui valide (respectivement invalide) une formule considérée. Par exemple, un monde  $w$  valide la formule  $\Box\varphi$  si tous ses mondes accessibles sont des mondes validants pour la formule  $\varphi$ .

Dans la suite, nous utilisons la notation  $R_w(\varphi)$  l'ensemble des mondes validant accessibles depuis  $w$ , pour une formule donnée  $\varphi$  :

$$R_w(\varphi) = \{w' \in R_w \mid \mathcal{M}, w' \vDash \varphi\} \quad (1.4)$$

### 1.1.3 Axiomatisation

Il existe plusieurs approches axiomatiques pour les logiques modales, la principale est basée sur un système formel qui étend le calcul de Hilbert, d'autres étendent par exemple le calcul des séquents (Poggioli, 2008). L'établissement d'axiomes modaux, ainsi que leur

Nom	Règle
Modus Ponens	$\frac{\vdash \varphi \rightarrow \psi \quad \vdash \varphi}{\vdash \psi}$
Nécessitation	$\frac{\vdash \varphi}{\vdash \Box \varphi}$

TABLEAU 1.1 – Règles d'inférence de la logique modale (Van Benthem, 1984), la conséquence syntaxique  $\vdash$  est utilisée classiquement.

Nom	Axiome
(K)	$\Box(\varphi \supset \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$
(D)	$\Box\varphi \rightarrow \Diamond\varphi$
(M)	$\Box\varphi \rightarrow \varphi$
(4)	$\Box\varphi \rightarrow \Box\Box\varphi$
(5)	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$

TABLEAU 1.2 – Principaux axiomes de la logique modale

exploitation, a fait l'objet de travaux à la fois logiques (voir par exemple Chellas (1980)) et philosophiques (Hintikka, 1962).

Le système normal de logique modale (voir par exemple Chellas (1980)) est un ensemble d'axiomes et de règles permettant de construire des preuves. Ce système est constitué de deux règles, le modus ponens et la nécessité, données dans le tableau 1.1, et un axiome modal qui s'ajoute à l'ensemble des axiomes de la logique propositionnelle classique :

$$(K) \quad \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$$

Il existe une série d'autres axiomes modaux qui sont communément associés à des notions logiques et philosophiques propres à la représentation de connaissances et de croyances non factuelles, voir section 1.1.5 ci-dessous et section 4.2.1 p. 72 pour le cas doxastique. Les principaux axiomes sont rappelés dans le tableau 1.2.

On peut ainsi effectuer un enchaînement d'inférences, en appliquant successivement les axiomes, et donc enrichir un ensemble de formules, dans un modèle de Kripke donné, sans s'intéresser précisément à la fonction de valuation des formules.

Nom	Axiome	Propriété de la relation associée
(D)	$\Box\varphi \rightarrow \Diamond\varphi$	Sérialité : $\forall u, \exists v uRv$
(M)	$\Box\varphi \rightarrow \varphi$	Réflexivité : $\forall u, uRu$
(4)	$\Box\varphi \rightarrow \Box\Box\varphi$	Transitivité : $\forall u, v, w (uRv \wedge vRw) \Rightarrow uRw$
(5)	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$	Euclidianité : $\forall u, v, w (uRv \wedge uRw) \Rightarrow vRw$

TABLEAU 1.3 – Principales correspondances entre axiomes et propriétés de la relation

### 1.1.4 Théorie de la correspondance

La théorie de la correspondance s'intéresse à la mise en relation des approches sémantique et axiomatique. Elle permet de relier un système formel de preuves au cadre sémantique de validité des formules. Pour la sémantique des mondes possibles, elle a d'abord été introduite par Sahlqvist (1975), puis établie par Van Benthem (1984), elle examine la correspondance entre axiomes et propriétés de la relation d'accessibilité.

Ainsi, elle permet de fixer des contraintes sur la relation des modèles de Kripke, comme la transitivité ou la symétrie. Les contraintes garantissent que les formules posées comme axiomes soient vraies dans des cadres respectant ces conditions et réciproquement. Ces axiomes sont donc acquis par la structure du modèle, plus précisément de par la relation d'accessibilité et ses propriétés. Par exemple, on peut montrer que pour tout cadre  $F = \langle W, R \rangle$  tel que  $R$  est sérielle, pour toute valuation  $s$ ,  $\langle W, R, s \rangle \models \Box\varphi \rightarrow \Diamond\varphi$  (Van Benthem, 1984).

La correspondance s'écrit comme un ensemble d'assertions du type :

$$\forall F = \langle W, R \rangle \quad (\forall s, \mathcal{M} \models \text{Axiome}) \Leftrightarrow R \text{ vérifie la propriété}$$

Cette correspondance induit donc une équivalence entre axiome et propriété de la relation. Cela signifie que l'on a l'implication « contrainte sur  $R \Rightarrow$  axiome » et sa réciproque « axiome  $\Rightarrow$  contrainte sur  $R$  »

Le tableau 1.3 résume les propriétés associées aux principaux axiomes présentés dans le tableau 1.2 (cf Chellas (1980)).

On note l'absence de l'axiome (K) dans le tableau 1.3 : ce dernier n'est relié à aucune propriété de la relation. En effet, cet axiome compose le système normal de logique modale, il est valide quel que soit le cadre  $F$  considéré.

### 1.1.5 Logiques aléthique, épistémique, doxastique, ...

Comme mentionné précédemment, la modalité peut exprimer différentes notions non factuelles : la nécessité, le savoir, la croyance, etc. Cette notion représentée par la modalité

Nom	Axiome	Propriété associée
(K)	$\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$	Rationalité
(D)	$\Box\varphi \rightarrow \Diamond\varphi$	Cohérence
(M)	$\Box\varphi \rightarrow \varphi$	Vérité
(4)	$\Box\varphi \rightarrow \Box\Box\varphi$	Introspection positive
(5)	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$	Introspection négative

TABLEAU 1.4 – Principaux axiomes de la logique modale avec la notion qui leur est associée, en logiques épistémique et doxastique

dépend directement des axiomes choisis : en effet, il est possible de ne sélectionner qu'une partie des axiomes et ainsi former une axiomatique qui correspond à une logique modale particulière. Ainsi, la logique épistémique, où  $\Box\varphi$  se lit comme « on sait que  $\varphi$  », et la logique doxastique, où  $\Box\varphi$  se lit comme « on croit que  $\varphi$  », sont deux systèmes modaux distincts dont l'axiomatique est différente : l'axiome (M) est présent dans la première mais pas dans la seconde (Fagin et al., 2004).

En effet, chaque axiome peut être interprété dans un cadre philosophique. Ainsi, (M) est décrit comme un axiome de vérité qui impose, dans la lecture épistémique, « si l'agent sait que  $\varphi$ , alors  $\varphi$  est vraie ». Il peut ainsi être interprété comme une infaillibilité du savoir (voir la section 4.2.1, p. 72, pour plus de détails sur le cas doxastique).

## 1.2 Axes de comparaison des logiques modales pondérées

Les logiques modales rappelées dans la section précédente ont fait l'objet de diverses extensions, nous considérons ici les logiques modales pondérées qui les enrichissent par l'introduction d'un coefficient numérique, comme présenté dans les sections 1.3 et 1.4. Syntaxiquement, ces modèles reposent sur des modalités  $\Box_\alpha$  et  $\Diamond_\alpha$  où  $\alpha$  est un nombre. Un tableau synthétique et récapitulatif est donné en bilan de ce chapitre, p. 24.

Auparavant, cette section présente les trois axes de comparaison que nous proposons afin de structurer ces modèles : le fait que la sémantique de Kripke soit préservée ou étendue, la position de la pondération, sur les mondes ou sur la relation, et enfin la sémantique de cette pondération, qui dépend du cadre arithmétique considéré.

### 1.2.1 Préservation vs extension de la sémantique de Kripke

Deux cas de figure se posent quand il s'agit de l'interprétation des formules modales pondérées : soit elle repose sur la sémantique classique des modalités, soit elle nécessite

d'adapter cette sémantique.

Dans le premier cas, les formules modales sont interprétés selon la sémantique de Kripke, rappelée en section 1.1, en utilisant la relation comme indicateur des mondes accessibles. Dans le second cas, les formules modales font appel à d'autres éléments que la relation pour leur interprétation. En particulier, le lien direct entre  $\Box$  et la quantification universelle des mondes accessibles n'est pas maintenu.

Un premier axe d'opposition apparaît ici correspondant respectivement à la préservation et à l'extension de la sémantique de Kripke. C'est l'axe principal que nous considérons, pour comparer les formalismes existants de logiques modales pondérées, il induit la distinction entre les sections 1.3 et 1.4.

### 1.2.2 Position de la pondération : monde vs relation

Lorsque le modèle de Kripke est étendu, il porte un ensemble de poids. Nous proposons un nouvel axe d'opposition possible, concernant l'effet induit par le placement de la pondération : celle-ci peut être portée par la relation ou par les mondes possibles.

D'un côté, la pondération de la relation étend l'interprétation de la modalité pondérée assez directement, en tenant compte de la nouvelle relation pondérée :  $R_\alpha$  remplace la relation  $R$  classique pour interpréter une modalité  $\Box_\alpha$  plutôt qu'une  $\Box$  classique. Elle remet néanmoins en cause l'ensemble des propriétés imposées à la relation, de type sérialité ou symétrie.

De l'autre côté, une pondération fixée sur les mondes n'implique pas nécessairement l'introduction de modalités indexées par un ensemble de degrés. En effet, le lien entre le poids porté par la modalité et celui porté par les mondes est moins direct.

Poser cette distinction comme nouvel axe de comparaison permet d'introduire une autre problématique autour de la pondération : indépendamment de son interprétation numérique discutée par la suite, se pose la question de l'expressivité que l'on peut lui accorder en fonction de son placement dans la sémantique de Kripke.

En effet, deux visions de la modalité s'opposent selon que la pondération se trouve sur les mondes ou sur la relation : on peut dire que les poids sur les mondes ont un effet *global* sur l'accessibilité entre les mondes, alors que le positionnement des poids sur la relation induit une influence *locale* de cette pondération, chaque accessibilité pondérée étant propre au monde de référence.

On peut considérer que l'existence d'une distribution de poids sur les mondes implique une idée de pondération globale : si un poids est associé à un monde, alors l'accessibilité vers celui-ci est identique quel que soit le monde d'origine. En d'autres termes, étant donné un monde  $w$  associé à un poids  $x$ , tout autre monde de l'univers qui a accès à  $w$  lui associe

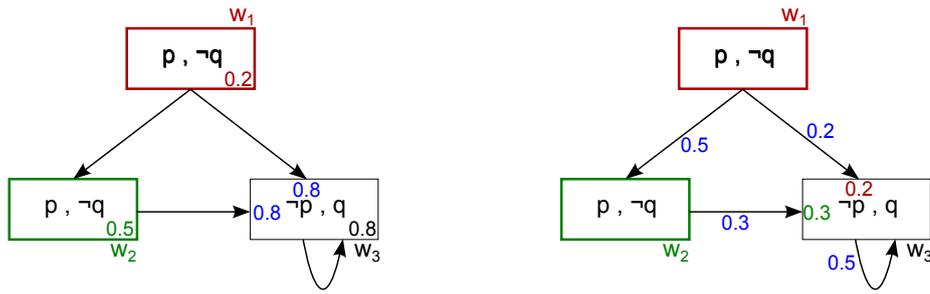


FIGURE 1.2 – Croyance globale et locale dans la sémantique de Kripke

la même valeur  $x$ . Ce principe est illustré sur le schéma de gauche de la figure 1.2 : il enrichit le modèle de Kripke de la figure 1.1 par des degrés portés ici par les mondes.

Ainsi, comme indiqué sur l'illustration à gauche de la figure 1.2, le degré affecté au monde  $w_3$  (de valeur 0.8) est identique pour les deux mondes de référence  $w_1$  et  $w_2$ .

Au contraire, une relation pondérée permet de moduler l'accessibilité en fonction du monde d'origine. Si l'on considère maintenant deux mondes  $w$  et  $w'$  qui ont pour successeur un même monde  $v$  mais avec une relation pondérée différente, alors chacun de ces mondes associe à  $v$  une valeur qui lui est propre. On peut donc interpréter cette pondération comme locale puisqu'elle dépend du monde de référence.

Le schéma à droite de la figure 1.2 illustre ce fait : le monde  $w_3$  n'est pas accessible au même degré depuis  $w_1$  (de valeur 0.2),  $w_2$  (de valeur 0.3) et lui-même (de valeur 0.5).

### 1.2.3 Sémantique de la pondération

Une troisième caractéristique des modèles présentés qui permet de les comparer concerne le sens donné à la pondération. En effet, selon la structure arithmétique sur laquelle les poids sont basés, ceux-ci peuvent être interprétés de différentes manières. Dans la suite, nous considérons deux cadres particuliers : la théorie des sous-ensembles flous (Zadeh, 1965) et la théorie des possibilités (Zadeh, 1978).

La différence entre ces deux formalismes concernant le sens donné aux poids réside dans l'interprétation de ces deux notions : les possibilités modélisent l'incertitude tandis que le formalisme flou capture les imprécisions. Incertitude et imprécision sont deux notions classiquement antagonistes (Bouchon-Meunier, 1993).

Lorsque les poids sont interprétés comme des valeurs floues, ils correspondent donc à des valeurs imprécises. Ainsi, ils permettent de considérer des mondes plus ou moins accessibles, plus ou moins possibles, etc. Cette imprécision dénote donc une modulation de la notion associée.

Lorsque les poids sont interprétés comme des degrés de possibilité, ils correspondent

à une mesure d'incertitude. Cela se différencie d'une imprécision de par le sens donné à ce type de mesures : une mesure incertaine n'est pas liée à l'imprécision autour d'une valeur mais à la véracité de la valeur que l'on lui attribue. En d'autres termes, l'incertitude concerne le fait que la mesure prenne réellement cette valeur plutôt que la valeur elle-même.

Par exemple, dire qu'une personne « *semble mesurer 1m70* » illustre l'absence d'information sur la taille de cette personne, l'estimation proposée étant peu certaine. Au contraire, l'imprécision indique que l'information est inexacte. On peut, par exemple, dire qu'une personne « *mesure environ 1m70* », ce qui induit que la taille évaluée est imprécise mais révèle une approximation de son estimation.

Il est toutefois important d'indiquer que l'interprétation du degré peut être indépendante de ces deux formalismes : nous présentons, en section 1.4, un modèle dont les poids sont liés à la non-représentativité des mondes possibles et ne sont vus ni comme des degrés flous, ni comme des degrés de possibilités.

Enfin, le cadre arithmétique choisi joue également un rôle dans la construction syntaxique de la logique modale pondérée : selon les relations entre les degrés, qui peuvent être définis de manière ordinale ou cardinale, le système modal pondéré sous-jacent est multi-modal avec ou sans relations inter-modalités, comme discuté dans les deux sections suivantes.

### 1.3 Modèles basés sur la préservation de la sémantique de Kripke

Comme indiqué dans la section précédente, nous opposons principalement les logiques modales pondérées suivant qu'elle préservent ou étendent la sémantique des mondes possibles.

Dans la première catégorie présentée dans cette section, des poids sont intégrés au modèle de Kripke, comme détaillé ci-dessous. L'interprétation classique des modalités est cependant conservée : la modalité pondérée est définie en regard des mêmes principes de quantification universelle ou existentielle des mondes accessibles validant la formule considérée (voir équations (1.2) et (1.3)). Néanmoins, ces modèles tiennent compte d'autres définitions de l'accessibilité entre les mondes.

Cette section est organisée comme suit : les deux premières sous-sections présentent une pondération possibiliste, respectivement sur les mondes et sur la relation, la troisième présente un modèle de pondération floue sur la relation, la dernière sous-section décrit d'autres types de pondération.

### 1.3.1 Pondération possibiliste des mondes

Le modèle proposé par Boutilier (1994) est basé sur une sémantique de Kripke enrichie par une distribution de possibilités, notée  $\pi$ , définie sur l'ensemble des mondes  $W$  : chaque monde possible  $w \in W$  est associé à un degré de possibilité  $\pi(w)$ .

Ce degré induit un ordre sur les mondes, la relation d'accessibilité est ensuite définie en fonction de cet ordre : un monde plus possible qu'un autre est accessible depuis celui-ci. Formellement :

$$wRw' \Leftrightarrow \pi(w) \leq \pi(w')$$

La figure 1.3 illustre un tel modèle de Kripke dans lequel les mondes sont pondérés, avec l'ordre induit :  $\pi(w_1) \leq \pi(w_2) \leq \pi(w_3)$  et la relation définie à partir de cet ordre.

La sémantique des deux modalités  $\Box$  et  $\Diamond$  est ensuite définie comme dans le cas classique, mais utilise cette relation d'accessibilité basée sur  $\pi$ . L'intégration du poids influence donc seulement la définition de la relation entre les mondes. Aussi, les modalités ne sont pas pondérées. La formule modale  $\Box\varphi$  est donc interprétée classiquement mais avec la relation d'ordre :

$$\begin{aligned} w \models \Box\varphi &\Leftrightarrow \forall w' \in R_w, w' \models \varphi \\ &\Leftrightarrow \forall w' \text{ tel que } \pi(w) \leq \pi(w'), w' \models \varphi \\ &\Leftrightarrow \varphi \text{ est vraie dans tous les mondes au moins aussi possibles que } w \end{aligned}$$

Du fait de cette nouvelle interprétation, ce modèle a la particularité de dessiner une organisation hiérarchique des mondes induite par la nouvelle relation d'accessibilité. En effet, l'utilisation d'une relation d'ordre comme relation d'accessibilité permet d'introduire l'idée qu'un monde se trouvant à une extrémité, c'est-à-dire qui n'a pas de successeur, est plus possible que tous les autres

On peut également noter que cette relation est, par définition, nécessairement antisymétrique, transitive et réflexive, ce qui est restrictif du point de vue de l'expressivité de la modalité : celle-ci ne peut pas représenter n'importe quelle notion modale, cf. section 1.1.5, page 13.

On peut noter que, réciproquement, une relation d'accessibilité entre des mondes permet d'induire une distribution de possibilités qualitatives sur les mondes.

Un autre modèle de pondération possibiliste des mondes est proposé par Dubois et al. (2012). Il enrichit également les modèles de Kripke à l'aide d'une distribution de possibilités sur l'ensemble  $W$ , également notée  $\pi$  : les mondes sont à nouveau considérés comme plus ou moins possibles. L'interprétation des formules repose cependant sur une extension de  $\pi$  aux formules : elles définissent  $\Pi(\varphi) = \max_{w \in W} \{\pi(w) \mid \mathcal{M}, w \models \varphi\}$ . Ce modèle permet essentiellement de généraliser la logique possibiliste et en propose une interprétation épistémique.

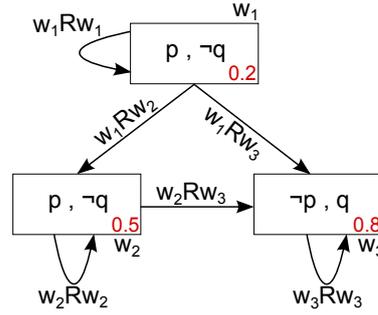


FIGURE 1.3 – Un modèle de Kripke pour une logique modale possibiliste qualitative,  $\pi(w)$  est donnée en rouge pour tout monde  $w$  et  $R$  est induite par l'ordre donné dans la définition

### 1.3.2 Pondération possibiliste de la relation

Le modèle proposé par Fariñas del Cerro et Herzig (1991) est aussi une extension possibiliste de la sémantique de Kripke, mais dans laquelle la pondération est appliquée à la relation : la distribution de possibilités est définie sur la relation, ce qui permet de modéliser une incertitude sur l'accessibilité entre deux mondes.

La relation  $R$  associée à une possibilité de valeur  $\alpha$  est notée  $R_\alpha$ . La relation entre deux mondes est donc pondérée par un degré de possibilité. Étant donné deux mondes  $v$  et  $w$ , l'expression  $wR_\alpha v$  désigne le fait que le monde  $v$  soit possiblement accessible, à un degré  $\alpha$ , depuis le monde  $w$ . Un tel modèle est illustré sur la figure 1.4.

Contrairement au modèle précédent, cette sémantique basée sur une relation pondérée induit la construction de nouveaux opérateurs modaux. Chaque relation  $R_\alpha$  définit une modalité  $\Box_\alpha$ , et  $\Diamond_\alpha$  par dualité, suivant l'interprétation classique des formules modales :

$$\forall \alpha \in [0, 1], \quad w \models \Box_\alpha \varphi \Leftrightarrow \forall w', (wR_\alpha w' \Rightarrow w' \models \varphi) \quad (1.5)$$

Cette définition permet donc de considérer une infinité de modalités pondérées, chacune engendrée par la relation pondérée par une valeur. Un système de logique multi-modale est ainsi défini. On ne peut déterminer de relation de dépendance entre ces modalités, car les relations indexées utilisées sont elles-mêmes indépendantes : la théorie des possibilités n'offre pas les outils nécessaires pour créer de telles dépendances.

On peut noter que ces indépendances entre  $R_\alpha$  pour des valeurs distinctes de *alpha* impliquent que les propriétés topologiques de type symétrie, transitivité ou euclidianité sont à considérer séparément pour chacune d'entre elles. Ainsi, sur l'exemple de la figure 1.4, l'existence des relations  $w_1R_{0.5}w_2$  et  $w_2R_{0.3}w_3$  peut conduire à s'interroger sur la pondération portée par la relation entre  $w_1$  et  $w_3$  par transitivité entre les deux relations précédemment mentionnées, alors que l'on a déjà  $w_1R_{0.2}w_3$ .

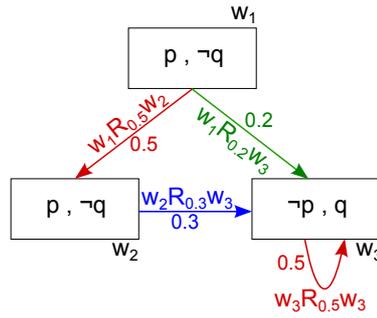


FIGURE 1.4 – Un modèle de Kripke pour une logique modale possibiliste quantitative, chaque relation pour une valeur est représentée dans une couleur donnée

### 1.3.3 Pondération floue de la relation

Le modèle proposé par Bou et al. (2009) repose sur le même principe de relation pondérée que ci-dessus. Le cadre formel utilisé est cependant différent : les poids se placent dans le contexte de la théorie des sous-ensembles flous (Zadeh, 1965).

Aussi ils représentent ici le degré d'imprécision de l'accessibilité et formalisent l'importance de l'accessibilité entre deux mondes. La pondération est donc un degré d'accessibilité entre les mondes : la relation étant plus ou moins importante suivant la pondération qu'elle porte. La figure 1.5 illustre la différence entre ce modèle et celui de Fariñas del Cerro et Herzig (1991) : la pondération est représentée par une différence sur l'épaisseur des flèches et non une différence de couleur (figure 1.4).

En se basant sur des mondes plus ou moins accessibles, l'extension floue de la modalité  $\Box_\alpha$ , et par dualité  $\Diamond_\alpha$ , est interprétée de la même manière que l'extension possibiliste, selon l'équation (1.5), page 19. Toutefois, l'interprétation, passant d'incertitude à imprécision, en est différente.

En particulier, il faut noter une différence cruciale concernant les dépendances des relation pondérées  $R_\alpha$  dans le cas flou : elles vérifient une propriété essentielle d'inclusion :

#### Propriété d'inclusion

$$\forall w_1, w_2 \in W, \forall \alpha, \beta \in [0, 1], \text{ si } \alpha \geq \beta \text{ alors } w_1 R_\alpha w_2 \Rightarrow w_1 R_\beta w_2$$

Ainsi, la présence d'une relation au degré  $\alpha$  indique, implicitement, la présence de toutes les relations  $R_\beta$  pour  $\beta \leq \alpha$  : les relations ne sont pas indépendantes les unes des autres, à la différence du cas possibiliste. Cette propriété induit une gradualité qui a pour conséquence une relation entre modalités, pour toute formule  $\varphi$  et pour tout modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $w \in W$  :

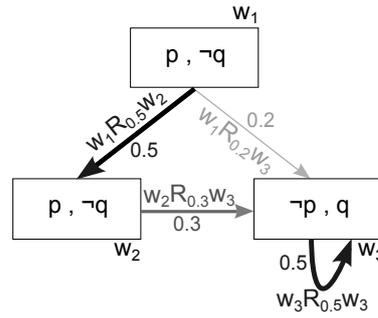


FIGURE 1.5 – Un modèle de Kripke pour une logique modale floue, le degré d'accessibilité est représenté par un assombrissement progressif de la flèche

### Propriété de gradualité

$$\forall \alpha, \beta \in [0, 1], \text{ si } \alpha \geq \beta \text{ alors } w \models \Box_\alpha \varphi \Rightarrow w \models \Box_\beta \varphi \quad (1.6)$$

Ainsi, la logique modale pondérée obtenue est, elle aussi, une logique multi-modale mais dont les modalités sont dépendantes les unes des autres. En effet, la relation  $R_\alpha$  est indexée selon son degré, et contrairement à la relation utilisée dans l'équation (1.5), celle-ci ne conduit pas à un modèle multi-relationnel : chaque relation  $R_\alpha$  correspond à l' $\alpha$ -coupe de la relation floue  $R$  et les relations  $R_\alpha$  ne sont pas indépendantes.

#### 1.3.4 D'autres types de pondération

Il existe d'autres modèles pondérés, qui sortent des cadres arithmétiques précédents. En particulier, Demolombe (2004) propose un modèle basé sur un ensemble de relations indexées  $R_\alpha$ , avec une sémantique équivalente à celles présentées pour l'approche possibiliste et pour l'approche floue (équation (1.5), page 19) mais pour laquelle l'intégration d'une notion d'agent plutôt que dans la signification du degré est primordiale et prioritaire sur l'interprétation du degré. Cette proposition est appliquée à la logique épistémique multi-agents (Fagin et al., 2004).

Ce modèle comporte un ensemble d'agents tel que pour chaque agent, il existe un ensemble de relations pondérées  $\{R_\alpha | \alpha \in [0, 1]\}$ . On note  $R_\alpha^m$  la relation de degré  $\alpha$  pour l'agent  $m$ . Étant donné un monde  $w$  (le monde de référence) et un agent  $m$ , cette définition tient compte de l'ensemble des mondes en relation avec  $w$  :

$$w \models \Box_\alpha^m A \Leftrightarrow \forall w', w R_\alpha^m w' \Rightarrow w' \models A \quad (1.7)$$

Il existe également des logiques modales pondérées dédiées à des types de modalités spécifiques. Ainsi, il existe aussi des logiques modales probabilistes, qui s'inscrivent dans le cadre particulier d'interprétation proposé par la théorie des probabilités. Une comparaison complète de ces modèles a été faite par Shirazi et Amir (2007). Au-delà de la logique modale pondérée, de nombreux travaux ont été fait concernant l'utilisation d'une probabilité pour induire la non-binarité des notions modales.

Enfin, Thiele et Kalenka (1993) proposent une logique temporelle floue, basée sur des concepts linguistiques permettant de nuancer les modalités. La logique temporelle considère la validité des formules dans le temps, et s'intéresse par exemple à des formules qui seront vraies à toutes les dates futures ( $\Box\varphi$ ) ou pour lesquelles il existe une date future à laquelle elles seront vraies ( $\Diamond\varphi$ ). Dans l'extension floue de cette logique modale, Thiele et Kalenka (1993) proposent de distinguer la validité des formules dans un passé ou futur qui est soit proche, soit lointain. Ce modèle n'est pas exposé ici car il est trop spécifique à la logique temporelle.

## 1.4 Modèles basés sur des extensions de la sémantique de Kripke

Les modèles présentés dans la section précédente conservent l'interprétation classique des formules modales, en utilisant une sémantique de Kripke enrichie, par exemple par une relation d'accessibilité différente. Les modèles présentés dans cette section proposent d'autres définitions des modalités pondérées  $\Box_\alpha$  et  $\Diamond_\alpha$ , associées à des poids dont les valeurs sont basées sur la comparaison avec une quantité dérivée de la structure du modèle de Kripke. En particulier, l'un d'eux est fondé sur la somme de degrés associés aux mondes, selon un cadre arithmétique différent des modèles présentés précédemment. Le nombre de mondes accessibles peut également être l'une de ces quantités dérivées sur lesquelles la nouvelle sémantique se base.

Cette section présente les modèles qui étendent la sémantique de Kripke, elle est organisée en deux sous-sections : la première présente un modèle basé sur l'exceptionnalité des mondes, la seconde repose sur un décompte des mondes.

### 1.4.1 Degré d'exceptionnalité et mondes invalidants

Le modèle proposé par Laverny et Lang (2004; 2005) enrichit la sémantique de Kripke en associant à chaque monde  $w$  une valeur d'exceptionnalité noté  $except(w)$ , qui représente à quel point ledit monde est différent, ou au contraire représentatif, de l'ensemble des mondes

qui composent le modèle. Un degré d'exceptionnalité est ensuite étendu aux formules :

$$\text{except}(\varphi) = \min_{w \in W} \{\text{except}(w) \mid \mathcal{M}, w \models \varphi\}$$

Une formule modale pondérée est ensuite interprétée de la manière suivante :

$$\mathcal{M}, w \models \Box_\alpha \varphi \quad \Leftrightarrow \quad \text{except}(\neg \varphi) \geq \alpha$$

Cette définition tient compte de la représentativité des mondes invalidants : plus un monde est exceptionnel, moins il est normal, c'est-à-dire qu'il n'est que peu représentatif de l'ensemble des mondes. Ainsi  $\Box_\alpha \varphi$  est vraie si  $\varphi$  est fausse seulement dans des mondes très exceptionnels. Cette exceptionnalité est une borne supérieure pour la pondération affectée à la modalité. Notons que l'on ne se place pas dans un monde de référence puisque le degré d'exceptionnalité d'un monde est global et ne dépend d'aucun autre.

Deux propriétés découlant de cette définition sont à souligner, on peut les désigner par les termes de généralité et gradualité. D'abord, la validité d'une formule modale ne dépend pas du monde de référence considéré, elle est générale à l'ensemble des mondes composant le modèle :  $\mathcal{M}, w \models \Box_\alpha \varphi \Leftrightarrow \mathcal{M} \models \Box_\alpha \varphi$ .

De plus, par définition de l'inégalité utilisée pour l'interprétation de la modalité pondérée, on retrouve une dépendance graduelle entre modalités pondérées, identique au cas de la pondération floue de Bou et al. (2009), présentée dans l'équation (1.6), page 21.

### 1.4.2 Comptage de mondes invalidants

Un modèle basé sur un comptage des mondes a été proposé par Fine (1972), puis de façon indépendante par Fattorosi-Barnaba et De Caro (1985). Il a notamment été repris par van der Hoek et Meyer (1992) puis exploité dans d'autres travaux (van der Hoek & De Rijke, 1995; Pacuit & Salame, 2004). Ce modèle ne propose pas de modification de la définition classique du modèle de Kripke : il n'intègre de pondération ni sur la relation ni sur les mondes. La modalité pondérée est définie au niveau de la quantification des mondes accessibles validants. La sémantique proposée introduit d'abord l'interprétation des formules modales contenant la modalité  $\Diamond_\alpha$  et pose :

$$\mathcal{M}, w \models \Diamond_\alpha \varphi \quad \Leftrightarrow \quad |R_w(\varphi)| \geq \alpha$$

Cette interprétation de  $\Diamond_\alpha$  est basée sur un durcissement du quantificateur existentiel :  $\Diamond_\alpha$  ne repose pas seulement sur la nécessité d'avoir *au moins un* monde accessible validant la formule  $\varphi$ , mais d'en avoir *au moins  $\alpha$* .

Par dualité entre les modalités, on a l'interprétation suivante pour  $\Box_\alpha$  :

$$\mathcal{M}, w \models \Box_\alpha \varphi \quad \Leftrightarrow \quad |R_w(\neg \varphi)| < \alpha$$

Ainsi, la modalité  $\Box_\alpha$  est pondérée par le nombre de mondes invalidants. La quantité  $\alpha$  représente une borne supérieure du nombre de mondes invalidants. Cela rejoint l'idée de l'exceptionnalité : si les mondes invalidants sont sur-représentés, alors la formule modale est invalide.

Cette approche par comptage modifie la quantification des mondes accessibles sur laquelle se base la sémantique classique des modalités : dans le cas de  $\Diamond$ , la quantification est enrichie puisqu'il est possible de moduler la modalité par le nombre de mondes validants au lieu de se contenter de la binarité classique.

Il faut noter que dans ce modèle, les poids sont donc nécessairement à valeur entière, comprise entre 1 et  $\max\{|R_w|, w \in W\}$ . De plus, ce modèle respecte également la propriété de gradualité définie dans l'équation (1.6), p. 21.

Cette définition s'appuie sur un comptage absolu des mondes. La logique majoritaire (Pacuit et Salame 2004; 2006) s'intéresse à un cas particulier de comptage relatif : l'opérateur modal majoritaire définit  $\Box\varphi$  est vraie si  $\varphi$  est formule vraie dans plus de la moitié des mondes accessibles. Elle repose donc sur une approche par comptage relatif mais définit des modalités non pondérées. Elle aborde également la question d'une sémantique dans laquelle l'ensemble des mondes  $W$  est infini.

Contrairement aux approches décrites dans la section précédente, basées sur une modification des modèles de Kripke, l'approche par comptage est de plus axiomatisée, tant dans le cas absolu que relatif (De Caro, 1988; Pacuit & Salame, 2006) : ces modèles proposent un ensemble de règles de manipulation de la modalité pondérée.

## 1.5 Bilan

L'état de l'art présenté dans cette section apporte une vision d'ensemble des différentes techniques de pondération des logiques modales. Nous avons proposé une étude comparative des modèles présentés, structurée suivant différents axes : le type de pondération (flou, possibiliste, autre), le mode de pondération dans la sémantique de Kripke (mondes vs relation) et le traitement réservé à la sémantique des formules modales (préservée vs modifiée). Le tableau 1.5 donne une vue synthétique de l'ensemble des approches citées, organisées selon les trois axes proposés.

Les modèles présentés se situent au niveau sémantique, en étendant les modèles de Kripke au cas pondéré ou en modifiant l'interprétation des modalités selon la sémantique de Kripke : ils ont tous pour particularité de ne pas traiter en profondeur la question de l'axiomatisation jusqu'au cas de la correspondance avec la sémantique (à l'exception de la logique majoritaire de Pacuit et Salame (2006) qui néanmoins ne définit pas de modalité pondérée).

Référence	Sémantique	Position	Cadre arithmétique
Boutilier (1994)	préservée	mondes	possibiliste
Fariñas del Cerro et Herzig (1991)	préservée	relation	possibiliste
Bou et al. (2009)	préservée	relation	flou
Demolombe (2004)	préservée	relation	
Thiele et Kalenka (1993)	étendue		flou
Laverny et Lang (2004)	étendue		autre : exceptionnalité
Fine (1972) Fattorosi-Barnaba et De Caro (1985) van der Hoek et Meyer (1992)	étendue		autre : comptage absolu
Pacuit et Salame (2004)	étendue		fréquentiste

TABLEAU 1.5 – Ensemble des modèles de logique modale pondérée présentés dans cet état de l’art.

Dans le chapitre suivant, nous présentons la sémantique que nous proposons pour une logique modale pondérée qui se place dans le cadre des approches par extension de la sémantique, par comptage relatif de mondes accessibles validants. Le chapitre 3 examine la question de la définition d’axiomes pondérés pour cette sémantique.



## Chapitre 2

# Proposition d'une logique modale pondérée : sémantique

L'adaptation des logiques modales au cas pondéré passe, dans un premier temps, par la création de modalités pondérées. Le poids, défini et intégré syntaxiquement, doit être pris en compte dans l'interprétation des formules modales formées à partir de la nouvelle syntaxe.

Dans ce chapitre, nous donnons une syntaxe qui assigne un poids compris dans l'intervalle réel  $[0, 1]$  à la modalité classique  $\Box$ , conduisant à la modalité  $\Box_\alpha$ . Ce chapitre décrit ensuite la sémantique que nous proposons, qui s'inscrit dans le cadre des approches par comptage, qui étendent la sémantique des modalités (cf. section 1.4.2, page 23, du chapitre précédent sur l'état de l'art). Sa particularité, détaillée et discutée, est de reposer sur un comptage relatif, et non absolu, des mondes accessibles. Elle permet ainsi de moduler le quantificateur universel classiquement associé à la modalité  $\Box$ .

Les propriétés qui découlent de cette définition sont ensuite présentées et analysées.

La structure du chapitre suit cette organisation : la section 2.1 décrit la syntaxe et la section 2.2 la sémantique que nous proposons. La section 2.3 discute cette dernière en termes de quantificateurs généralisés et la section 2.4 présente les propriétés qui découlent de la sémantique proposée, constituant une base préliminaire en travail sur les axiomes présentés dans le chapitre 3. Les travaux présentés dans ce chapitre et le suivant ont fait l'objet de publications (Legastelois et al., 2015b; Legastelois et al., 2015c).

Nous considérons, dans la suite de ce chapitre, que le modèle  $\mathcal{M} = \langle W, R, s \rangle$  est fixé et défini comme dans le cas classique de sémantique des mondes possibles (cf. chapitre 1, section 1.1.2, page 10).

## 2.1 Syntaxe

Cette section présente le cadre formel d'écriture de formules proposé pour une logique modale pondérée.

Nous définissons le langage suivant, qui étend le langage modal rappelé en section 1.1, page 9, en utilisant une grammaire formelle :

$$\begin{aligned}
\mathcal{P} &::= p, q, \dots \\
\mathcal{F} &::= \mathcal{P} \\
&| \mathcal{F} \wedge \mathcal{F} \\
&| \mathcal{F} \vee \mathcal{F} \\
&| \neg \mathcal{F} \\
&| \mathcal{F} \rightarrow \mathcal{F} \\
&| \Box_{\alpha} \mathcal{F}, \forall \alpha \in [0, 1] \\
&| \Diamond_{\alpha} \mathcal{F}, \forall \alpha \in [0, 1]
\end{aligned}$$

Nous précisons que la modalité  $\Diamond_{\alpha}$  est, comme dans le cas classique, duale par rapport à  $\Box_{\alpha}$  : la formule  $\Diamond_{\alpha} \varphi$  est une réécriture de  $\neg \Box_{\alpha} \neg \varphi$ . On appelle formule de logique modale pondérée toute formule appartenant à ce langage.

Il est important de noter le domaine de définition du poids  $\alpha$  proposé : définie dans l'intervalle  $[0, 1]$ , la pondération associée à la modalité peut être considérée comme un degré et exprime les graduations possibles de la notion portée par la modalité.

## 2.2 Définition de la sémantique proposée : comptage relatif

Cette section présente la sémantique des modalités pondérées que nous proposons. Celle-ci repose sur une approche fréquentiste du comptage des mondes. La section 2.2.1 en décrit les principes, la section 2.2.2 la définit formellement et la section 2.2.3 l'illustre dans un cas particulier.

### 2.2.1 Principes

Nous proposons une sémantique permettant d'interpréter les formules modales définies dans le langage de la section précédente. Cette sémantique suit le même principe que les approches par comptage, rappelées dans la sous-section 1.4, p. 22 et basées sur le nombre de mondes validant la formule concernée par la modalité.

Comme ces dernières, elle repose sur la définition classique des modèles de Kripke  $\mathcal{M} = \langle W, R, s \rangle$  (cf. section 1.1.2), non enrichis par des poids associés à  $w$  ou  $R$ .

Notre approche s'appuie sur deux caractéristiques différentes. D'abord, la sémantique est définie pour la modalité  $\Box_\alpha$ , le cas de  $\Diamond_\alpha$  découlant de cette première définition. Ensuite, et principalement, elle repose sur un comptage relatif et non un comptage absolu.

En effet, si la cardinalité des ensembles de mondes validants est prise en compte, elle est normalisée : les poids associés aux modalités sont définis en terme de proportions des mondes accessibles. La motivation de cette approche vient du choix de normalisation qui présente l'avantage de détacher la valeur du poids de la connexité du monde de référence : savoir que deux mondes accessibles valident la formule considérée n'a pas la même signification si le monde de référence possède deux ou mille mondes accessibles. Ainsi, une valeur absolue peut être interprétée comme trop peu informative : connaître le nombre total de mondes accessibles est nécessaire pour évaluer si les mondes accessibles validants sont représentatifs. En d'autres termes, savoir qu'un monde sur deux valide la formule est plus informatif que de seulement savoir qu'un monde accessible la valide, et permet de distinguer cette configuration du cas où un monde sur mille est validant : une valeur relative donne une information sur la représentativité des mondes accessibles validants par rapport à l'ensemble des mondes accessibles.

Il faut noter sur l'approche relative repose sur l'hypothèse, considérée dans la suite, selon laquelle l'ensemble des mondes accessibles est fini, afin de pouvoir calculer la fraction. S'il est infini, le comptage relatif des mondes accessibles validants doit être généralisé dans un cadre probabiliste.

### 2.2.2 Définition formelle

Formellement, le poids attribué à la modalité pondérée  $\Box_\alpha$  est défini comme borne inférieure de la proportion de mondes accessibles qui valident la formule associée :

**Définition 1** *Pour tout  $\varphi$  :*

$$\forall \alpha \in [0, 1]$$

$$\text{Si } R_w \neq \emptyset : \quad \mathcal{M}, w \models \Box_\alpha \varphi \quad \Leftrightarrow \quad \frac{|R_w(\varphi)|}{|R_w|} \geq \alpha \quad (2.1)$$

$$\text{Si } R_w = \emptyset : \quad \mathcal{M}, w \models \Box_\alpha \varphi \quad (2.2)$$

Il est important de souligner que l'on est obligé ici de distinguer le cas où  $R_w$  est vide, c'est-à-dire le cas où  $w$  est un monde aveugle : dans ce cas, la fraction n'est pas définie. De ce fait nous proposons, comme dans le cas classique, qu'une formule modale pondérée de la forme  $\Box_\alpha \varphi$  est valide dans un monde aveugle.

Par dualité, on a la relation entre les deux modalités pondérées  $\Diamond_\alpha = \neg \Box_\alpha \neg$ , ce qui induit la définition suivante.

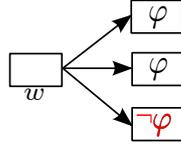


FIGURE 2.1 – Sémantique pondérée : si certains mondes accessibles invalident  $\varphi$ , il existe tout de même un  $\alpha = \frac{2}{3}$  tel que  $w \models \Box_\alpha \varphi$

**Définition 2** Pour tout  $\varphi$  :

$$\forall \alpha \in [0, 1]$$

$$\text{Si } R_w \neq \emptyset : \quad \mathcal{M}, w \models \Diamond_\alpha \varphi \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} > 1 - \alpha \quad (2.3)$$

$$\text{Si } R_w = \emptyset : \quad \mathcal{M}, w \not\models \Diamond_\alpha \varphi \quad (2.4)$$

**Démonstration :** L'interprétation de la modalité pondérée  $\Diamond_\alpha$  découle directement de la dualité entre  $\Diamond_\alpha$  et  $\Box_\alpha$ .

Pour tout  $\alpha \in [0, 1]$  et  $w$  tel quel  $R_w \neq \emptyset$ ,

$$\mathcal{M}, w \models \Diamond_\alpha \varphi \Leftrightarrow \mathcal{M}, w \not\models \Box_\alpha \neg \varphi \Leftrightarrow \frac{|R_w(\neg \varphi)|}{|R_w|} < \alpha \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} > 1 - \alpha$$

Si  $R_w = \emptyset$ ,  $\mathcal{M}, w \not\models \Box_\alpha \neg \varphi$  d'après l'équation 2.2, donc  $\mathcal{M}, w \models \neg \Box_\alpha \neg \varphi$ , soit  $\mathcal{M}, w \not\models \Diamond_\alpha \varphi$ . ■

### 2.2.3 Exemple illustratif

La figure 2.1 illustre la sémantique proposée dans le cas très simple d'un modèle de Kripke à quatre mondes, tel que  $w$  est en relation avec les trois autres.

Comme l'un de ces mondes accessibles ne valide pas  $\varphi$ , la contrainte imposée par le quantificateur universel de la définition classique n'est pas vérifiée et on a seulement  $w \not\models \Box \varphi$ .

Dans le cas pondéré, qui offre une expressivité accrue, on a également  $w \not\models \Box \varphi$ ; cependant, il permet de fournir l'information que  $w \models \Box_{\frac{2}{3}} \varphi$ .

## 2.3 Interprétation de la sémantique proposée en terme de quantification

Comme illustré dans l'exemple ce-dessus, la sémantique pondérée présente la particularité de modifier la quantification de mondes accessibles validant.

Cette section ouvre la discussion sur la sémantique proposée en terme de quantification des mondes. Dans un premier temps, nous détaillons les mécanismes associés à chacune des modalités : l'assouplissement et le durcissement du quantificateur pour respectivement les modalités  $\Box$  et  $\Diamond$ . Ensuite, nous discutons de l'utilisation de quantificateurs généralisés dans l'interprétation des formules modales. Ceux-ci représentent en effet une autre forme d'extension de la sémantique.

### 2.3.1 Assouplissement du $\forall$ et durcissement du $\exists$

Les définitions données pour interpréter les modalités pondérées sont toutes deux caractérisées en terme de quantification des mondes. Ainsi, la définition de  $\Box_\alpha$ , donnée par l'équation (2.1), page 29, ne requiert pas que tous les mondes accessibles valident la formule  $\varphi$  mais seulement une partie ; plus  $\alpha$  est élevé et plus la modalité est contrainte.

Classiquement, la sémantique de la modalité  $\Box$  repose sur la notion d'universalité des mondes validants, formellement exprimée par le quantificateur universel  $\forall$ , cf. équation (1.2) page 11. Dans le cas de la logique modale pondérée que nous proposons, la nécessité de validation de  $\varphi$  est maintenue mais seulement pour une proportion de mondes accessibles : cela induit un assouplissement du quantificateur universel.

Il est nécessaire de préciser que, si l'utilisation de cette proportion est plus riche qu'un modèle avec comptage absolu, comme discuté dans la section 2.2, p. 28, elle perd en information par rapport au quantificateur universel : on sait qu'il y a des mondes accessibles validant la formule mais on ne sait pas précisément lesquels. L'usage de la notion de proportion permet en effet de dénombrer mais pas de désigner les mondes concernés. Cela a une certaine influence sur la façon d'ébaucher une analyse de la théorie de la correspondance dans le chapitre suivant.

La modalité  $\Diamond_\alpha$ , quant à elle, requiert au moins une proportion  $1 - \alpha$  de mondes accessibles qui valident la formule  $\varphi$ , au lieu d'un seul monde accessible dans le cas classique. De la même façon que les modèles présentés dans la section 1.4.2, page 23, cela durcit le quantificateur existentiel : l'interprétation de  $\Diamond_\alpha$  peut nécessiter plus d'un seul monde validant. Pour la modalité  $\Diamond_\alpha$ , de par sa définition, plus  $\alpha$  est élevé et moins la condition est contrainte.

### 2.3.2 Discussion autour des quantificateurs généralisés

Nous proposons de mettre en perspective la sémantique proposée et l'utilisation de quantificateurs généralisés, du type quantificateur proportionnel (Keenan, 2004), qui étendent  $\forall$  et  $\exists$ . Cela consiste en une comparaison entre la sémantique classique et la sémantique pondérée, selon la quantification des mondes accessibles pris en compte pour l'interpréta-

tion d'une formule modale.

L'idée d'utiliser un quantificateur généralisé semble naturelle en raison de la définition ensembliste de l'interprétation de  $\Box_\alpha$ , qui est basée sur une proportion de mondes : on tient compte du cardinal de l'ensemble  $R_w$  pour calculer cette proportion.

Les quantificateurs proportionnels sont une catégorie des quantificateurs généralisés, qui désignent aussi des quantités abstraites comme « beaucoup », « la plupart » ou encore « pratiquement aucun ». Il est donc possible d'opposer ces quantificateurs linguistiques aux quantificateurs proportionnels.

Par exemple, le cas du *most versus more than half* est représentatif de la problématique liée à la relation entre ces deux types de quantificateurs : il s'agit de s'interroger sur la façon dont l'élément linguistique « la plupart » peut concrètement signifier « plus de la moitié » (Partee, 1989; Hackl, 2009). Il n'y a cependant pas de réel consensus quant aux termes linguistiques à employer pour désigner ces quantificateurs.

Aussi, il existe des sémantiques de logique modale basées sur des quantificateurs généralisés : par exemple, Askounis et al. (2012) en proposent un modèle dans lequel ils distinguent une modalité avec la quantification universelle « Knowledge means 'all' » de la quantification majoritaire, sous l'hypothèse que « la plupart » correspond à « plus de la moitié », associé à une autre modalité « Belief means 'most' ». Les auteurs proposent de cette façon un système bi-modal marqué par l'utilisation d'un quantificateur affaibli pour assouplir la notion de savoir en une croyance.

## 2.4 Propriétés

Cette section présente les principales propriétés de la sémantique proposée. Dans un premier temps, en section 2.4.1, nous définissons les cas limites pour les poids, en étudiant les cas où  $\alpha = 0$  et  $\alpha = 1$ . Nous établissons ensuite, dans la section 2.4.2, la propriété de gradualité décroissante, qui est une propriété essentielle par la suite, à la fois dans l'étude de la théorie de la correspondance (cf. chapitre 3) et dans le cadre de la logique doxastique (cf. Partie II). La dernière propriété présentée dans la section 2.4.3 concerne la relation entre les deux modalités  $\Box_\alpha$  et  $\Diamond_\alpha$ .

### 2.4.1 Cas limites

Cette section examine les deux cas limites,  $\alpha = 1$  et  $\alpha = 0$ . Nous montrons ici que le cas  $\alpha = 1$  correspond à la définition classique, et le cas  $\alpha = 0$  est une tautologie pour  $\Box$  et une contradiction pour  $\Diamond$  :

**Propriété 1** *Cas limites*

$$\begin{aligned}\Box_1\varphi = \Box\varphi & \quad \vDash \Box_0\varphi \\ \Diamond_1\varphi = \Diamond\varphi & \quad \vDash \neg\Diamond_0\varphi\end{aligned}$$

**Démonstration :** Soit  $\mathcal{M} = \langle W, R, s \rangle$  un modèle et  $w$  un monde de  $W$ , soit  $\varphi$  une formule quelconque

1) Montrons que  $\mathcal{M}, w \vDash \Box_1\varphi \Leftrightarrow \mathcal{M}, w \vDash \Box\varphi$ , ce qui permet d'établir l'égalité  $\Box_1\varphi = \Box\varphi$  :

Si  $R_w = \emptyset$ , alors  $w \vDash \Box_1\varphi$  et  $w \vDash \Box\varphi$ , donc on a bien  $w \vDash \Box_1\varphi \Leftrightarrow w \vDash \Box\varphi$ .

Sinon :

$$\begin{aligned}w \vDash \Box_1\varphi & \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} = 1 \\ & \Leftrightarrow R_w(\varphi) = R_w \text{ car } R_w(\varphi) \subseteq R_w \\ & \Leftrightarrow \forall w' \in R_w, w' \vDash \varphi \\ & \Leftrightarrow w \vDash \Box\varphi\end{aligned}$$

2) Montrons, de même, que  $\Diamond_1\varphi = \Diamond\varphi$  :

$$\begin{aligned}w \vDash \Diamond_1\varphi & \Leftrightarrow R_w \neq \emptyset \text{ et } \frac{|R_w(\neg\varphi)|}{|R_w|} < 1 \\ & \Leftrightarrow R_w \neq \emptyset \text{ et } |R_w(\neg\varphi)| < |R_w| \\ & \Leftrightarrow \exists w' \in R_w, w' \not\vDash \neg\varphi \\ & \Leftrightarrow \exists w' \in R_w, w' \vDash \varphi \\ & \Leftrightarrow w \vDash \Diamond\varphi\end{aligned}$$

3) On prouve que  $\vDash \Box_0\varphi$  et  $\vDash \neg\Diamond_0\varphi$  directement en utilisant les définitions : Si  $R_w = \emptyset$  alors, par définition, on a  $w \vDash \Box_0\varphi$  et  $w \vDash \neg\Diamond_0\varphi$ .

Sinon :

$$\begin{aligned}w \vDash \Box_0\varphi & \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} \geq 0 \\ w \vDash \Diamond_0\varphi & \Leftrightarrow \frac{|R_w(\varphi)|}{|R_w|} > 1 - 0\end{aligned}$$

La première assertion est toujours vérifiée tandis que la seconde est toujours fausse car  $R_w(\varphi) \subseteq R_w$ .

■

La valeur nulle possède donc le caractère particulier de ne porter aucune information quant à la valuation des formules dans les mondes accessibles. Du fait de cette propriété, on exclut autant que possible, par la suite, les cas où la modalité est pondérée par la valeur nulle. Elle est essentielle pour les travaux sur la théorie de la correspondance, présentés dans le chapitre suivant.

### 2.4.2 Gradualité décroissante

En raison de la transitivité de l'inégalité sur laquelle repose la sémantique proposée, on a la propriété de gradualité décroissante. Déjà présente dans les modèles de Bou et al. (2009) et Laverny et Lang (2004; 2005), cf. pages 20 et 22, cette propriété est rappelée ici :

#### Propriété 2 *Gradualité décroissante*

$$\forall \alpha, \beta \quad \alpha \geq \beta \quad \Rightarrow \quad (w \models \Box_\alpha \varphi \Rightarrow w \models \Box_\beta \varphi)$$

**Démonstration :** La démonstration de cette propriété est immédiate par transitivité des inégalités sur les proportions de mondes validants.

Soit  $\alpha$  et  $\beta$  tels que  $\alpha \geq \beta$  et soit un modèle  $\mathcal{M}$ . Supposons qu'il existe  $w \in W$  tel que  $w \models \Box_\alpha \varphi$ .

Alors  $\frac{|R_w(\varphi)|}{|R_w|} \geq \alpha$ . Comme  $\alpha \geq \beta$ , alors  $\frac{|R_w(\varphi)|}{|R_w|} \geq \beta$ , donc  $w \models \Box_\beta \varphi$

■

Cette propriété implique que si une formule  $\Box_\alpha \varphi$  est valide, alors c'est aussi le cas pour toutes les modalités  $\Box_\beta$  de degré inférieur. En d'autres termes, si  $\Box_\alpha \varphi$  est vraie dans un monde  $w \in W$ , alors le poids considéré  $\alpha$  peut être affaibli et prendre toute valeur plus petite. Ainsi, en reprenant l'exemple illustré en section 2.2.3, comme  $w \models \Box_{\frac{2}{3}} \varphi$  alors on a aussi, par exemple,  $w \models \Box_{0,5} \varphi$ .

Ainsi, le poids le plus informatif pour la modalité  $\Box_\alpha$  est donc la valeur maximale que peut prendre  $\alpha$ , puisque tous les poids inférieurs peuvent en être déduits. Par dualité, la valeur de  $\alpha$  la plus informative pour la modalité  $\Diamond_\alpha$  est sa valeur minimale. Ces deux remarques sont indispensables et exploitées pour répondre à la problématique étudiée en section 3.2.1.

### 2.4.3 Relation entre $\Box_\alpha$ et $\Diamond_\alpha$

Comme dans le cas classique, il existe une relation de dualité entre les deux modalités  $\Box_\alpha$  et  $\Diamond_\alpha$ . Nous montrons ici qu'il est possible de les lier autrement, en proposant une

relation entre les poids qu'elles portent. Il faut noter que, en raison de l'utilisation des inégalités respectivement large et stricte dans les interprétations de  $\Box_\alpha$  et  $\Diamond_\alpha$ , le lien de dualité entre les poids de ces deux modalités pondérées n'est pas réciproque :

**Propriété 3** *Relation entre  $\Box_\alpha$  et  $\Diamond_\alpha$*

$$\forall \alpha \in ]0, 1] \quad \models \Diamond_\alpha \varphi \rightarrow \Box_{1-\alpha} \varphi$$

$$\forall \alpha \in ]0, 1] \quad \not\models \Box_\alpha \varphi \rightarrow \Diamond_{1-\alpha} \varphi$$

**Démonstration :** Soit  $\alpha \in ]0, 1]$ . Soit  $\mathcal{M} = \langle W, R, s \rangle$  quelconque et  $w \in W$ .

1) Montrons que  $\models \Diamond_\alpha \varphi \rightarrow \Box_{1-\alpha} \varphi$  :

$$\begin{aligned} w \models \Diamond_\alpha \varphi &\Leftrightarrow R_w \neq \emptyset \text{ et } \frac{|R_w(\varphi)|}{|R_w|} > 1 - \alpha \\ &\Rightarrow R_w \neq \emptyset \text{ et } \frac{|R_w(\varphi)|}{|R_w|} \geq 1 - \alpha \\ &\Rightarrow w \models \Box_{1-\alpha} \varphi \end{aligned}$$

2) Pour montrer que  $\not\models \Box_\alpha \varphi \rightarrow \Diamond_{1-\alpha} \varphi$ , il suffit de construire un contre-exemple : considérons par exemple le modèle défini graphiquement sur la figure 2.1 de la page 30. On a  $\frac{|R_w(\varphi)|}{|R_w|} = \frac{2}{3}$ , donc  $w \models \Box_{\frac{2}{3}} \varphi$  car  $\frac{|R_w(\varphi)|}{|R_w|} \geq \frac{2}{3}$ . Toutefois  $w \not\models \Diamond_{\frac{1}{3}} \varphi$  car  $\frac{|R_w(\varphi)|}{|R_w|} \not> \frac{2}{3}$ . On en déduit que  $w \models \Box_{\frac{1}{3}} \varphi \not\models \Diamond_{\frac{2}{3}} \varphi$ .

De façon générale, l'asymétrie vient de ce qu'une inégalité large définit la sémantique de  $\Box_\alpha$  alors qu'une inégalité stricte intervient dans celle de  $\Diamond_\alpha$ . ■

## 2.5 Bilan

Ce chapitre a présenté notre proposition de sémantique pour les formules comportant une modalité pondérée  $\Box_\alpha$ , et par extension  $\Diamond_\alpha$ . Elle repose sur un assouplissement du quantificateur universel, permettant de ne considérer qu'une proportion plutôt que l'ensemble des mondes accessibles.

Nous avons également détaillé des propriétés découlant de cette sémantique. En effet, l'introduction d'un poids sur les modalités pose la question des liens entre, d'une part, les valeurs binaires et le cas classique et, d'autre part entre tous les degrés intermédiaires. Enfin, nous avons étudié les valeurs des poids vis-à-vis de la relation de dualité entre les modalités.

Ces travaux permettent de poser un cadre pour l'étude des axiomes et de la théorie de la correspondance, effectuée dans le chapitre suivant.

## Chapitre 3

# Proposition d'une logique modale pondérée : axiomes et correspondance

Alors que le chapitre précédent a adopté un point de vue purement sémantique pour une proposition de pondération d'une logique modale, ce chapitre traite la question de la pondération des axiomes modaux. Nous proposons de considérer ces derniers comme des règles de manipulation des modalités : associés au *modus ponens* classique, ils permettent d'établir des relations entre des formules dans lesquelles les opérateurs modaux apparaissent une fois, de manière répétée ou combinée. Par exemple, l'axiome (4)  $\vdash \Box\varphi \rightarrow \Box\Box\varphi$  établit une relation d'implication entre une unique occurrence et une répétition de  $\Box$  ; l'axiome (D)  $\vdash \Box\varphi \rightarrow \Diamond\varphi$ , également, exprime une implication entre les deux opérateurs modaux.

Ce chapitre étudie la transposition de ces axiomes au cas pondéré, pour définir des règles de manipulation et combinaison entre les opérateurs modaux pondérés  $\Box_\alpha$  et  $\Diamond_\alpha$ , en exploitant la sémantique proposée dans le chapitre précédent. La section 3.1 décrit plus formellement la problématique considérée et la section 3.2 présente les principes proposés pour construire les axiomes pondérés. La section 3.3 donne une vue d'ensemble des résultats obtenus, que nous proposons d'organiser selon une typologie contenant quatre catégories, dépendant des relations entre l'axiome pondéré, sa variante classique et la correspondance avec les contraintes sur les cadres de Kripke associés. Les sections 3.4 à 3.7 décrivent successivement chacun de ces quatre types. Enfin, la section 3.8 propose d'analyser les différences entre les axiomes classiques et les axiomes pondérés, notamment en terme de conséquences sur la théorie de la correspondance.

Les travaux présentés dans ce chapitre ont fait l'objet de plusieurs publications (Legastelois et al. 2015c; 2015b; 2017b).

### 3.1 Problématique

**Problématique considérée** Les axiomes ont un rôle fondamental dans la manipulation des opérateurs modaux : ils permettent de réécrire une formule sous une forme différente sans en altérer la valeur de vérité. L'objectif de ces travaux est d'étudier la transposition des axiomes classiques au cas pondéré : l'approche sémantique permet de fournir des justifications quant au choix des valeurs et relations entre les poids. De ce fait, nous ne proposons pas d'axiomatiser la logique modale pondérée indépendamment des axiomes modaux classiques, mais d'adapter ceux-ci dans le contexte de la sémantique que nous proposons.

Nous proposons donc de définir ce que nous appelons les règles de manipulation des opérateurs modaux pondérés, pour établir les liens et combinaisons possibles entre  $\Box_\alpha$  et  $\Diamond_\alpha$ . Nous considérons des axiomes pondérés génériques comme point de départ du processus de leur définition : ils sont obtenus à partir des axiomes classiques en remplaçant chaque modalité par un équivalent pondéré, chacun ayant son propre poids.

Ainsi, le candidat initial pour la variante pondérée de l'axiome (4), qui s'écrit  $\vdash \Box\varphi \rightarrow \Box\Box\varphi$ , est de la forme  $\vdash \Box_\alpha\varphi \rightarrow \Box_\beta\Box_\gamma\varphi$  où  $\alpha, \beta, \gamma \in [0, 1]^3$ .

**Valeurs pertinentes** La problématique considérée est alors d'établir les valeurs pertinentes pour les poids  $\alpha, \beta, \gamma$ , qui dépendent les uns des autres. La définition même de la pertinence des poids ouvre la voie à des discussions. Nous proposons d'adopter une approche sémantique basée sur l'interprétation des modalités présentée dans le chapitre précédent. Cela permet de motiver et justifier les valeurs choisies pour les coefficients, en exploitant le fait que  $\Box_\alpha$  et  $\Diamond_\alpha$  sont définies selon le comptage relatif des mondes accessibles.

Cette sémantique offre en effet l'avantage d'être suffisamment informative pour conduire à la définition d'axiomes pondérés. La connaissance d'une proportion de mondes validants permet, en terme sémantique, de traiter les questions d'union et d'intersection des ensembles utilisés. Cette connaissance est essentielle pour déterminer les valeurs des poids issus de transformations de diverses modalités. Les principes exploités sont décrits dans la section suivante, page 39.

L'approche que nous proposons permet d'identifier des dépendances entre les poids, valables dans tout cadre de Kripke, ou seulement dans des cadres dont la relation d'accessibilité vérifie des contraintes particulières. Comme détaillé par la suite, elle peut conduire à plusieurs variantes pondérées d'un axiome modal classique, comme c'est le cas pour 4 et C4. De plus, nous étudions également la réciproque, afin d'examiner si les cadres dans lesquels un axiome est vérifié satisfont des conditions spécifiques : les résultats obtenus peuvent être considérés comme ouvrant la voie à la définition d'une théorie de la correspondance.

**Établissement d'une typologie** Les travaux présentés dans ce chapitre établissent une typologie des axiomes modaux pondérés qui permet de les répartir dans quatre catégories, dépendant de leur réécriture du cas classique et selon les choix effectués pour la propriété du cadre classiquement associée, qui peut être conservée ou modifiée.

Ainsi, le type I regroupe les axiomes qui ne peuvent pas être transposés au cas pondéré en raison du trop grand degré de liberté offert par l'affaiblissement de la modalité classique. Le type II comporte les axiomes pondérés qui préservent la condition classique. Les types III et type IV contiennent les axiomes pondérés qui requièrent une modification de la condition, avec une correspondance dans la sémantique qui est respectivement non prouvée et prouvée.

## 3.2 Principes proposés pour la construction d'axiomes modaux

Cette section présente les principes généraux que nous proposons pour identifier des valeurs de poids pertinentes pour les axiomes pondérés candidats, trois types sont considérés. Les deux premiers arguments sont basés sur l'interprétation des axiomes comme les éléments d'un système d'inférence : la plupart des axiomes classiques (voir tableau 1.2, page 12) prennent la forme d'implications et sont d'autant plus exploitables qu'ils permettent, à l'aide du *modus ponens*, des inférences riches. Ainsi, il est attendu que les axiomes aient des prémisses faciles à vérifier et des conclusions informatives. Ce principe permet de guider la définition de valeurs de poids pertinentes selon la position des modalités, à la fois par rapport à la structure de l'axiome, comme décrit dans la section 3.2.1, et par rapport aux combinaisons de modalités, comme décrit dans la section 3.2.2. La section 3.2.3 présente un autre type d'outil, basé sur les propriétés des cadres de Kripke associés aux axiomes classiques que l'on cherche à pondérer.

### 3.2.1 Exploitation de la position des modalités

Dans le cas où ils contiennent une implication, nous définissons des axiomes comme utiles s'ils permettent une inférence riche. Afin de construire de tels axiomes, nous proposons de distinguer deux cas : d'une part, lorsque la prémisse de l'axiome est faible, donc facile à satisfaire, ce qui rend l'axiome facile à déclencher ; d'autre part, lorsque la conclusion est forte.

Nous soulignons que cette idée d'avoir une formule modale forte ou faible est particulièrement interprétable dans le cadre de la sémantique proposée, décrite dans le chapitre 2), notamment en raison de la propriété cruciale de gradualité décroissante (voir Prop. 2, p. 34).

Nous détaillons successivement ses conséquences pour la présence de modalités dans la conclusion puis la prémisse, en considérant d'abord le cas  $\Box_\alpha$  puis celui de  $\Diamond_\alpha$ .

**Présence de  $\Box_\alpha$  dans la conclusion** Un implication dont la conclusion contient  $\Box_\alpha$  est d'autant plus riche que  $\alpha$  est élevé.

En effet, toute valeur inférieure pouvant être déduite par gradualité décroissante, le cas le plus informatif, c'est-à-dire la conclusion la plus riche, est le cas où la valeur  $\alpha$  est la plus élevée possible.

Par exemple pour  $\varphi$  et  $\psi$  deux formules quelconques, si l'on a  $F_1 = \varphi \rightarrow \Box_{0.8}\psi$  et  $F_2 = \varphi \rightarrow \Box_{0.3}\psi$ , alors  $F_1$  est préférée à  $F_2$ , car plus informative. En effet,  $F_2$  peut être déduite de  $F_1$ , grâce à cette propriété, qui indique que  $\Box_{0.8}\psi \rightarrow \Box_{0.3}\psi$ , et grâce à la transitivité de l'implication.

On peut noter que, par conséquent, nous cherchons à éviter les cas où la modalité est pondérée par un degré nul, qui correspond à une implication non informative puisque sa conclusion est toujours vérifiée (cf. propriété 1, p. 32,  $\Box_0\varphi$  est une tautologie).

**Présence de  $\Box_\alpha$  dans la prémisse** Une implication dont la prémisse contient  $\Box_\alpha$  est d'autant plus riche que  $\alpha$  est faible.

En effet, une faible valeur définit une prémisse faible à satisfaire, et le cas le plus informatif : elle peut être déduite de toute valeur supérieure et reste donc déclenchable pour une observation contenant un poids plus élevé.

Par exemple, si l'on a  $F_3 = \Box_{0.8}\varphi \rightarrow \psi$  et  $F_4 = \Box_{0.3}\varphi \rightarrow \psi$ , alors  $F_4$  est préférée à  $F_3$  car plus informative. En effet,  $F_4$  peut être déduite de  $F_3$  en combinant la gradualité décroissante et la transitivité de l'implication :  $F_4$  est plus facilement déclenchable que  $F_3$  et définit de la sorte une inférence plus riche.

Par conséquent, pour une formule requise  $\Box_\alpha\varphi$ , comme chaque formule de la forme  $\Box_\beta\varphi$  avec un  $\beta$  plus élevé l'induit, l'axiome indiqué est déclenché. De cette façon, nous cherchons au maximum à éviter les cas où la modalité est pondérée par un degré nul, en particulier si le seul  $\alpha$  établi est du type  $\Box_0$  en conclusion.

**Cas de  $\Diamond_\alpha$**  Par dualité, pour la modalité  $\Diamond_\alpha$  qui est liée à une propriété de gradualité croissante, l'inverse est établi pour caractériser la pertinence des valeurs des poids associés :  $\alpha$  doit être minimal pour  $\Diamond_\alpha$  dans la conclusion et maximal dans la prémisse.

### 3.2.2 Cas des modalités successives

Certains des axiomes modaux classiques contiennent des modalités combinées par succession, du type  $\Box\Box$ ,  $\Box\Diamond$  et  $\Diamond\Box$  : c'est par exemple le cas de (B), (4) ou encore (C4). Dans

ces cas, le principe de maximisation/minimisation, discuté dans la section précédente, devient plus complexe et requiert la définition d'échange entre les poids respectifs. En effet, considérant par exemple le cas où  $\Box_\alpha\Box_\beta$ , l'informativité est maximale pour la combinaison  $\Box_1\Box_1$ , mais toutes les combinaisons ne peuvent pas être comparées. Ainsi, parmi les deux cas  $\Box_{0.5}\Box_{0.8}$  et  $\Box_{0.8}\Box_{0.5}$ , aucun ne peut être privilégié par rapport à l'autre. Plus généralement, faire décroître la valeur de  $\alpha$  peut être considéré comme informatif si cela permet d'augmenter la valeur de  $\beta$ .

Le cas de  $\Box_\alpha\Diamond_\beta$  et  $\Diamond_\gamma\Box_\delta$  peut être discuté selon le même principe, à la différence que le poids de  $\Diamond$ ,  $\beta$ , doit être minimisé. À nouveau, faire décroître  $\alpha$  dans le but de faire décroître  $\beta$ , ou réciproquement, est pertinent : parmi  $\Box_{0.9}\Diamond_{0.5}$  et  $\Box_{0.5}\Diamond_{0.3}$ , aucun ne peut être considéré comme plus informatif que l'autre.

Par conséquent, dans le cas où plusieurs modalités se succèdent, le principe général consistant à maximiser le poids associé à  $\Box_\alpha$  et minimiser celui de  $\Diamond_\beta$  lorsque les deux modalités apparaissent en conclusion d'une implication, devient plus complexe. Il en va de même pour le principe consistant à minimiser  $\alpha$  et maximiser  $\beta$ , pour  $\Box$  dans la prémisse. Cette discussion justifie l'établissement de plusieurs variantes des axiomes contenant des successions de modalités, comme détaillé en section 3.7, page 59.

### 3.2.3 Exploitation des contraintes sur la relation

Un second outil qui permet d'établir les relations entre les poids des axiomes modaux pondérés s'appuie la propriété de la relation associée à la version classique de l'axiome : la définition de la modalité pondérée étant basée sur l'ensemble des mondes accessibles, elle tient compte de la relation du modèle. Il est donc naturel de nous intéresser à l'influence des propriétés des relations sur la valeur des poids.

Étant donné que l'approche sémantique est au cœur de ces travaux sur les axiomes modaux, nous proposons d'exploiter la théorie de la correspondance (Van Benthem, 1984) pour élaborer l'extension pondérée des axiomes. En effet, l'équivalent sémantique des axiomes modaux est lié à une classe particulière de cadres, contrainte par des propriétés sur la relation d'accessibilité, à laquelle il est par exemple imposée d'être réflexive ou symétrique.

Le tableau 3.1 rappelle les définitions des principales propriétés possibles pour une relation d'accessibilité.

Nous considérons que privilégier la contrainte classiquement associée à la relation offre un point de départ à ces travaux : en faisant cette seule hypothèse sur la relation, il est intéressant d'analyser comment elle influence les valeurs des poids. Néanmoins, il se peut que la contrainte classique ne donne pas de garanties suffisantes pour pondérer les axiomes

sérielle	$\forall u$	$\exists v uRv$
réflexive	$\forall u$	$uRu$
symétrique	$\forall u, v,$	$uRv \Rightarrow vRu$
shift-réflexive	$\forall u, v,$	$uRv \Rightarrow vRv$
transitive	$\forall u, v, w,$	$(uRv \wedge vRw) \Rightarrow uRw$
euclidienne	$\forall u, v, w,$	$(uRv \wedge uRw) \Rightarrow vRw$
dense	$\forall u, v,$	$uRv \Rightarrow \exists w (uRw \wedge wRv)$
convergente	$\forall u, v, w,$	$(uRv \wedge uRw) \Rightarrow \exists x (vRx \wedge wRx)$
fonctionnelle	$\forall u, v, w,$	$(uRv \wedge uRw) \Rightarrow v = w$

TABLEAU 3.1 – Propriétés éventuelles d'une relation  $R$  définie sur  $W \times W$ , avec  $u, v, w \in W$ .

de manière pertinente. En considérant un point de vue sémantique pour l'interprétation des axiomes pondérés, nous ne considérons donc pas seulement les cadres qui satisfont la propriété correspondant à sa version classique : nous proposons d'examiner si d'autres propriétés sont pertinentes. Ce principe peut remettre en cause la compatibilité avec le cas classique pour lequel tous les poids sont égaux à 1.

Par ailleurs, lorsqu'elle apparaît en prémisse, la modalité  $\Box_\alpha$ , qui présente une interprétation plus expressive que  $\Box$ , est aussi moins spécifique. En effet, savoir qu'une proposition de mondes accessibles valide une formule ne permet pas d'avoir des informations sur la validité de cette formule dans chaque monde individuellement :  $\Box_\alpha$  donne une indication globale et mène à une incertitude sur la valuation propre à chaque monde, alors que la modalité classique  $\Box$  donne une indication sur tous les mondes accessibles. Par conséquent, il est naturel d'attendre des extensions graduelles des axiomes qu'elles imposent des contraintes supplémentaires sur la relation. Plus précisément, il est possible que l'axiome obtenu n'exclue pas le cas où la conclusion est de la forme  $\Box_0\varphi$  qui, comme discuté en section 2.4.1 ne serait pas informative. L'approche que nous proposons consiste donc à observer et déterminer les propriétés de la relation qui permettent d'exclure de tels cas. Pour cela, il est souvent nécessaire de renforcer la contrainte sur la relation.

D'autre part, lorsqu'elle apparaît dans la conclusion, la modalité assouplie  $\Box_\alpha$  avec  $\alpha < 1$ , porte moins d'information. Cela rejoint le principe de maximisation du poids en conclusion, décrit en section 3.2.1 page 39. Pour compenser cet affaiblissement de la modalité, nous proposons d'éventuellement relâcher la contrainte associée : avec une contrainte plus faible, la prémisse peut être affaiblie, ce qui rend possible l'enrichissement de la conclusion de l'axiome. Dans certains cas, la version pondérée de l'axiome permet donc de supprimer la propriété de la relation correspondant classiquement, comme détaillé en

section 3.3.2, page 44.

Enfin, une fois qu'une extension pondérée suffisamment pertinente est établie, selon une contrainte sur la relation qui peut être renforcée ou relâchée, nous étudions si la réciproque est vraie : cela revient à montrer que si un axiome est supposé, alors n'importe quel modèle respecte nécessairement la contrainte associée.

### 3.3 Typologie des axiomes pondérés établis

Cette section donne une vue d'ensemble des résultats obtenus : la section 3.3.1 présente la forme générale des théorèmes établis et de leur démonstration, en appliquant les principes de pondération décrits dans la section précédente pour définir des valeurs de poids pertinentes. La section 3.3.2, quant à elle, présente la typologie que nous proposons pour organiser les axiomes pondérés.

#### 3.3.1 Forme générale des théorèmes prouvés et leur démonstration

La forme générale des théorèmes établis par la suite est exprimée de la façon suivante : on note  $\varphi_\alpha$  l'axiome pondéré, c'est-à-dire une formule contenant des modalités pondérées. Un théorème s'écrit alors :

$$\forall \langle W, R \rangle \\ R \text{ satisfait les propriétés } P \Rightarrow \forall \alpha \in [0, 1], \forall s, \quad \langle W, R, s \rangle \models \varphi_\alpha$$

avec  $P$  un ensemble de propriétés de la relation, parmi celles données dans le tableau 3.1.

Comme illustré dans la section suivante, un tel théorème ne peut pas toujours être prouvé : cela signifie qu'il peut être montré qu'il existe un cadre  $\langle W, R \rangle$  dont la relation  $R$  satisfait les propriétés  $P$ ,  $\alpha \in [0, 1]$ , une valuation  $s$  et un monde  $w \in W$  tels que  $\langle W, R, s \rangle, w \not\models \varphi_\alpha$ . Dans la plupart des cas, un résultat plus général encore peut être établi, pour lequel le poids est universellement quantifié, ce qui a pour forme :

$$\forall \alpha \in [0, 1] \\ \exists \langle W, R \rangle \text{ tel que } R \text{ satisfait } P, \exists s \exists w \in W \quad \langle W, R, s \rangle, w \not\models \varphi_\alpha$$

$\langle W, R, s \rangle, w$  est un contre-exemple que l'on construit explicitement et qui permet de prouver ce résultat.

Lorsque le premier théorème peut être établi, sa réciproque peut être examinée. Elle s'écrit :

$$\forall \langle W, R \rangle \\ (\forall \alpha \in [0, 1], \forall s, \langle W, R, s \rangle \models \varphi_\alpha) \Rightarrow R \text{ satisfait } P$$

Un tel résultat est généralement montré en utilisant la contraposée :

$$\forall \langle W, R \rangle$$

$$R \text{ ne satisfait pas } P \Rightarrow \exists \alpha \in [0, 1], \exists s, \exists w \in W \langle W, R, s \rangle, w \not\models \varphi_\alpha$$

La preuve associée demande de construire, pour un cadre  $\langle W, R \rangle$  avec une relation  $R$  qui ne satisfait pas l'ensemble de propriétés, une valeur pour  $\alpha$ , une valuation  $s$  et un monde  $w \in W$  tel que  $\varphi_\alpha$  n'est pas valide dans  $w$ .

### 3.3.2 Quatre types d'axiomes

L'ensemble des résultats obtenus, démontrés et détaillés dans les sections 3.4 et 3.7, est présenté dans le tableau 3.2 : il liste les extensions pondérées des axiomes modaux classiques en indiquant les valeurs pertinentes des poids et les contraintes sur la relation de Kripke associées. démonstrations en sont données dans les sections suivantes.

Nous proposons d'organiser ces axiomes pondérés selon une typologie formée de quatre catégories discutées et détaillées dans les sections suivantes et indiquées dans la dernière colonne du tableau 3.2 : elles distinguent les quatre cas suivants :

- (I) axiomes impondérables
- (II) axiomes pondérés avec correspondance classique
- (III) axiomes pondérés sans correspondance
- (IV) axiomes pondérés avec correspondance enrichie

Ces quatre types dépendent de la relation entre l'axiome pondéré et les contraintes sur la relation d'accessibilité qui interviennent dans les théorèmes, ainsi que la relation entre ces contraintes et celles en correspondance avec l'axiome modal classique qui lui correspond.

Le type I regroupe les axiomes qui ne peuvent pas exploiter la liberté offerte par l'intégration d'un poids, en raison d'un trop grand relâchement de l'axiome classique. Le type II est, pour sa part, composé des axiomes pondérés qui préservent la contrainte usuellement associée, sans la modifier. Les types III et IV regroupent les axiomes pondérés pour lesquels la contrainte associée doit être modifiée afin d'intégrer des poids pertinents, comme défini dans la section 3.2. Ces contraintes permettent de prouver s'il y a correspondance entre l'axiome pondéré et la contrainte enrichie (III) ou non (IV).

Il est important de noter que nous autorisons les axiomes classiques à avoir plusieurs extensions pondérées, selon les variantes possibles d'enrichissement de la contrainte.

## 3.4 Type I : axiomes impondérables

La première catégorie d'axiomes regroupe ceux pour lesquels le seul poids possible se ramène au cas classique, avec la valeur 1 pour tous les poids. Ces axiomes ne peuvent

(K $_{\alpha}$ )	$\Box_{\alpha}(\varphi \rightarrow \psi) \rightarrow (\Box_{\beta}\varphi \rightarrow \Box_{\max(0,\alpha+\beta-1)}\psi)$		(II)
(NA1 $_{\alpha}$ )	$(\Box_{\alpha}\varphi \wedge \Box_{\beta}\psi) \rightarrow \Box_{\max(0,\alpha+\beta-1)}(\varphi \wedge \psi)$		(II)
(NA1 $_{\alpha}$ )	$\Box_{\alpha}(\varphi \wedge \psi) \rightarrow (\Box_{\alpha}\varphi \wedge \Box_{\alpha}\psi)$		(II)
(NA2 $_{\alpha}$ )	$(\Box_{\alpha}\varphi \vee \Box_{\beta}) \rightarrow \Box_{\min(\alpha,\beta)}(\varphi \vee \psi)$		(II)
(CD $_{\alpha}$ )	$\Diamond_{\alpha}\varphi \rightarrow \Box_{1-\alpha}\varphi$		(IV)
(M $_{\alpha}$ )	$\Box_1\varphi \rightarrow \varphi$	réflexive	(I)
(B $_{\alpha}$ )	$\varphi \rightarrow \Box_1\Diamond_1\varphi$	symétrique	(I)
( $\Box$ M $_{\alpha}$ )	$\Box_1(\Box_1\varphi \rightarrow \varphi)$	shift-réflexive	(I)
(5 $_{\alpha}$ )	$\Diamond_1\varphi \rightarrow \Box_1\Diamond_1\varphi$	euclidienne	(I)
(D $_{\alpha}$ )	$\Box_{\alpha}\varphi \rightarrow \Diamond_{1-\alpha+\varepsilon}\varphi$	sérielle	(II)
(C $_{\alpha}$ )	$\Diamond_1\Box_{\alpha}\varphi \rightarrow \Box_{1-\alpha+\varepsilon}\Diamond_1\varphi$	convergente	(III)
(4 $_{\alpha}$ )	$\Box_{\alpha}\varphi \rightarrow \Box_1\Box_{\alpha}\varphi$	transitive & euclidienne	(III)
(C4 $_{\alpha}$ )	$\Box_{\alpha}\Box_{\beta}\varphi \rightarrow \Box_{\beta}\varphi$	transitive & euclidienne	(IV)
(C4 $_{\alpha}$ )	$\Box_{\alpha}\Box_1\varphi \rightarrow \Box_{\alpha}\varphi$	shift-réflexive	(IV)

TABLEAU 3.2 – Axiomes pondérés obtenus avec la contrainte associée (non nécessairement correspondante) et leur type, comme défini dans la section 3.3.2, p. 44. Les coefficients  $\alpha, \beta$  sont définis dans l'intervalle réel  $[0,1]$  et  $\varepsilon \in ]0,\alpha]$ .

pas bénéficier du relâchement de la modalité par l'intégration d'une pondération, pour des raisons liées à la pertinence des poids.

Il existe deux arguments qui justifient de ne pas pondérer un axiome : d'une part, nous prouvons pour l'axiome (M) qu'une pondération à valeur inférieure à 1 n'est pas assez pertinente. D'autre part, le relâchement de l'axiome n'est pas souhaitable pour les axiomes ( $\Box$ M), (B) et (5), car cela peut affecter leur pouvoir de manipulation. Ces deux cas sont successivement décrits par la suite.

### 3.4.1 Axiome (M)

L'axiome (M) classique s'écrit  $\vdash \Box\varphi \rightarrow \varphi$  et, d'après la théorie de la correspondance, est associé à la contrainte de réflexivité de la relation d'accessibilité. Sa version pondérée s'écrit  $\vdash \Box_{\alpha}\varphi \rightarrow \varphi$ . La question est ici, et pour chacun des axiomes étudiés par la suite, d'examiner les valeurs pertinentes pour le poids  $\alpha$ .

Or, dans le cas où la pondération est  $\alpha < 1$ , la contrainte de réflexivité ne permet pas, à elle seule, de conclure : si l'on considère un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  réflexive et  $w \in W$ , alors si  $w \models \Box_{\alpha}\varphi$ , il existe au moins un monde accessible depuis  $w$  qui ne valide pas  $\varphi$ . La réflexivité n'offre aucune garantie pour que  $w$  ne fasse pas partie de l'ensemble de ces mondes invalidants.

Formellement, on peut établir le théorème suivant :

**Théorème 1 ( $M_\alpha$ )**  $\forall \alpha \in [0, 1[$ , il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  réflexive et  $w \in W$  tels que  $\mathcal{M}, w \models \Box_\alpha \varphi$  mais  $\mathcal{M}, w \not\models \neg \varphi$ .

**Démonstration :** Soit  $\alpha \in [0, 1[$ . La preuve consiste à construire un tel modèle  $\mathcal{M}$ . Soit  $W$  un ensemble de  $n$  mondes, avec  $n$  tel que  $(n-1)/n \geq \alpha$  (un tel  $n$  existe puisque par hypothèse  $\alpha < 1$ ) et  $w$  un monde de référence dans  $W$ . Soit  $R$  une relation réflexive telle que  $R_w = W$ . Soit  $s$  la valuation suivante :

- (i)  $x \models \varphi$  pour tout  $x \in W \setminus \{w\}$
- (ii)  $w \models \neg \varphi$

En posant  $\mathcal{M} = \langle W, R, s \rangle$ , on a  $\mathcal{M}, w \models \Box_\alpha \varphi$  car  $|R_w| = n$ ,  $|R_w(\varphi)| = n-1$  et  $(n-1)/n \geq \alpha$ . Cependant  $\mathcal{M}, w \not\models \varphi$ . ■

De plus, l'ajout de contraintes sur la relation (comme la transitivité, la symétrie ou l'eucledianité par exemple) n'apporterait pas l'information nécessaire sur la valuation du monde de référence pour garantir qu'il ne soit pas invalidant. Aussi, il semble difficile de proposer une variante enrichie de la contrainte qui permette de contrebalancer la perte d'information qui survient lorsque l'on emploie une modalité pondérée plutôt que classique, comme discuté dans la section 3.2.3, page 41.

Afin de respecter le cas classique, nous conservons la réflexivité mais avec la seule valuation pertinente du poids :  $\alpha = 1$ . Cela correspond donc au cas classique, rendant cet axiome impondérable.

### 3.4.2 Axiome ( $\Box M$ )

L'axiome classique ( $\Box M$ ) s'écrit  $\vdash \Box(\Box \varphi \rightarrow \varphi)$  et, d'après la théorie de la correspondance, est associé à la shift-réflexivité de la relation d'accessibilité. Sa forme pondérée est notée  $\vdash \Box_\alpha(\Box_\beta \varphi \rightarrow \varphi)$ . Comme détaillé ci-dessous, nous prouvons que la seule valuation pertinente des poids est 1 pour  $\alpha$  et pour  $\beta$ , à condition de respecter la contrainte de shift-réflexivité.

De manière évidente, le cas où  $\alpha = 0$  est une tautologie, mais ne présente que peu d'intérêt, comme discuté dans la section 2.2, page 28.

Dans le cas où  $\beta = 1$ , en utilisant la shift-réflexivité, alors  $\langle W, R \rangle \models \Box_1(\Box_1 \varphi \rightarrow \varphi)$ , ce qui induit n'importe quelle valeur  $\alpha$ , grâce à la gradualité décroissante.

Dans tous les autres cas, nous avons le théorème :

**Théorème 2** ( $\square\mathbf{M}_\alpha$ )  $\forall\alpha \in ]0, 1], \forall\beta \in [0, 1[$ , il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  shift-réflexive et  $w \in W$  tels que  $\mathcal{M}, w \not\models \square_\alpha(\square_\beta\varphi \rightarrow \varphi)$

**Démonstration :** Soit  $\alpha \in ]0, 1]$  et  $\beta \in [0, 1[$ . La preuve consiste, là aussi, à construire un tel modèle  $\mathcal{M}$ , qui est très similaire à celui proposé dans le preuve du Théorème 1, page 46.

Soit  $W$  un ensemble de  $n + 1$  mondes, où  $n$  est tel que  $(n - 1)/n \geq \beta$ . Un tel  $n$  existe puisque par hypothèse  $\beta < 1$ . Soit  $u \in W$  et  $w \in W$ . Soit  $R$  définie de la façon suivante :  $R_u = \{w\}$ ,  $R_w = W \setminus \{u\}$  et  $\forall x \in W \setminus \{u, w\}$ ,  $R_x = \{x\}$ . Par construction,  $R$  est shift-réflexive. Soit  $s$  une valuation telle que (i)  $x \models \varphi$  pour tout  $x \in W \setminus \{w\}$  et (ii)  $w \models \neg\varphi$ .

Ainsi, en utilisant la même preuve que pour le Théorème 1,  $w \models \neg(\square_\beta\varphi \rightarrow \varphi)$ . Aussi  $u \models \square_{1-\beta}(\square_\beta\varphi \rightarrow \varphi)$  : pour tous les  $\alpha > 0$ ,  $u \not\models \square_\alpha(\square_\beta\varphi \rightarrow \varphi)$ . ■

Comme précédemment, l'ajout de contraintes sur la relation (comme par exemple la transitivité, la symétrie ou l'eulidianité) n'apporterait pas d'information supplémentaire sur la valuation du monde de référence  $w$ .

### 3.4.3 Axiome (B)

L'axiome (B) classique s'écrit  $\vdash \varphi \rightarrow \square\Diamond\varphi$  et est associé à la contrainte de symétrie de la relation. Sa version pondérée ( $\mathbf{B}_\alpha$ ) se note  $\vdash \varphi \rightarrow \square_\alpha\Diamond\varphi$ . Elle présente deux caractéristiques qui la différencient des axiomes (M) et ( $\square\mathbf{M}$ ) précédemment considérés : la prémisse de l'implication ne contient aucune modalité et la conclusion contient une combinaison de modalités. La définition de poids pertinents conduit alors à considérer que faire décroître la valeur de  $\alpha$  est possible à condition de diminuer la valeur de  $\beta$ , comme discuté en section 3.2.2, p. 40. Par souci de compatibilité avec le cas classique, la partie sémantique de cet axiome pondéré considère une relation symétrique. Nous montrons que, à l'exception du cas non informatif pour lequel  $\alpha = 0$ ,  $\beta$  ne peut avoir d'autre valuation que 1 :

**Théorème 3** ( $\mathbf{B}_\alpha$ )  $\forall\alpha \in ]0, 1], \forall\beta \in [0, 1[$ , il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  symétrique et  $w \in W$  tels que  $\mathcal{M}, w \models \varphi$  mais  $\mathcal{M}, w \not\models \square_\alpha\Diamond\varphi$

**Démonstration :** Soit  $\alpha \in ]0, 1]$  et  $\beta \in [0, 1[$ . La preuve consiste à construire un tel modèle  $\mathcal{M}$ . Soit  $F = \langle W, R \rangle$  un cadre contenant  $n = \lfloor \frac{1}{1-\beta} \rfloor + 1$  mondes. Soit  $w, w' \in W$  deux mondes de référence et  $R$  la relation d'accessibilité symétrique telle que  $R_w = \{w'\}$  et  $R_{w'} = W$ . Soit  $s$  la valuation telle que :

- (i)  $x \models \neg\varphi$  pour tout  $x \in W \setminus \{w\}$
- (ii)  $w \models \varphi$

Ainsi  $\mathcal{M}, w' \models \neg\Diamond_\beta\varphi$  : en effet,  $\frac{|R_{w'}(\varphi)|}{|R_{w'}|} = \frac{1}{n}$ . Par définition de  $n$ ,  $n > \frac{1}{1-\beta}$ , et donc  $\frac{1}{n} < 1-\beta$ . Par conséquent, comme  $w'$  est l'unique successeur de  $w$ ,  $\mathcal{M}, w \not\models \Box_\alpha\Diamond_\beta\varphi$ . ■

### 3.4.4 Axiome (5)

L'axiome (5) classique s'écrit  $\vdash \Diamond\varphi \rightarrow \Box\Diamond\varphi$  et, d'après la théorie de la correspondance, est associé à une relation euclidienne. L'axiome pondéré (5 $_\alpha$ ) s'écrit  $\vdash \Diamond_\alpha\varphi \rightarrow \Box_\beta\Diamond_\gamma\varphi$  : il contient une combinaison de modalités et possède la spécificité d'avoir une modalité  $\Diamond$  en prémisse.

Si l'on considère le cas où  $\alpha = 1$ , on peut en conclure que la compatibilité avec le cas classique impose d'avoir  $\beta = \gamma = 1$ .

Notre objectif est donc d'examiner s'il est possible d'avoir une meilleure valuation, notamment parce qu'il serait plus pertinent que  $\alpha$  soit augmenté (cf. section 3.2). Par exemple, pour avoir une conclusion plus riche, il est possible d'augmenter la valeur de  $\gamma$ , mais cela nécessite de diminuer  $\beta$ , ou éventuellement d'augmenter  $\alpha$  ce qui imposerait une prémisse plus informative.

Comme pour les théorèmes précédents, nous considérons une relation euclidienne, par compatibilité avec le cas classique. Considérer un ajout de contraintes constitue une problématique supplémentaire.

Le théorème 4 ci-dessous montre que  $\beta$  est nécessairement égal à 1 ; ensuite, le théorème 5 montre que le cas où  $\alpha < 1$  ne permet pas de faire varier  $\gamma$  et n'est donc pas pertinent.

**Théorème 4**  $\forall\beta \in ]0, 1], \forall\gamma \in [0, 1],$

$\forall\bigcirc \in \{\Box, \Diamond\}, \forall\langle W, R \rangle,$

$R$  euclidienne  $\Rightarrow \forall w \in W, \text{ si } w \models \Box_\beta \bigcirc_\gamma \varphi \text{ alors } w \models \Box_1 \bigcirc_\gamma \varphi$

La preuve de ce théorème repose sur le lemme suivant :

**Lemme 1**  $\forall\langle W, R \rangle$  tel que  $R$  est euclidienne,  $\forall w \in W :$

$$\forall u, v \in R_w, R_u = R_v$$

**Démonstration :** [du lemme 1] Soit un modèle  $\langle W, R \rangle$  tel que  $R$  est euclidienne, un monde  $w$  dans ce modèle et  $u, v \in R_w$ . Comme  $wRu$  et  $wRv$ , alors par définition de l'euclidianité  $uRv$ . Ensuite, pour tout  $x \in R_u$ , la combinaison de  $uRv$  avec  $uRx$  implique que  $vRx$ , c'est-à-dire  $x \in R_v$ , et ainsi  $R_u \subseteq R_v$ . On montre de la même façon que  $R_v \subseteq R_u$ . ■

**Démonstration :** [du théorème 4] Soit  $\beta \in ]0, 1]$  et  $\gamma \in [0, 1]$  deux valeurs numériques. Soit un modèle  $\mathcal{M} = \langle \langle W, R \rangle, s \rangle$  avec  $R$  euclidienne, et soit  $w \in W$ . Comme  $R$  est euclidienne, on a  $\forall u, v \in R_w, R_u = R_v$  par le lemme 1.

Par conséquent, toute formule modale satisfaite dans un successeur de  $w$  est satisfaite dans tous ses autres successeurs, en raison des ensembles de mondes accessibles qui sont identiques pour tous ces mondes. Ainsi, toute proportion de successeurs de  $w$  est représentative de l'ensemble des successeurs. De ce fait, si  $w \models \Box_\beta \bigcirc_\gamma \varphi$  alors  $w \models \Box_1 \bigcirc_\gamma \varphi$ . ■

Ce théorème montre donc qu'il n'est pas intéressant de modifier la valeur de  $\beta$  ou d'augmenter la valeur de  $\gamma$  :  $\beta = 1$  peut être déduit de tout  $\beta$  si la relation est euclidienne.

La question qui se pose alors est de savoir si faire décroître la valeur de  $\alpha$  permet d'enrichir la conclusion, de la forme  $\Box_1 \diamond_\gamma \varphi$  avec  $\gamma < 1$ . Le théorème suivant répond par la négative :

**Théorème 5 ( $5_\alpha$ )**  $\forall \alpha \in [0, 1[, \forall \gamma \in [0, 1[,$  il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  euclidienne et  $w \in W$  tels que  $\mathcal{M}, w \models \diamond_\alpha \varphi$  mais  $\mathcal{M}, w \not\models \Box_1 \diamond_\gamma \varphi$ .

**Démonstration :** Nous construisons un tel modèle pour prouver le théorème. Le principe repose sur le fait qu'une relation euclidienne apporte une information riche : pour tout  $w' \in R_w, R_w \subseteq R_{w'}$ . Ainsi, il n'existe pas de garantie que les mondes validants appartiennent exclusivement à  $R_{w'}$  et non à  $R_w$ .

Plus précisément, étant donné  $\alpha$  et  $\gamma$  dans  $[0, 1[$ , soit  $n, m \in \mathbb{N}^*$  tels que  $\frac{m}{n+m} > 1 - \gamma$  et  $F = \langle W, R \rangle$  un cadre contenant  $n + m + 1$  mondes. Soit  $w$  un monde de référence dans  $W, R$  et  $s$  tels que :

- (i)  $|R_w| = n$
- (ii)  $\forall w' \in R_w, R_{w'} = W \setminus \{w\}$  et  $w' \models \varphi$
- (iii)  $\forall w' \notin R_w, R_{w'} = W \setminus (R_w \cup \{w\})$  et  $w' \models \neg \varphi$

Par construction,  $R$  est euclidienne.

On a  $\frac{|R_w(\varphi)|}{|R_w|} = 1 > 1 - \alpha$  et donc  $w \models \diamond_\alpha \varphi$ .

De plus,  $\forall w' \in R_w$ ,  $\frac{|R_{w'}(\varphi)|}{|R_{w'}|} = \frac{n}{n+m} \leq 1 - \gamma$ . Ainsi  $w' \not\models \diamond_\gamma \varphi$  et  $w \not\models \square_1 \diamond_\gamma \varphi$ . ■

Pour améliorer ce résultat et proposer une valuation pertinente, il est possible d'ajouter des contraintes sur la relation. Néanmoins l'influence d'un tel changement sur la relation, et donc le cadre, est conséquente. Par exemple, imposer la transitivité force la relation entre un monde et les successeurs de ses propres successeurs. Plus formellement, par transitivité et euclidianité, on a le lemme suivant :

**Lemme 2**  $\forall \langle W, R, s \rangle$ , tel que  $R$  est transitive et euclidienne :

$$\forall w \in W, \forall w' \in R_w, R_{w'} = R_w$$

**Démonstration :** Soit un modèle  $\langle W, R, s \rangle$  tel que  $R$  est transitive et euclidienne, et soit  $w \in W$ . La preuve de ce lemme découle directement de la définition des deux propriétés :

- si  $R$  est transitive :  $\forall v \in R_w, R_v \subseteq R_w$
- si  $R$  est euclidienne :  $\forall v \in R_w, R_w \subseteq R_v$

■

Ainsi, si un monde  $w$  valide  $\diamond_\alpha \varphi$ , tous les mondes  $w' \in R_w$  valident également cette formule, donc  $w \models \square_1 \diamond_\alpha \varphi$ .

Toutefois, un tel choix signifie que l'on examine, par rapport à l'axiome classique, une éventuelle variante dans laquelle on impose des contraintes additionnelles à la fois sur la prémisse, de la forme  $\diamond_\alpha \varphi$  avec  $\alpha < 1$ , et sur les cadres de Kripke considérés, de la forme  $R$  euclidienne *et* transitive.

De ce fait, et bien que cette éventuelle variante puisse peut-être conduire à une conclusion plus riche  $\square_1 \diamond_\beta \varphi$ , avec  $\beta < 1$ , elle réduit sévèrement le pouvoir de satisfiabilité de l'axiome, dont les conditions d'application semblent trop restrictives et difficile à valider. Par conséquent, nous excluons cette proposition.

### 3.5 Type II : axiomes pondérés avec correspondance classique

Les axiomes appartenant au type II donnent une forme relâchée de leur version classique et peuvent être établis selon la même contrainte sémantique : la contrainte classiquement

associée est préservée et suffisante pour donner aux poids des valeurs pertinentes. Ce principe s'applique tout particulièrement au cas des axiomes qui sont associés à une relation non contrainte, l'axiome (K) et les axiomes normaux, mais aussi l'axiome (D), comme discuté successivement dans les sections suivantes.

### 3.5.1 Axiome (K)

La variante pondérée de l'axiome (K) prend la forme suivante :

**Théorème 6** ( $K_\alpha$ )  $\forall \alpha, \beta \in [0, 1]$

$$\models \Box_\alpha(\varphi \rightarrow \psi) \rightarrow (\Box_\beta\varphi \rightarrow \Box_\gamma\psi) \quad \text{avec } \gamma = \max(0, \alpha + \beta - 1)$$

**Démonstration :** Soit  $F = \langle W, R \rangle$  un cadre,  $s$  une valuation et  $w \in W$ .

Si  $R_w = \emptyset$ ,  $w$  valide trivialement les trois formules modales de la formule, qui est valide.

Si  $|R_w| > 0$ , la preuve consiste à appliquer le *modus ponens* dans les mondes pour lesquels  $\varphi \rightarrow \psi$  et  $\varphi$  sont toutes les deux validées :  $R_w(\varphi \rightarrow \psi) \cap R_w(\varphi) \subseteq R_w(\psi)$ .

Par définition du cardinal de l'intersection ensembliste :

$$|R_w(\varphi \rightarrow \psi) \cap R_w(\varphi)| = |R_w(\varphi \rightarrow \psi)| + |R_w(\varphi)| - |R_w(\varphi \rightarrow \psi) \cup R_w(\varphi)|$$

Comme  $|R_w| \geq |R_w(\varphi \rightarrow \psi) \cup R_w(\varphi)|$ , on a :

$$\begin{aligned} |R_w(\psi)| &\geq |R_w(\varphi \rightarrow \psi) \cap R_w(\varphi)| \\ &\geq |R_w(\varphi \rightarrow \psi)| + |R_w(\varphi)| - |R_w| \end{aligned}$$

Donc  $\frac{|R_w(\psi)|}{|R_w|} \geq \alpha + \beta - 1$ . ■

### 3.5.2 Axiomes normaux

Les axiomes normaux offrent des règles de manipulation de la modalité pour la factoriser et la distribuer sur la conjonction et la disjonction. Nous les étudions selon l'opérateur logique qu'ils manipulent.

**Conjonction : distribution et factorisation** L'axiome classique (NA1) s'écrit  $\vdash \Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi)$ . La version pondérée de cet axiome est donnée dans le théorème 7.

**Théorème 7** ( $NA1_\alpha$ )

$$\begin{aligned} \forall \alpha \in [0, 1], \quad &\models \Box_\alpha(\varphi \wedge \psi) \rightarrow (\Box_\alpha\varphi \wedge \Box_\alpha\psi) \\ \forall \beta, \gamma \in [0, 1], \quad &\models (\Box_\beta\varphi \wedge \Box_\gamma\psi) \rightarrow \Box_\alpha(\varphi \wedge \psi) \quad \text{avec } \alpha = \max(\beta + \gamma - 1, 0) \end{aligned}$$

**Démonstration :** Soit  $F = \langle W, R \rangle$  un cadre,  $s$  une valuation et  $w \in W$ . Si  $R_w = \emptyset$ ,  $w$  valide trivialement chaque implication.

Si  $w \models \Box_\alpha(\varphi \wedge \psi)$ , alors comme  $R_w(\varphi \wedge \psi) = R_w(\varphi) \cap R_w(\psi) \subseteq R_w(\varphi)$ , on obtient

$$\alpha \leq \frac{|R_w(\varphi \wedge \psi)|}{|R_w|} \leq \frac{|R_w(\varphi)|}{|R_w|}$$

ce qui implique que  $w \models \Box_\alpha \varphi$ . De la même manière, comme,  $R_w(\varphi \wedge \psi) \subseteq R_w(\psi)$ , on a  $w \models \Box_\alpha \psi$ . Ainsi  $w \models \Box_\alpha \varphi \wedge \Box_\alpha \psi$ .

Supposons que  $w \models \Box_\beta \varphi \wedge \Box_\gamma \psi$ . En utilisant la définition du cardinal de l'intersection ensembliste, on a :

$$\begin{aligned} \frac{|R_w(\varphi \wedge \psi)|}{|R_w|} &= \frac{|R_w(\varphi) \cap R_w(\psi)|}{|R_w|} = \frac{|R_w(\varphi)| + |R_w(\psi)| - |R_w(\varphi) \cup R_w(\psi)|}{|R_w|} \\ &\geq \frac{|R_w(\varphi)|}{|R_w|} + \frac{|R_w(\psi)|}{|R_w|} - \frac{|R_w|}{|R_w|} \\ &\geq \beta + \gamma - 1 \end{aligned}$$

Cela implique que  $w \models \Box_\alpha(\varphi \wedge \psi)$  avec  $\alpha = \max(\beta + \gamma - 1, 0)$ . ■

**Factorisation de la disjonction** L'axiome (NA2) traite du cas disjonctif, permettant de manipuler les modalités par factorisation :  $\vdash (\Box\varphi \vee \Box\psi) \rightarrow \Box(\varphi \vee \psi)$ . L'extension pondérée de cet axiome est donnée par le théorème 8.

**Théorème 8 (NA2<sub>αf</sub>)**  $\forall \alpha, \beta \in [0, 1]$

$$\models (\Box_\alpha \varphi \vee \Box_\beta \psi) \rightarrow \Box_\gamma (\varphi \vee \psi) \quad \text{avec } \gamma = \min(\alpha, \beta)$$

**Démonstration :** Soit un cadre  $F = \langle W, R \rangle$ ,  $s$  une valuation et  $w \in W$ . Si  $R_w = \emptyset$ ,  $w$  valide trivialement les trois formules modales et donc l'implication.

Si  $|R_w| > 0$ , la preuve découle des faits suivants :  $R_w(\varphi \vee \psi) = R_w(\varphi) \cup R_w(\psi)$ ,  $R_w(\varphi) \subseteq R_w(\varphi) \cup R_w(\psi)$  et  $R_w(\psi) \subseteq R_w(\varphi) \cup R_w(\psi)$ .

Si  $w \models \Box_\alpha \varphi \vee \Box_\beta \psi$ , deux cas peuvent être distingués : soit  $w \models \Box_\alpha \varphi$  et donc  $\frac{|R_w(\varphi) \cup R_w(\psi)|}{|R_w|} \geq \frac{|R_w(\varphi)|}{|R_w|} \geq \alpha$ ; soit  $w \models \Box_\beta \psi$ , et, de manière identique,  $\frac{|R_w(\varphi) \cup R_w(\psi)|}{|R_w|} \geq \beta$ .

Par conséquent, dans les deux cas,  $\frac{|R_w(\varphi) \cup R_w(\psi)|}{|R_w|} \geq \min(\alpha, \beta)$ . ■

**Distribution de la disjonction** La réciproque, qui correspond à une distribution de l'opérateur  $\vee$  n'est pas valide en logique modale :  $\not\models \Box(\varphi \vee \psi) \rightarrow (\Box\varphi \vee \Box\psi)$ . Néanmoins, nous choisissons de nous intéresser à une version pondérée de cette formule, qui peut s'écrire  $\Box_\alpha(\varphi \vee \psi) \rightarrow (\Box_\beta\varphi \vee \Box_\gamma\psi)$ . En effet, un tel axiome permet de discuter le triplet  $(\alpha, \beta, \gamma)$ , dans une perspective de caractérisation de la relation entre ces trois valeurs. Une forme générale peut, dans cette optique, s'écrire :

**Théorème 9 (NA2 $_\alpha$ d)**

$$\models \Box_\alpha(\varphi \vee \psi) \rightarrow (\Box_\beta\varphi \vee \Box_\gamma\psi) \quad \text{avec } \alpha \leq \beta + \gamma$$

**Démonstration :** Soit un cadre  $F = \langle W, R \rangle$ ,  $s$  une valuation et un monde  $w \in W$ . Si  $R_w = \emptyset$ ,  $w$  valide toute formule modale par définition, et donc l'implication.

Sinon, la preuve se base sur les propriétés ensemblistes de cardinalité :

$$|R_w(\varphi \vee \psi)| = |R_w(\varphi) \cup R_w(\psi)| = |R_w(\varphi)| + |R_w(\psi)| - |R_w(\varphi) \cap R_w(\psi)|$$

ce qui induit que  $\alpha \leq \beta + \gamma$  en divisant les deux côtés de l'égalité par  $|R_w|$ .

■

Néanmoins, on ne peut garantir une valuation plus précise des poids : la distribution de la disjonction nécessite de répartir  $\alpha$  entre la valeur de  $\beta$  et celle de  $\gamma$ . Il n'existe en effet aucun moyen pour décider quelle quantité revient auquel de ces degrés.

### 3.5.3 Axiome (D)

L'extension pondérée de l'axiome (D), qui correspond classiquement à une relation sérielle, permet de mettre en relation les deux modalités pondérées,  $\Box_\alpha$  et  $\Diamond_\alpha$ , complétant la propriété donnée par la Proposition 2.4.3, p. 34, en traitant sa réciproque :

**Théorème 10 (D $_\alpha$ )**  $\forall F = \langle W, R \rangle$ ,

*R est sérielle*

$$\Leftrightarrow \forall \alpha \in ]0, 1] \text{ et } \forall \varepsilon \in ]0, \alpha] F \models \Box_\alpha\varphi \rightarrow \Diamond_{1-\alpha+\varepsilon}\varphi$$

**Démonstration :** Soit  $F = \langle W, R \rangle$  un cadre avec  $R$  sérielle. Soit  $s$  une valuation,  $w \in W$  et  $\alpha \in ]0, 1]$ . Si  $w \models \Box_\alpha\varphi$ , comme  $R$  est sérielle, alors on a  $R_w \neq \emptyset$  et  $\frac{|R_w(\varphi)|}{|R_w|} \geq \alpha$ . De ce fait,  $\frac{|R_w(\varphi)|}{|R_w|} > \alpha - \varepsilon$  pour tout  $\varepsilon$  positif. Cela implique que  $w \models \Diamond_{1-\alpha+\varepsilon}\varphi$ .

Réciproquement, si on suppose que (D $_\alpha$ ) est valide pour tout  $\alpha$  et tout  $\varepsilon$ , la relation est nécessairement sérielle. En effet, en appliquant la théorie de la correspondance classique, si  $R$  n'est pas sérielle, l'axiome est contredit par la valeur  $\alpha = 1 + \varepsilon$ .



Il faut toutefois noter que, la valeur  $\alpha = 0$  étant non informative, en raison de la sémantique associée comme détaillé en section 3.2, nous n'en tenons pas compte. Le cas où  $\alpha$  constitue un cas trivial pour lequel aucune valeur de  $\varepsilon$  ne peut être considéré.

Cette remarque nous permet de considérer que la version pondérée de cet axiome en améliore l'informativité : même dans le cas où l'informativité de la prémisse est réduite, c'est-à-dire pour un monde  $w$  tel que  $w \models \Box_\alpha \varphi$  avec  $\alpha < 1$ , la conclusion n'en est pas moins riche, au contraire  $w \models \Diamond_{1-\alpha+\varepsilon} \varphi$  est plus informative que le  $\Diamond$  classique. Cela est lié au fait que les modalités classiques ont un pouvoir d'expressivité plus faible que les modalités pondérées, car elle ne portent aucune précision sur la quantité de mondes validant ou invalidant une formule.

### 3.6 Type III : axiomes pondérés sans correspondance

Les axiomes appartenant à cette troisième catégorie sont définis comme pondérés sans correspondance, c'est-à-dire que s'ils permettent d'assouplir la modalité, c'est au détriment de la théorie de la correspondance : l'axiome pondéré n'est ne suffit pas à imposer aux cadres de Kripke de respecter les contraintes liées. En effet, malgré l'ajout de contraintes nécessaire à l'établissement des variantes pondérées des axiomes, l'équivalence n'est pas maintenue. Nous présentons successivement le cas de l'axiome (C) et celui de l'axiome (4).

#### 3.6.1 Axiome (C)

L'axiome (C)  $\vdash \Diamond \Box \varphi \rightarrow \Box \Diamond \varphi$ , classiquement en correspondance avec une relation convergente, contient une double combinaison de modalité en prémisse et en conclusion. L'extension pondérée de cet axiome comporte donc quatre poids et s'écrit :  $\Diamond_\alpha \Box_\beta \varphi \rightarrow \Box_{\beta'} \Diamond_{\alpha'} \varphi$ .

Tout d'abord, afin d'affaiblir la contrainte de la prémisse, le cas où  $\alpha = 1$  est considéré : ainsi, la seconde modalité peut être pondérée de manière pertinente, ce qui enrichit l'axiome par rapport au cas classique tout en conservant la contrainte classique. Cela fait une différence avec l'axiome (4) discuté par la suite, pour lequel la contrainte sur la relation doit être renforcée. Le théorème 11 présente cet axiome pondéré.

**Théorème 11**  $(C_\alpha) \forall F = \langle W, R \rangle,$

*R est convergente*

$\Rightarrow \forall s, \forall w \in W$  avec  $w$  non aveugle,  $\forall \beta \in ]0, 1], \forall \varepsilon \in ]0, \beta]$

$\langle W, R, s \rangle, w \models \Diamond_1 \Box_\beta \varphi \rightarrow \Box_{\frac{1}{|R_w|}} \Diamond_{1-\beta+\varepsilon} \varphi$

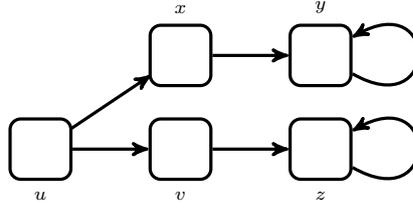


FIGURE 3.1 – Exemple de cadre pour montrer que la réciproque du théorème 11 n'est pas vraie, comme établi dans le théorème 12.

Notons que si  $w$  est aveugle, alors on a  $w \models \Box_\alpha \psi$  pour toutes les valeurs  $\alpha$  et toutes les formules  $\psi$ , en particulier  $\psi = \Diamond_{1-\beta+\varepsilon} \varphi$ . De ce fait,  $w$  valide l'implication du théorème, bien que la valeur du poids  $\frac{1}{|R_w|}$  ne soit pas définie.

**Démonstration :** Soit un modèle  $\mathcal{M} = \langle W, R, s \rangle$  tel que  $R$  est convergente. Soit un monde  $w \in W$  non aveugle,  $\beta \in ]0, 1]$ ,  $\varepsilon \in ]0, \beta]$ .

Si  $w \models \Diamond_1 \Box_\beta \varphi$ , alors  $R_w \neq \emptyset$  et  $\exists w' \in R_w$  tel que  $w' \models \Box_\beta \varphi$ .

Comme  $R$  est convergente,  $R_{w'} \neq \emptyset$  : en considérant deux fois  $wRw'$  et  $wRw'$ , il existe  $x$  tel que  $w'R_x$ . Comme  $\beta > 0$  et  $\varepsilon \in ]0, \beta]$ , par une preuve similaire à celle utilisée pour le Théorème 10 ( $D_\alpha$ ) page 53, en utilisant une hypothèse de sérialité locale, ce qui signifie qu'aucun successeur n'est aveugle,  $w' \models \Box_\beta \varphi$  implique que  $w' \models \Diamond_{1-\beta+\varepsilon} \varphi$ .

De cette façon,  $R_w(\Diamond_{1-\beta+\varepsilon} \varphi) \neq \emptyset$ , donc  $w \models \Box_{\frac{1}{|R_w|}} \Diamond_{1-\beta+\varepsilon} \varphi$ .

■

Néanmoins, la réciproque n'est pas vraie : si l'on suppose l'axiome ( $C_\alpha$ ), la relation n'est pas nécessairement convergente. Ce résultat est donné par le théorème :

**Théorème 12 ( $C_\alpha$ )** *Pour tout  $\varphi$ , il existe un cadre  $F = \langle W, R \rangle$  dont la relation  $R$  n'est pas convergente et tel que  $\forall \beta \in [0, 1], \forall \varepsilon \in ]0, \beta]$ , pour tout monde  $w \in W$ ,  $F \models \Diamond_1 \Box_\beta \varphi \rightarrow \Box_{\frac{1}{|R_w|}} \Diamond_{1-\beta+\varepsilon} \varphi$ .*

**Démonstration :** La figure 3.1 donne un exemple de tel cadre, pour lequel  $W = \{u, v, z, x, y\}$  : la relation  $R$  n'est pas convergente puisqu'il n'existe pas de successeur commun aux deux mondes  $x$  et  $v$ . Soit  $\varphi$  une formule quelconque.

Soit une valuation  $s$  et deux poids  $\beta$  et  $\varepsilon$  dont les valeurs sont telles que  $\beta \in ]0, 1]$  et  $\varepsilon \in ]0, \beta]$ . Soit la formule  $\psi = \Diamond_1 \Box_\beta \varphi \rightarrow \Box_{\frac{1}{|R_u|}} \Diamond_{1-\beta+\varepsilon} \varphi$  qui reprend la formule considérée. La démonstration qui suit montre que  $F \models \psi$  en considérant chaque monde successivement.

Si  $u \not\models \Diamond_1 \Box_\beta \varphi$ , alors  $u \models \psi$ . Dans le cas contraire, comme  $R_u = \{x, v\}$ , on a soit  $x \models \Box_\beta \varphi$  soit  $v \models \Box_\beta \varphi$ . Dans le premier cas, on montre que  $x \models \Diamond_{1-\beta+\varepsilon} \varphi$ , par une preuve similaire à celle du théorème 10 ( $D_\alpha$ ) page 53, en considérant que la relation d'accessibilité est localement sérielle. On en déduit que  $u \models \Box_{\frac{1}{|R_u|}} \Diamond_{1-\beta+\varepsilon} \varphi$ . Dans le second cas, on a de même  $v \models \Diamond_{1-\beta+\varepsilon} \varphi$  et donc  $u \models \Box_{\frac{1}{|R_u|}} \Diamond_{1-\beta+\varepsilon} \varphi$ .

Si  $x \not\models \Diamond_1 \Box_\beta \varphi$ , alors  $x \models \psi$ . Dans le cas contraire, comme  $y$  est le seul successeur de  $x$ , alors  $y \models \Box_\beta \varphi$ . On a ainsi, par la même preuve que pour le Théorème 10 ( $D_\alpha$ ) page 53, mais en considérant une sérialité locale pour la relation,  $y \models \Diamond_{1-\beta+\varepsilon} \varphi$ . Donc  $x \models \Box_{\frac{1}{|R_x|}} \Diamond_{1-\beta+\varepsilon} \varphi$ .

La même preuve peut être donnée pour le monde  $v$  dont le seul successeur est  $z$ , tout comme  $y$  et  $z$  qui sont leur propre successeur. ■

### 3.6.2 Axiome (4)

L'axiome (4), qui s'écrit  $\vdash \Box \varphi \rightarrow \Box \Box \varphi$ , est classiquement équivalent à la transitivité de la relation. Nous proposons d'en étudier l'extension pondérée :  $\vdash \Box_\alpha \varphi \rightarrow \Box_\beta \Box_\gamma \varphi$ . Le problème qui se pose est donc de déterminer les valeurs appropriées pour  $\beta$  et  $\gamma$ , étant donné une valeur de  $\alpha$ . Le théorème ci-dessous montre que la seule condition de transitivité de  $R$  n'est pas suffisante pour que les valeurs des poids soit pertinentes.

**Théorème 13**  $\forall \alpha \in [0, 1[$ , il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  transitive et  $w \in W$  tel que  $\mathcal{M}, w \models \Box_\alpha \varphi$  et  $\mathcal{M}, w \not\models \Box_1 \Box_1 \neg \varphi$ .

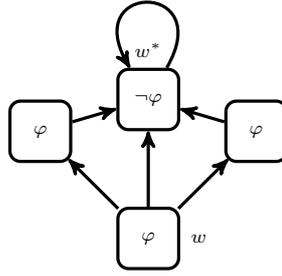
**Démonstration :** La démonstration consiste à construire un tel modèle  $\mathcal{M}$ . Étant donné  $\alpha < 1$ , soit  $m, q \in \mathbb{N}^*$  tels que  $\alpha \leq m/(m+q)$  (un tel couple existe car  $\alpha \neq 1$ ). Soit  $W$  un ensemble de  $1+q+m$  mondes et  $w \in W$  un monde de référence. Enfin, soit  $R$  la relation binaire définie sur  $W \times W$  et  $s$  la valuation telles que :

- (i)  $R_w = W \setminus \{w\}$
- (ii)  $|R_w(\varphi)| = m$
- (iii)  $|R_w(\neg \varphi)| = q$ , on note  $w_n$  un monde de  $R_w(\neg \varphi)$ .
- (iv)  $\forall u \in R_w, R_u = \{w_n\}$

Par construction,  $R$  est transitive. Un tel modèle est illustré par la figure 3.2 pour  $\alpha = 2/3$ , avec  $m = 2$  et  $q = 1$ .

En notant  $\mathcal{M} = \langle W, R, s \rangle$ , on a  $\mathcal{M}, w \models \Box_{m/(m+q)} \varphi$ . Ainsi, par gradualité décroissante,  $\mathcal{M}, w \models \Box_\alpha \varphi$ .

De plus, comme  $\forall u \in R_w, \mathcal{M}, u \not\models \Box_1 \neg \varphi$ , on a  $\mathcal{M}, w \not\models \Box_1 \Box_1 \neg \varphi$ .

FIGURE 3.2 – Modèle illustrant la démonstration du théorème 13 pour  $\alpha = \frac{2}{3}$ 

■

Ce théorème induit que pour  $\gamma > 0$  il n'existe pas de  $\beta > 0$  tel que  $\mathcal{M}, w \models \Box_\beta \Box_\gamma \varphi$ .

Par conséquent, la transitivité n'est pas une condition suffisante pour avoir des garanties sur les valeurs de  $\beta$  et  $\gamma$ . Il est donc nécessaire de renforcer la contrainte de la relation en ajoutant une nouvelle propriété : le théorème 14 ci-dessous considère le cas de l'eucledienne. En effet, cette propriété permet d'éviter l'apparition de mondes puits, comme ceux appartenant à  $R_w(\neg\varphi)$  dans la démonstration précédente. La transitivité est conservée afin de préserver la compatibilité avec le cas classique, obtenu lorsque les poids sont égaux à 1. Cela amène au théorème :

**Théorème 14** ( $4_\alpha$ )  $\forall F = \langle W, R \rangle$ ,

*R est transitive et euclidienne*

$\Rightarrow \forall \alpha \in [0, 1], F \models \Box_\alpha \varphi \rightarrow \Box_1 \Box_\alpha \varphi$

**Démonstration :** Cette démonstration repose sur le fait que, pour une relation euclidienne et transitive, le modèle possède la propriété  $\forall w' \in R_w, R_{w'} = R_w$  (voir lemme 2 p. 50). De ce fait, pour toute valuation  $s$  et pour tout  $\alpha \in [0, 1]$ , si  $w \models \Box_\alpha \varphi$ , alors tous les mondes accessibles  $w' \in R_w$  valident également  $w' \models \Box_\alpha \varphi$ . Par conséquent, on a  $w \models \Box_1 \Box_\alpha \varphi$

■

L'axiome alors obtenu est particulièrement fort, étant donné la valeur du poids de la première modalité en conclusion, qui est maximale. La seconde modalité de la conclusion, quant à elle, porte la valeur  $\alpha$  qui correspond au poids donné en prémisse de l'implication. Cette valeur  $\alpha$  est pertinente et ne peut qu'être diminuée par gradualité décroissante. De fait, une version de cet axiome dont la dernière modalité pondérée est d'une valeur plus

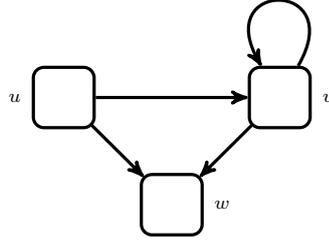


FIGURE 3.3 – Exemple de cadre qui contredit la réciproque du théorème 14, comme montré par le théorème 15.

élevée ne peut pas être considéré, ce qui signifie qu'il n'y a pas d'amélioration possible du pouvoir de manipulation de la version proposée.

Il est important de noter qu'à l'inverse du cas de l'axiome (5), discuté en section 3.4.4, p. 48, la contrainte de la relation est renforcée mais la prémisse de l'implication est relâchée : elle ne nécessite que  $\Box_\alpha\varphi$  au lieu de  $\Box_1\varphi$ .

Néanmoins, la réciproque n'est pas vraie :

**Théorème 15** ( $4_\alpha$ ) *Il existe un cadre  $F = \langle W, R \rangle$  avec une relation  $R$  non euclidienne tel que pour toute valuation  $s$  et  $\forall \alpha \in [0, 1]$ ,  $\langle F, s \rangle \models \Box_\alpha\varphi \rightarrow \Box_1\Box_\alpha\varphi$*

**Démonstration :** La démonstration construit un tel cadre, illustré sur la figure 3.3 qui représente un exemple de modèle avec  $W = \{u, v, w\}$  : la relation est transitive mais non euclidienne. En effet, la relation  $wRv$ , qui est requise car on a  $uRw$  et  $uRv$ , n'est pas présente.

Soit  $s$  une valuation quelconque et  $\alpha \in [0, 1]$ , notons  $\psi = \Box_\alpha\varphi \rightarrow \Box_1\Box_\alpha\varphi$ . Dans cette preuve, chaque monde est considéré successivement. Tout d'abord, nous examinons le monde  $v$  : si  $v \not\models \Box_\alpha\varphi$ ,  $v \models \psi$ . Dans le cas contraire, comme  $R_v = \{v, w\}$  et  $w \models \Box_\alpha\varphi$  car  $w$  est aveugle, tous les successeurs de  $v$  valident la formule  $\Box_\alpha\varphi$ . De ce fait,  $v \models \Box_1\Box_\alpha\varphi$  et donc  $v$  valide  $\psi$ .

Le cas du monde  $u$  est identique : si  $u \not\models \Box_\alpha\varphi$ , alors  $u \models \psi$ . Au contraire, comme  $R_u = R_v$ ,  $u \models \Box_\alpha\varphi$  induit que  $v \models \Box_\alpha\varphi$ . De plus,  $w \models \Box_\alpha\varphi$  car  $w$  est aveugle. Ainsi tous les successeurs de  $u$  valident  $\Box_\alpha\varphi$ , et donc  $u \models \psi$ .

Enfin, puisque  $w$  est aveugle,  $w$  valide toutes les formules modales de l'implication.

■

### 3.7 Type IV : axiomes pondérés avec correspondance enrichie

Les axiomes pondérés de type IV sont définis comme des extensions pour lesquelles une contrainte modifiée peut être considérée. La différence avec le type III réside dans le fait que, pour le type IV, la correspondance peut être établie. Deux types de modification sont envisagés : la propriété classique peut être supprimée, menant à une relation moins contrainte, ou de nouvelles propriétés peuvent être ajoutées afin d'enrichir la contrainte.

#### 3.7.1 Axiome (CD)

L'axiome (CD), classiquement en correspondance avec la fonctionnalité de la relation, établit que  $\vdash \Diamond\varphi \rightarrow \Box\varphi$ .

Les versions pondérées de cet axiome sont établies par les propositions données en section 2.4.3, p. 34. En effet, conformément à la sémantique proposée, la proposition 2.4.3 permet d'établir  $\models \Diamond_\alpha\varphi \rightarrow \Box_{1-\alpha}\varphi$ . Comme cette version ne nécessite pas de contrainte sur la relation, la correspondance est trivialement établie.

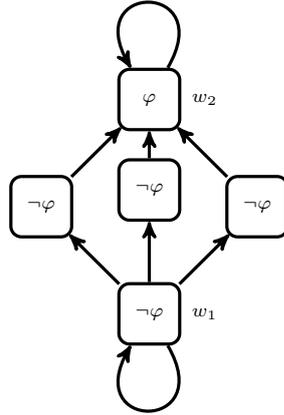
La deuxième proposition de la section 2.4.3 indique que l'on a  $\models \Diamond_\alpha\varphi \rightarrow \Box_{\frac{1}{|R_w|}}\varphi$  pour toute relation et tous les mondes  $w$  (si  $w$  est aveugle, le coefficient  $1/|R_w|$  n'est pas défini, mais  $w$  valide toute formule modale  $\Box_\alpha\psi$ ). Comme établi, le cas où  $\alpha = 1$  donne un résultat bien plus informatif : il rejoint l'axiome classique dans lequel la modalité en conclusion est équivalente à  $\Box_1$ . L'axiome établi dans cette section, cependant, pour  $\alpha = 1$ , indique que  $\Diamond_1\varphi \rightarrow \Box_0\varphi$ . Cela illustre les différents facteurs d'influence des pondérations sur les axiomes classiques : si l'ajout d'un poids peut assouplir un axiome, le cas présenté ci-dessus montre qu'un axiome trop assoupli perd en information. En effet  $\Diamond_1$  est moins contraint que  $\Diamond_\alpha$  mais induit une conclusion trop pauvre.

Notons que dans les deux cas cités, renforcer la prémisse, imposant d'avoir  $\Diamond_\alpha\varphi$  au lieu de  $\Diamond\varphi$ , permet de supprimer la propriété de fonctionnalité de la relation : si  $\alpha \neq 1$  alors il existe nécessairement plus d'un successeur au monde de référence, ce qui est incompatible avec la fonctionnalité.

#### 3.7.2 Axiome (C4)

L'axiome ( $C4_\alpha$ ) illustre le cas d'un renforcement de la contrainte de la relation : dans le cas classique, cet axiome, noté  $\vdash \Box\Box\varphi \rightarrow \Box\varphi$ , est équivalent à la densité de la relation. Nous proposons d'en donner la forme pondérée  $\Box_\alpha\Box_\beta\varphi \rightarrow \Box_\gamma\varphi$ .

Comme exprimé par le théorème ci-dessous, la densité seule ne suffit pas à garantir des valeurs pertinentes pour les poids. En effet, pour toute valeur de  $\alpha$  et  $\beta$ , on peut construire un modèle pour lequel  $\gamma = 0$ .

FIGURE 3.4 – Contre-exemple pour le théorème 16 pour  $\alpha = \frac{3}{4}$ 

**Théorème 16**  $\forall \alpha \in [0, 1[, \forall \beta \in [0, 1]$ , il existe un modèle  $\mathcal{M} = \langle W, R, s \rangle$  avec  $R$  dense et  $w \in W$  tels que  $\mathcal{M}, w \models \Box_\alpha \Box_\beta \varphi$  et  $\mathcal{M}, w \models \Box_1 \neg \varphi$ .

**Démonstration :** Dans cette preuve, nous proposons de construire un tel modèle. Étant donné  $\alpha \in [0, 1[$  et  $\beta \in [0, 1]$ , soit  $n \in \mathbb{N}$  tel que  $n \geq \alpha/(1 - \alpha)$ . Soit  $W$  un ensemble de  $n + 2$  mondes,  $w_1$  et  $w_2$  deux mondes distincts de  $W$ ,  $R$  la relation et  $s$  la valuation telles que :

- (i)  $\mathcal{M}, w_2 \models \varphi$
- (ii)  $\forall w \in W \setminus \{w_2\}, \mathcal{M}, w \models \neg \varphi$
- (iii)  $R_{w_1} = W \setminus \{w_2\}$
- (iv)  $\forall w \in W \setminus \{w_1\}, R_w = \{w_2\}$

Par construction,  $R$  est dense. La figure 3.4 donne une illustration de ce contre-exemple, dans le cas où  $\alpha = 0.75$ .

Ainsi  $\forall w \in W \setminus \{w_1\}$ , on a  $\mathcal{M}, w \models \Box_1 \varphi$ , qui implique, par gradualité décroissante,  $\mathcal{M}, w \models \Box_\beta \varphi$ . De ce fait :

$$\begin{aligned} |R_{w_1}(\Box_\beta \varphi)| = n &\Rightarrow \frac{|R_{w_1}(\Box_\beta \varphi)|}{|R_{w_1}|} = \frac{n}{n+1} \geq \alpha \\ &\Rightarrow \mathcal{M}, w_1 \models \Box_\alpha \Box_\beta \varphi \end{aligned}$$

Toutefois,  $\mathcal{M}, w_1 \models \Box_1 \neg \varphi$ , ce qui induit que  $\nexists \gamma > 0$  tel que  $\mathcal{M}, w_1 \models \Box_\gamma \varphi$ .

■

Par conséquent, pour garantir une valeur strictement positive pour  $\gamma$ , nous considérons des contraintes additionnelles sur la relation. dans le tableau ?? p. ??, Deux ensembles de

propriétés peuvent être ajoutés, menant à deux extensions pondérées possibles pour  $(C4_\alpha)$ . Ces deux versions diffèrent par l'informativité de leur conclusion et la contrainte de leur prémisse, ces éléments étant nécessairement corrélés en raison de la succession de modalités qui apparaît dans la forme de l'axiome, comme discuté dans la section 3.2.2, p. 40.

La version la plus contrainte est établie dans le théorème ci-dessous. Nous soulignons que la compatibilité avec le cas classique est préservée, la densité étant conséquence de l'eulidianité.

**Théorème 17**  $(C4_\alpha) \forall F = \langle W, R \rangle,$

*R est transitive et euclidienne*

$$\Leftrightarrow \forall \alpha \in ]0, 1], \forall \beta \in [0, 1], F \models \Box_\alpha \Box_\beta \varphi \rightarrow \Box_\beta \varphi$$

**Démonstration :** Soit  $\langle W, R \rangle$  tel que  $R$  est transitive et euclidienne, soit  $s$  une valuation et  $\alpha, \beta \in ]0, 1]$  (si  $\beta = 0$  alors  $\Box_\beta \varphi$  est valide, et donc l'implication l'est aussi). Enfin, soit un monde  $w \in W$ .

Si  $R_w = \emptyset$ ,  $w$  valide trivialement toutes les formules de l'axiome et donc l'implication. Pour cette raison, on suppose par la suite que  $|R_w| > 0$ .

Si  $w \not\models \Box_\alpha \Box_\beta \varphi$ , alors  $w$  valide l'implication.

Si  $w \models \Box_\alpha \Box_\beta \varphi$ , alors une proportion  $\alpha > 0$  de mondes accessibles depuis  $w$  valide  $\Box_\beta \varphi$ . Soit  $u$  un tel monde :  $\frac{|R_u(\varphi)|}{|R_u|} \geq \beta$ . Comme  $R$  est transitive et euclidienne, on a  $\forall w' \in R_w, R_{w'} = R_w$  (cf. lemme 2, p. 50). En particulier,  $R_u = R_w$  et donc  $w \models \Box_\beta \varphi$ .

La réciproque est montrée par contraposée : pour tout cadre  $\langle W, R \rangle$ , si la relation  $R$  n'est pas transitive ou n'est pas euclidienne, il existe  $\alpha, \beta$  ainsi qu'une valuation  $s$  et  $w \in W$  tels que  $w \not\models \Box_\alpha \Box_\beta \varphi \rightarrow \Box_\beta \varphi$ .

Soit  $W$  un ensemble fini de mondes et  $R$  une relation non transitive (il existe donc  $u, v, w \in W$  tels que  $uRv \wedge vRw \wedge \neg uRw$ ) ou non euclidienne (dans ce cas aussi, il existe  $u, v$  et  $w$  sont aussi tels que  $vRu \wedge vRw \wedge \neg uRw$ ). On peut traiter similairement les deux cas qui sont identiques. Soit  $\alpha = \frac{1}{|R_u|}$  et  $\beta = \frac{1}{|R_v|}$ . Enfin, soit une valuation  $s$  telle que :

- (i)  $w \models \varphi$
- (ii)  $\forall x \in W \setminus \{w\}, x \not\models \varphi$ .

La figure 3.5 donne un exemple d'un modèle.

Les assertions suivantes sont donc vraies :

- $u \models \Box_1 \neg \varphi$  car le seul monde validant  $\varphi$ ,  $w$ , n'est pas accessible depuis  $u$ . Cependant  $u \not\models \Box_\beta \varphi$ .
- $v \models \Box_\beta \varphi$  parce que  $w \in R_v$  et  $w \models \varphi$
- $u \models \Box_\alpha \Box_\beta \varphi$  parce que  $v \in R_u$  et  $v \models \Box_\beta \varphi$

■

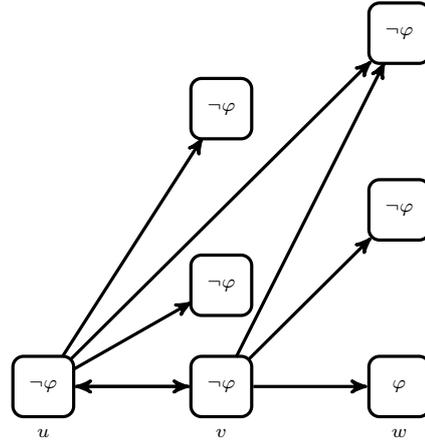


FIGURE 3.5 – Cadre illustrant la preuve du théorème 17.

Nous soulignons le fait que sans la transitivité et l'euclidianité, d'autres valeurs pertinentes sont possibles, mais s'appuient sur une valeur plus élevée en prémisse : le prochain théorème traite du cas de la shift-réflexivité (qui induit la densité). La prémisse de l'implication devient ici de la forme  $\Box_\alpha \Box_1 \varphi$ , ce qui en fait une version plus riche que le cas général  $\Box_\alpha \Box_\beta \varphi$  considéré dans le théorème 17.

**Théorème 18** ( $C4_\alpha$ )  $\forall F = \langle W, R \rangle$ ,

*R est shift-réflexive*

$$\Leftrightarrow \forall \alpha \in ]0, 1], F \models \Box_\alpha \Box_1 \varphi \rightarrow \Box_\alpha \varphi$$

**Démonstration :** Soit  $\langle W, R \rangle$  un modèle tel que  $R$  est shift-réflexive. Soit  $s$  une valuation et  $\alpha \in ]0, 1]$ . Soit un monde  $w \in W$ . Si  $R_w = \emptyset$ , alors  $w$  valide trivialement toutes les formules de l'implication. Nous considérons dans la suite que  $|R_w| > 0$ .

Si  $w \not\models \Box_\alpha \Box_1 \varphi$ , alors  $w \models \Box_\alpha \Box_1 \varphi \rightarrow \Box_\alpha \varphi$ .

Si  $w \models \Box_\alpha \Box_1 \varphi$ , comme  $R$  est shift-réflexive, pour tout  $w' \in R_w$  tel que  $w' \models \Box_1 \varphi$ , on a  $w' \models \varphi$ . Par conséquent,  $R_w(\Box_1 \varphi) \subseteq R_w(\varphi)$ . On peut en déduire que :

$$\alpha \leq \frac{|R_w(\Box_1 \varphi)|}{|R_w|} \leq \frac{|R_w(\varphi)|}{|R_w|}$$

Aussi  $w \models \Box_\alpha \varphi$ .

La réciproque est montrée par contraposée : si la relation n'est pas shift-réflexive, alors l'axiome ( $C4_\alpha$ ) n'est pas validé. Soit  $W$  un ensemble fini de mondes et  $R$  une relation non shift-réflexive. Soit  $w \in W$  et  $w' \in R_w$  tels que  $w' \notin R_{w'}$ . Soit  $s$  une valuation telle que  $w' \not\models \varphi$  et  $\forall x \in W \setminus \{w'\}, x \models \varphi$ .

De fait, on a  $w \not\models \Box_1\varphi$  à cause de  $w'$ , cependant on a aussi  $w \models \Box_1\Box_1\varphi$ . Ainsi, en posant  $\alpha = 1$ , alors  $w \models \Box_\alpha\Box_1\varphi$  mais  $w \not\models \Box_\alpha\varphi$ , ce qui implique que  $w \not\models (C4_\alpha)$ .

■

### 3.8 Discussion

L'analyse de la question de la pondération des axiomes et de leur correspondance sémantique présentée dans ce chapitre permet d'établir l'ensemble des formules que le tableau 3.2, p. 45, récapitule en leur associant la contrainte correspondante sur la relation ainsi que leur type.

Cette étude a été effectuée en considérant chaque axiome individuellement, indépendamment des autres. Nous ébauchons ici l'examen de leur prise en compte simultanée, et les conséquences qu'elle provoque sur d'éventuelles versions pondérées des systèmes modaux classiques : fondés sur un sous-ensemble d'axiomes (cf. section 1.1.5, page 13) : on peut constater que dans les cas où la correspondance classique est modifiée, les systèmes modaux le sont également.

C'est par exemple le cas du système classique S4, qui regroupe les axiomes K, D, T et 4. Nous avons établi que l'axiome 4 n'est pondérable qu'à condition d'ajouter la contrainte d'euclidianité à la transitivité qui lui est classiquement associée. Considérer l'euclidianité impose d'avoir également l'axiome 5 ajouté au système. Même si ce dernier est impondérable, cette observation conduit à s'interroger sur la définition d'un éventuel système S4 pondéré, d'autant plus que la correspondance n'est pas établie pour l'axiome  $4_\alpha$  : la question des propriétés des modèles dans une axiomatique l'incluant est à traiter avec rigueur.

Selon des principes similaires, on peut noter que tout système modal pondéré inclut d'office l'axiome  $CD_\alpha$ , car celui-ci n'est conditionné par aucune contrainte sur la relation. Le passage entre  $\Diamond$  et  $\Box$  permis par l'axiome CD classique est sémantiquement très expressif, mais associé à une contrainte forte sur la relation, qui réduit le nombre de mondes accessibles à un seul.  $CD_\alpha$ , qui permet le passage entre  $\Diamond_\alpha$  et  $\Box_{1-\alpha}$ , a un pouvoir d'inférence moins riche mais il est intégré à tous les systèmes modaux, dont la richesse d'expression doit être considérée globalement.

Ainsi, plus généralement, ces axiomes ouvrent la voie à l'étude de systèmes modaux pondérés, notamment sous l'angle de leur expressivité.

### 3.9 Bilan

Ce chapitre a présenté l'étude approfondie que nous avons faite des axiomes de la logique modale afin d'en proposer des variantes pondérées, vis-à-vis de la sémantique proposée dans le chapitre précédent. En effet, il est indispensable de tenir compte de cette sémantique pour définir des valeurs pertinentes des poids, ouvrant à la voie à l'examen des répercussions de cette sémantique sur la théorie de la correspondance.

Nous avons organisé ces axiomes pondérés dans une typologie définissant quatre catégories, autour de trois axes de comparaison : la possibilité ou non d'effectuer la pondération, l'éventuelle modification de la propriété de la relation classiquement associée à l'axiome, que la pondération nécessite ou autorise et, réciproquement, l'éventuelle équivalence entre tout modèle respectant la ou les contraintes et l'axiome.

Nous avons effectué cette analyse dans le cadre des logiques modales pondérées générales, sans nous restreindre à une lecture spécifique de la modalité. Dans la partie suivante, nous nous concentrons sur le cas d'une modalité de croyances afin de définir et manipuler formellement des croyances graduelles.

Deuxième partie

Logique doxastique graduelle



# Introduction

Alors que la première partie a considéré la question des logiques modales pondérées de façon générale, nous nous intéressons, dans cette seconde partie, au cas particulier de la manipulation de croyances pondérées : nous considérons la question de la définition d'une logique des croyances dont nous voulons proposer une extension graduelle, en nous inspirant de la logique doxastique. En effet, celle-ci ne permet que de distinguer les deux cas « croire que » et « ne pas croire que ». La formalisation de croyances pondérées doit permettre de représenter « croire à un degré  $\alpha$  » où  $\alpha$  est une valeur numérique, permettant de moduler finement les croyances et d'accroître l'expressivité du modèle, à la fois en termes de représentation et de raisonnement.

Le modèle de représentation que nous proposons repose sur une sémantique ensembliste des croyances, qui se situe dans le courant représentationaliste de l'épistémologie, en considérant le cadre pondéré de la théorie des sous-ensembles flous : les degrés de croyance sont interprétés comme des degrés d'appartenance à l'ensemble des croyances de l'agent, défini sur l'ensemble des formules bien formées de la logique propositionnelle.

Nous examinons ensuite les règles de manipulation des croyances graduelles qui découlent de cette représentation, à la fois en donnant une lecture doxastique des propriétés issues de la théorie des sous-ensembles flous et en étudiant l'expression, dans le formalisme flou, d'une extension graduelle du raisonnement doxastique, en particulier les axiomes normaux ainsi que K, D, 4 et 5.

De plus, nous mettons en regard les formules ainsi obtenues avec les variantes pondérées d'axiomes établies dans le chapitre 3, en particulier  $(KD45)_\alpha$  : nous comparons ainsi une approche dédiée au cas de la manipulation de croyances graduelles et une approche consistant à particulariser les définitions associées aux logiques modales pondérées générales.

Cette partie est organisée comme suit : le chapitre 4 enrichit l'état de l'art de la partie précédente en détaillant les formalismes existants pour la représentation et la manipulation de croyances non binaires, à la fois de points de vue philosophiques et logiques. Le chapitre 5 formalise le modèle ensembliste flou que nous proposons pour représenter des croyances

graduelles et de discuter l'aspect doxastique des propriétés qui en découlent

Le chapitre 6 examine les conséquences de l'interprétation proposée pour la construction du raisonnement doxastique graduel : il propose un ensemble de règles de manipulation pour les croyances graduelles dans le formalisme flou, en considérant la conjonction, la disjonction, l'implication et l'introspection, c'est-à-dire en étudiant les axiomes classiques de la logique doxastique (les axiomes normaux et K, D, 4 et 5) dans le cadre flou et en les comparant à leurs variantes pondérés  $(KD45)_\alpha$ . Enfin, le chapitre 7 étudie le cas particulier de la négation appliquée à des croyances graduelles.

## Chapitre 4

# État de l'art : formalisation de la croyance

La seconde partie de ces travaux étant centrée sur la formalisation de croyances graduelles, nous proposons d'enrichir l'état de l'art présenté dans le chapitre 1 : nous nous intéressons à la notion même de croyance et à la façon dont les croyances sont conçues et représentées en épistémologie et en logique. Nous examinons également les différents mécanismes de graduation des croyances, en passant en revue les principaux formalismes qui permettent de représenter des degrés de croyance.

Ce chapitre est organisé en trois sections. La première présente un aperçu des travaux épistémologiques consacrés à la représentation des croyances binaires et non binaires. La section 4.2 décrit des modèles formels, logiques, pour la modélisation des croyances binaires et la section 4.3 présente ceux pour la modélisation des croyances non-binaires.

### 4.1 Aperçu épistémologique

La croyance est une attitude cognitive essentielle pour l'être humain, qui permet, par exemple, d'acquérir des informations, se construire une opinion ou améliorer ses interactions sociales (Fishbein & Ajzen, 1976). Cela en fait une notion primordiale en philosophie de l'esprit et en épistémologie. Cette section traite à la fois du cas binaire et du cas pondéré.

Il existe de multiples approches, présentées de façon détaillée dans l'encyclopédie de philosophie de Stanford (Schwitzgebel, 2015). Un rapide survol est présenté dans la section 4.1.1, le courant du représentationalisme (Dretske, 1988; Fodor, 1990) est détaillé dans la section 4.1.2 en raison de son rôle pour le modèle que nous proposons dans le chapitre suivant. Enfin, la section 4.1.3 considère le cas non binaire avec des degrés de croyance.

### 4.1.1 Des conceptions de croyance

Comme détaillé dans Schwitzgebel (2015), il existe de multiples modèles de croyances, qui diffèrent selon leur interprétation, leur représentation interne, leurs causes et leurs conséquences sur le comportement humain. Sans prétendre à l'exhaustivité, et de façon simplifiée, on peut par exemple citer :

- le dispositionalisme (Marcus, 1990) qui voit la croyance comme influençant la *disposition* de l'agent à des comportements ou actions : la croyance concerne uniquement son attitude quant aux observations au moment où il les rencontre. Elle n'est pas internalisée, contrairement au représentationalisme (cf. section 4.1.2).
- l'interprétationisme (Davidson, 2001) selon lequel la croyance permet d'*interpréter* et de prédire des systèmes non modélisables d'un point de vue physique ou biologique : il conserve l'aspect actif de la croyance entre agent et observation du dispositionalisme et l'enrichit par une dimension supplémentaire, car il rend possible l'observation de l'attitude de l'agent par un agent extérieur.
- le fonctionnalisme (Armstrong, 1968), compatible avec le représentationalisme comme avec le dispositionalisme, selon lequel l'observation de l'environnement conduit à construire les croyances.

Il existe également des postures plus radicales : l'éliminativisme (Churchland, 1981) réfute les croyances comme objet d'étude en niant leur existence, l'instrumentalisme (Dennett, 1991), moins extrême, considère que les croyances sont sans intérêt en elles-mêmes, mais qu'elles peuvent être utiles comme intermédiaires pour d'autres analyses.

### 4.1.2 Représentationalisme : une conception ensembliste des croyances

Nous proposons de nous concentrer sur le représentationalisme (Dretske, 1988; Fodor, 1990) dans la suite de ce chapitre : ce courant considère les croyances d'une manière adaptée à la formalisation logique.

Selon cette approche, la croyance est une attitude qui concerne la représentation d'un fait : une expérience, ou observation, induit la construction d'un ensemble de propriétés subjectives permettant de *représenter* l'observation dans l'esprit de l'individu. Celui-ci constitue ainsi sa propre description de son environnement, par un ensemble de représentations mentales des faits observés.

Ainsi, lorsqu'un agent rationnel observe un fait dans son environnement, il construit mentalement une représentation interne de ce fait et la conserve de son esprit. Ainsi, il peut se faire son propre avis sur le fait en question et alors exprimer sa croyance quant à l'observation initiale. De cette façon, un ensemble de représentations de faits constitue une description de ce que l'agent perçoit comme étant sa réalité.

Cette internalisation des faits induit une conception de la croyance sur laquelle nous basons nos propositions décrites dans le chapitre suivant : nous proposons que chaque représentation mentale d'un fait soit formellement considérée comme une formule de la logique propositionnelle et que les croyances soient conçus comme un ensemble de formules logiques constituant la représentation de l'environnement comme détaillé dans le chapitre 5, qui inclut de plus une gradualité. Nous proposons la construction d'un ensemble de croyances comme sous-ensemble flou de l'ensemble des formules bien formées.

### 4.1.3 Introduction de la gradualité

Il existe de nombreux travaux en épistémologie qui discutent de l'intérêt de raisonner avec des croyances modulées. Elles sont par exemple réfutées par Harman (2008) ou Moon (2017), mais soutenues par Sturgeon (2008) ou Staffel (2013). En particulier, Sturgeon (2008) développe une argumentation en considérant la croyance comme un concept au cœur d'un ensemble d'attitudes épistémiques propres à l'humain liés à la conceptualisation de la confiance et de la conviction, et montre qu'il est indispensable de la nuancer.

Staffel (2013) illustre l'intérêt de gradualiser le raisonnement sur les croyances par plusieurs exemples. En effet, l'utilisation de croyances binaires ne permet pas de refléter la réelle complexité des problématiques liées au raisonnement sur les croyances.

Un tel exemple a été évoqué dans l'introduction de ce mémoire, par les cryptozoologues, qui, respectivement, croient « dur comme fer », « très peu » ou « pas du tout » à l'existence des licornes. De même, on peut considérer un astrophysicien qui possède des croyances au sujets de plusieurs éléments liés à l'univers : il croit que le centre de la voie lactée contient un trou noir supermassif, appelé Sagittarius A\*. Il croit également que la vie est possible sur une autre planète de l'univers, peut-être dans une autre galaxie. Ces deux observations ne sont pas également admises par la communauté scientifique, de ce fait, l'astrophysicien eut ne pas vouloir y croire avec la même intensité. Cet exemple illustre le fait que la binarité des croyances est trop limitée pour être réaliste.

## 4.2 Modélisation de croyances binaires : formalisations logiques

Il existe de nombreux formalismes permettant de représenter formellement les croyances. Dans cette section, nous les passons brièvement en revue. Nombre d'entre eux dérivent de la notion d'*état de croyance* d'un agent, ou état épistémique, qui permet de modéliser sa perception d'une information. De cette façon, les modèles présentés peuvent être mis en perspective avec le formalisme modal en général et le formalisme doxastique ou épistémique

en particulier.

Cette section présente, dans un premier temps, la logique doxastique et décrit les propriétés de ce cas particulier des logiques modales en complétant les rappels effectués en première partie de ce mémoire. Ensuite, elle présente brièvement un modèle de représentation des croyances, aux propriétés différentes de la logique doxastique : la logique méta-épistémique, qui permet à un agent de se représenter les croyances d'un autre agent. Nous omettons la révision de croyance (Gärdenfors, 2003), qui offre un cadre pour représenter les croyances dont la particularité est d'être dynamique : ce modèle est consacré au raisonnement séquentiel ou temporel sur les croyances.

### 4.2.1 Logique doxastique

La logique doxastique (Hintikka, 1962), déjà évoquée dans la section 1.1.5, p. 13, est une logique modale qui interprète la modalité  $\Box$  comme une modalité de croyance :  $\Box\varphi$  se lit « l'agent croit  $\varphi$  ». Les axiomes doxastiques, qui sont communément admis constituent le système appelé KD45.

Ce système d'axiomes est classiquement composé des quatre schémas d'axiome, présentés ci-dessous, ainsi que leur interprétation doxastique associée.

(K)	$\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$	Rationalité
(D)	$\Box\varphi \rightarrow \neg\Box\neg\varphi$	Cohérence
(4)	$\Box\varphi \rightarrow \Box\Box\varphi$	Introspection positive
(5)	$\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$	Introspection négative

L'axiome (K) décrit le phénomène d'inférence interne d'un agent rationnel au sujet de ses propres croyances : s'il croit une règle, alors s'il croit la prémisse, il effectue l'inférence pour croire aussi la conclusion. L'axiome (D) impose la cohérence d'un agent vis-à-vis de ses croyances : il lui interdit de croire à la fois une chose et son contraire. Enfin, les axiomes (4) et (5) expriment le processus d'introspection : ils décrivent la conscience de l'agent pour ses propres croyances, c'est-à-dire s'il croit, ou non, une chose, alors il croit qu'il le croit, ou non.

Associée à des règles d'inférence classique, comme le *modus ponens*, l'axiomatisation est classiquement considérée pour représenter le raisonnement d'un agent rationnel vis-à-vis de ses propres croyances (Hintikka, 1962).

La sémantique des mondes possibles de la logique modale peut, de notre point de vue, être mise en lien avec le représentationalisme. En effet, chaque monde possible peut être interprété comme une façon d'appréhender l'environnement par l'agent. Un monde étant un point de vue contextuel, chacun d'eux représente les déclinaisons possibles de la description

mentale d'un environnement. Le formalisme modal est donc cohérent avec l'interprétation représentationnaliste des croyances.

#### 4.2.2 Logique méta-épistémique

La logique méta-épistémique (Banerjee & Dubois, 2009) est une formalisation du raisonnement sur les croyances, basée sur une sémantique particulière. En effet, ce modèle se base sur un ensemble de croyances possibles, à partir desquelles des formules sont construites en utilisant les opérateurs logiques classiques (conjonction, disjonction et négation). Ce système logique considère les croyances comme atomes de son langage, rendant la modalité doxastique implicite, contrairement aux systèmes ajoutant une modalité de croyance aux formules de la logique propositionnelle.

La particularité de cette proposition est qu'un agent peut raisonner sur les croyances d'un autre agent : le modèle autorise la définition d'état de méta-croyances.

Une axiomatisation du système est aussi proposée, fournissant un ensemble de règles de manipulation pour ces méta-croyances. Cependant, elle exclut les axiomes d'introspection afin de ne pas permettre à un agent de raisonner sur ses propres croyances. Ce modèle s'inscrit donc plutôt dans une extension de la logique KD, choisissant d'exclure les axiomes classiques d'introspection, en refusant de considérer des croyances « d'ordre 2 », c'est-à-dire des croyances portant sur des croyances.

### 4.3 Modélisation de croyances non binaires : formalisations numériques

Comme évoqué dans la section 4.1.3, on peut reprocher aux modèles de croyances binaires de manquer d'expressivité. Cette section présente brièvement les principaux modèles de croyances graduelles, qui reposent essentiellement sur le choix d'une structure arithmétique pour représenter les degrés associés aux croyances : les règles de manipulation des croyances graduelles sont alors dérivées de la structure considérée. L'entrée « *Belief* » de l'encyclopédie de philosophie de Stanford (Schwitzgebel, 2015) propose une présentation de ces modèles selon un point de vue philosophique.

**Probabilités subjectives** Les probabilités subjectives constituent une interprétation des probabilités non basée sur la fréquence observée : elles tiennent compte du jugement propre à l'agent pour évaluer l'incertitude d'un événement. Elles conservent néanmoins les règles de manipulation et de combinaison des probabilités classiques. La notion de probabilités subjectives permet de représenter ce que l'agent perçoit comme l'éventualité

qu'un événement se produise et non sa probabilité réelle. De cette façon, il peut moduler sa perception de l'environnement. D'après Savage (1972); De Finetti (1974); Smith et al. (1991), les probabilités subjectives expriment en particulier le niveau des croyances d'un individu. Un modèle pour des croyances graduelles basé sur des probabilités a été proposé par Bacchus et al. (1996).

Il faut noter que, les probabilités subjectives sont fortement liées à la véracité des faits, contrairement au point de vue modal qui considère les croyances comme non factuelles et donc détachées de la véracité des faits. Castelfranchi et Falcone (2000), qui utilisent la notion de croyance pour définir une forme de confiance comme construction d'un ensemble de croyances complexes, soulignent que la confiance se définit au-delà des probabilités subjectives : celles-ci ne constituent en effet pas un outil suffisamment puissant pour représenter et manipuler les éléments qui influencent à la fois la perception et le jugement d'un agent vis-à-vis de son environnement.

En particulier, la complexité des volets internes et externes de l'état épistémologique d'un agent rationnel fait appel à la fois à une vision subjective des faits mais également à toute l'influence extérieure à laquelle il est soumis. Ces niveaux de complexité ne peuvent être seulement représentés par des probabilités subjectives car elles n'apportent qu'un point de vue factuel sur l'environnement.

**Possibilités** La logique possibiliste de Dubois et al. (1994a) exploite la théorie des possibilités (Zadeh, 1978) pour permettre la manipulation et la combinaison par des opérateurs logiques d'assertions incertaines, associées à des degrés de certitude.

La logique possibiliste généralisée (Dubois et al., 2012) l'étend et propose une axiomatisation mise en perspective avec les axiomes de la logique méta-épistémique évoquée dans la section 4.2.2. Elle offre de la sorte un cadre formel de manipulation de degrés de croyance.

**Fonctions de croyance** La théorie des fonctions de croyance (Shafer, 1976) constitue un compromis entre une approche ensembliste et une approche probabiliste de la représentation de l'incertitude : ce modèle propose de définir la mesure de croyance d'un ensemble comme la somme des masses de toutes ses parties. La valeur de cette fonction de croyance correspond à un degré que l'on peut assimiler au niveau de la croyance.

Bien que cette modélisation ne soit pas classiquement utilisée pour le raisonnement doxastique, elle est cependant proposée pour constituer une sémantique pour une logique modale des croyances graduelle : Godo et al. (2003) proposent d'utiliser les valeurs attribuées à la véracité d'une formule doxastique. Ainsi la formule  $B\varphi$  est plus ou moins vraie, et sa valeur associée s'inscrit dans le formalisme des fonctions de croyance. En terme

---

d'étude axiomatique, cette proposition adapte le système S5 et sort du cadre doxastique.

## 4.4 Bilan

Ce chapitre a présenté différentes formalisations de la croyance dans différents domaines, d'un point de vue philosophique et d'un point de vue logique, pour des croyances binaires ou graduelles.

Les modèles présentés pour la représentation de croyances graduelles, ont en commun d'être basés sur l'incertitude quant à la véracité des faits. Comme détaillé par la suite, la notion d'incertitude ne correspond pas à l'approche que nous souhaitons. De plus, nous nous intéressons à des éléments non factuels, car c'est ce qui fait l'essence, dans la logique modale, de la notion de croyance.

Considérant l'approche proposée par la suite pour les croyances graduelles, nous proposons d'adapter la conception de la croyance donnée par le représentationalisme avec un modèle adapté aux croyances graduelles et basé sur un ensemble de représentations internes de faits stockées dans l'esprit de l'agent.



## Chapitre 5

# Représentation de croyances graduelles : proposition d'un formalisme flou

Dans le but de formaliser les croyances graduelles, il faut à la fois pouvoir les représenter et les manipuler. Ce chapitre est consacré à la première de ces tâches et présente le modèle que nous proposons pour représenter les croyances graduelles. Formellement, nous notons une telle croyance graduelle  $B(\varphi, \alpha)$ , afin de représenter simultanément les trois éléments qui la constituent :

1. la formule considérée  $\varphi$ , au sujet de laquelle la croyance est exprimée
2. le type de modalité, une croyance notée  $B$ , exprimée au sujet de  $\varphi$
3. le degré de croyance  $\alpha$ , qui est un coefficient numérique

Avec cette notation, l'exemple « l'agent croit fermement que les licornes existent » s'écrit  $B(\varphi, \alpha)$ , où  $\varphi$  est la formule « les licornes existent » et  $\alpha$  un degré représentant « fermement ». La mise en correspondance des termes linguistiques avec des valeurs numériques sort du cadre des travaux présentés, on choisit ici, à titre d'exemple,  $\alpha = 0.8$ .

Le modèle que nous proposons pour interpréter une telle formule se situe, d'une part, dans un cadre ensembliste, en accord avec les principes du représentationnalisme présentés au chapitre précédent, et d'autre part dans le cadre de la théorie des sous-ensembles flous (Zadeh, 1965).

La première section discute des principes qui motivent l'utilisation d'un formalisme flou par rapport aux formalismes rappelés dans le chapitre précédent. La section 5.2 présente le modèle que nous proposons et la section 5.3 donne une lecture doxastique des propriétés découlant de l'interprétation floue des degrés de croyance.

Les travaux présentés dans ce chapitre ont fait l'objet des publications (Legastelois et al., 2016a; Legastelois et al., 2017a).

## 5.1 Motivations et principes

Nous présentons dans cette section les trois principes visant à établir des arguments pour l'interprétation des poids que nous proposons : la distinction entre croyance et incertitude, la pertinence d'une extension de l'approche représentationaliste des croyances et la définition de trois modalités induites.

### 5.1.1 Distinction entre croyance et incertitude

Comme évoqué dans la section 4.3 de l'état de l'art du chapitre précédent, il existe de nombreux travaux sur les croyances graduelles, qui travaillent principalement sur le degré de croyance, en considérant différentes structures arithmétiques pour le représenter, parmi lesquelles nous avons cité les probabilités, les probabilités subjectives, les possibilités ou les fonctions de croyance évidentielles. Cependant, ces modèles sont principalement basés sur la notion de certitude que l'agent associe à la véracité du fait considéré, voir en particulier (Bacchus et al., 1996), et ne se focalisent pas sur le niveau de croyance, potentiellement indépendant de cette véracité.

En effet, on peut s'attendre à ce que l'incertitude quant à la véracité d'une formule soit fortement corrélée au niveau de croyance que l'on peut lui associer : il peut sembler naturel de considérer que plus un degré d'incertitude est élevé, plus le niveau de croyance sera faible. Par exemple, un agent ayant des doutes sur le temps qu'il fait à Singapour ne pourrait pas avoir une profonde conviction sur le fait qu'il pleuve là-bas. Réciproquement, on peut considérer que si un fait est absolument certain, alors sa croyance associée est hautement valuée.

Toutefois, cette interprétation des croyances basée sur la corrélation entre la certitude et le niveau de croyance, semble restrictive. D'abord, elle ne tient pas compte de la subjectivité des croyances et de leur composante non factuelle : il se peut qu'un agent n'ait pas conscience qu'un fait est certain et n'y croie pas. Cette approche échoue également à mettre en avant d'autres aspects essentiels de la notion de croyance, nous discutons ici les cas de superstition et de dissonance cognitive.

D'une part, les superstitions conduisent les gens à croire fortement des événements particulièrement incertains. Par exemple, si un agent très superstitieux trouve un trèfle à quatre feuilles, il croira très fortement en sa chance de gagner à la loterie, ce qui est un fait intrinsèquement incertain.

D'autre part, la différence entre croyance et certitude peut provoquer une dissonance cognitive (Festinger, 1962) : cette section rend compte du fait que l'on peut être amené à déformer la réalité afin de la rendre conforme à ses croyances. En conséquence, si une information contredit ses convictions, un agent peut avoir tendance à la rejeter et à chercher plutôt d'autres sources d'information qui abondent dans son sens.

Un exemple classique de dissonance cognitive est l'effet placebo de certains médicaments : un médecin peut juger qu'il est très improbable qu'un traitement soit efficace mais le prescrire malgré tout. Dans ce cas, la démarche du médecin est rationnelle car il se base sur des études qui montrent un potentiel effet du médicament. Il veut croire en son efficacité malgré l'incertitude qui plane sur celle-ci, et dont il est conscient.

Pour ces raisons, nous considérons que les approches basées sur l'incertitude, présentées dans le chapitre précédent, ne peuvent représenter toutes les subtilités de la notion de croyance. Nous choisissons donc de faire la distinction entre des croyances basées sur l'incertitude et des croyances graduelles dont le niveau est imprécis.

Un type particulier est également observé lorsqu'un agent raisonne sur les croyances d'un autre (Banerjee & Dubois, 2009; Dubois et al., 2014) : cette incertitude modélisée dans un cadre possibiliste (Dubois & Prade, 2011; Dubois et al., 2011) offre une variante graduelle de la logique méta-épistémique (Banerjee & Dubois, 2009), là où les travaux présentés ici se concentrent sur les croyances d'un seul agent.

### 5.1.2 Extension dans le formalisme flou du représentationnalisme

La conception représentationnaliste des croyances, qui définit les croyances comme un ensemble de représentations de faits stocké dans l'esprit de l'agent (Fodor, 1981), peut être étendue au cas des croyances graduelles.

Pour cela, nous proposons de substituer l'ensemble classique, qui peut, dans un cadre formel, donner une représentation mathématique de l'ensemble des croyances de l'agent, par un sous-ensemble flou. Celui-ci permet d'associer à chaque description de fait un degré d'appartenance à l'ensemble des croyances.

En effet, l'ensemble des croyances stockées dans l'esprit d'un individu peut être étendu au cas d'une appartenance non binaire : chaque élément appartient plus ou moins à l'ensemble de croyance, le degré d'appartenance indiquant son degré de croyance associé.

D'un point de vue formel, le modèle proposé dans la section suivante repose sur un tel ensemble flou, considéré comme l'ensemble de croyances et pour lequel chaque valeur d'appartenance est interprétée comme un degré de croyance associé. Ainsi, formellement, les éléments de l'ensemble de croyances sont des formules logiques qui modélisent l'ensemble des faits observés par l'agent : les objets manipulés ici sont des représentations formelles

de ces observations. Par exemple, si l'agent croit fortement qu'une autre planète abrite des êtres vivants, alors la formule modélisant ce fait appartiendra à un degré correspondant à l'adverbe « fortement », par exemple 0.8, à l'ensemble de ses croyances.

### 5.1.3 Un système multimodal induit : croyance, semi-croyance et incroyance

La modélisation proposée pour les croyances graduelles repose sur l'utilisation d'un ensemble mathématique comme formalisation de la mémoire : les représentations de faits accumulées dans l'esprit de l'agent sont des éléments de cet ensemble, dit « de croyances ». Différents niveaux de croyance sont ainsi représentés, pour chaque degré d'appartenance dans l'intervalle réel  $[0, 1]$ .

Syntaxiquement, nous proposons de définir trois modalités différentes, correspondant aux cas de croyance, semi-croyance et incroyance. Ces trois notions représentent trois attitudes possibles d'un agent quant à son état de croyance.

Tout d'abord, il peut attribuer une croyance totale, et donc classique au sens modal du terme, ce qui conduit à une première modalité, notée  $B_1$ . Ensuite, on peut définir une modalité de semi-croyance, modulée par une imprécision quant à l'intensité de la croyance, celle-ci est notée  $B_\alpha$ , où  $\alpha$  est une valeur numérique encodant la modulation de la croyance. Enfin, l'incroyance, notée  $B_0$ , désigne le fait que l'agent n'a pas d'avis sur l'assertion considérée.

## 5.2 Proposition d'un modèle ensembliste flou

Nous proposons de formaliser l'interprétation floue de l'ensemble des croyances graduelles d'un agent introduite dans la section 5.1.3, dans laquelle une formule appartient plus ou moins à cet ensemble de croyances.

Cette section donne la définition que nous proposons puis discute de l'interprétation des degrés d'appartenance, dans les sections 5.2.1 et 5.2.2 respectivement.

### 5.2.1 Définitions

Formellement, nous interprétons l'ensemble  $\mathcal{B}$  des croyances graduelles d'un agent comme un sous-ensemble flou de l'univers des formules syntaxiquement correctes  $\mathcal{F}$  : chaque formule appartient plus ou moins à l'ensemble de croyances  $\mathcal{B}$  et le degré d'appartenance est interprété comme un degré de croyance.

Ainsi, en notant  $\mu_{\mathcal{B}} : \mathcal{F} \rightarrow [0, 1]$  la fonction d'appartenance à  $\mathcal{B}$ , on considère

$$B(\varphi, \alpha) \quad \text{si et seulement si} \quad \mu_{\mathcal{B}}(\varphi) = \alpha \quad (5.1)$$

L'état épistémique de l'agent est donc considéré dans une variante pondérée, notamment en basant le raisonnement de l'agent sur cet ensemble de croyances comme sous-ensemble flou.

### 5.2.2 Degrés de croyances et degrés d'appartenance

Il est important de noter que les poids ne sont pas considérés comme des degrés d'incertitude. En effet, l'utilisation de ce degré d'appartenance permet d'affirmer que le niveau de croyance indique un niveau d'imprécision quant à la quantification de cette croyance. Il faut également souligner que le degré d'appartenance n'est pas interprété comme le degré de vérité de la formule  $\varphi$ , mais comme la force de la croyance.

L'interprétation des degrés de croyance proposée est basée sur une lecture floue et la notion sous-jacente d'imprécision en terme de degrés de vérité.

Chaque formule, dans cette interprétation, appartient plus ou moins à l'ensemble de croyances  $\mathcal{B}$ . De cette façon, si l'agent croit totalement un fait, c'est-à-dire s'il le croit sans aucune restriction, alors la formule correspondante est associée à un degré de valeur 1. Cette formule appartient totalement à l'ensemble de croyance  $\mathcal{B}$ . Une formule associée à un degré de croyance plus faible appartient également plus faiblement à l'ensemble de croyances.

Cette interprétation floue invite à examiner, d'un point de vue doxastique, deux propriétés majeures de la théorie des sous-ensembles flous : l'inclusion des  $\alpha$ -coupes et les cas particuliers des valeurs aux bornes. Les conséquences de cette interprétation sont détaillées dans la prochaine section.

## 5.3 Lecture doxastique des propriétés floues

Cette section propose de s'intéresser à la signification des principales propriétés des sous-ensembles flous, en terme de degrés de croyance. Nous proposons donc une lecture doxastique de la notion d' $\alpha$ -coupe et des valeurs extrêmes 0 et 1.

### 5.3.1 $\alpha$ -coupes et gradualité

Pour tout sous-ensemble flou  $A$  d'un univers  $X$ , et pour tout  $\alpha \in [0, 1]$ , l' $\alpha$ -coupe de  $A$  est définie comme le sous-ensemble classique  $A_\alpha \subseteq X$  contenant tous les éléments dont le degré d'appartenance est supérieur ou égal à  $\alpha$ . Par définition, on a alors  $\forall \alpha, \beta \in [0, 1]$ , si  $\alpha \geq \beta$ ,  $A_\alpha \subseteq A_\beta$ .

En terme de degrés de croyance, l' $\alpha$ -coupe correspond à l'ensemble classique de formules que l'agent croit au moins au degré  $\alpha$  (expression abrégée en «  $\alpha$ -croit »). D'un point de

vue doxastique, la propriété d'inclusion indique que si  $\alpha \geq \beta$ , et si l'agent  $\alpha$ -croit une formule, alors il  $\beta$ -croit également cette même formule : cela correspond à la propriété de gradualité décroissante qui semble pertinente pour modéliser des croyances graduelles.

Formellement, cette propriété s'écrit :

$$\forall \varphi \in \mathcal{F}, \forall \alpha, \beta \in [0, 1], \quad B(\varphi, \alpha) \rightarrow B(\varphi, \beta) \quad (5.2)$$

Il est important de noter que cette propriété correspond à la propriété de gradualité décroissante donnée dans la propriété 2, page 34, dans le chapitre 2.

La notion d' $\alpha$ -coupes permet également de définir un raisonnement prudent de l'agent : il peut choisir d'effectuer un raisonnement binaire classique restreint aux croyances associées à un degré minimal qu'il fixe en omettant les distinctions qu'il peut faire entre elles, en utilisant l' $\alpha$ -coupe  $\mathcal{B}_\alpha$  de son ensemble de croyances.

Si l'agent choisit de raisonner de manière classique mais accepte de ne pas croire totalement un fait, alors il peut inclure toute croyance comprises entre le niveau maximal du degré 1 et le degré minimum acceptable pour lui.

### 5.3.2 Valeurs aux bornes et système multimodal induit

L'interprétation floue proposée amène également une discussion sur les valeurs limites des degrés de croyance possibles d'un agent. Les bornes du domaine de définition des degrés d'appartenance, 0 et 1, jouent un rôle spécifique en théorie des sous-ensembles flous. En effet, elles permettent d'assurer la correspondance avec les ensembles classiques : comme annoncé ci-dessus, une formule ayant un degré d'appartenance valant 1 appartient totalement à l'ensemble de croyance. Elle peut donc être interprétée comme une croyance dans le cadre de la logique doxastique classique ou, selon les trois modalités introduites dans la section 5.1.3, comme une croyance.

Les formules ayant un degré d'appartenance nul ont une interprétation plus délicate, qui peut se faire selon deux points de vue : d'abord la valeur 0 peut être lue comme une mesure de croyance négative, formalisant le fait que la formule associée est soumise à une modalité d'incroyance. Cette première approche permet d'enrichir la définition propre de l'ensemble de croyance car elle requiert l'établissement de règles pour raisonner à la fois sur des croyances positives et négatives.

La valeur nulle peut, toutefois, également être interprétée comme un manque d'information au sujet du degré de croyance réel : la 0-croyance représente alors la valeur par défaut du degré de croyance. Ainsi, cette valeur correspond au cas trivial où aucune information n'est apportée quant au niveau réel de la croyance considérée. Cette seconde interprétation est particulièrement pertinente vis-à-vis de la propriété d'inclusion des  $\alpha$ -coupes, qui

établit que la 0-coupe est égale à l'univers  $\mathcal{F}$ , dans lequel toutes les autres  $\alpha$ -coupes sont incluses, mais à partir duquel aucune nouvelle croyance ne peut être inférée.

Nous soulignons également la cohérence entre une croyance nulle de ce type et une incroyance comme définie précédemment : la 0-coupe ne permet aucune déduction supplémentaire car elle est égale à l'univers. Tout comme l'incroyance, il n'en découle donc pas d'inférence possible.

## 5.4 Bilan

Ce chapitre a discuté d'une formalisation dans le cadre flou de la conception représentationnelle des croyances : elle interprète les degrés de croyance associés à des assertions comme des degrés d'appartenance, mesurant l'appartenance des formules logiques qui les représentent à l'ensemble des croyances de l'agent.

Sur la base de ce choix de représentation des croyances, nous nous intéressons, dans les chapitres suivants, à la question de leurs manipulations et combinaisons par les opérateurs logiques classiquement appliqués aux croyances : conjonction, disjonction, implication et introspection dans le chapitre 6, le cas de la négation, plus complexe, étant examiné dans le chapitre 7.



## Chapitre 6

# Manipulation des croyances graduelles dans le formalisme flou

Ce chapitre étudie la question de la manipulation et de la combinaison de croyances graduelles représentées dans le formalisme flou proposé dans le chapitre précédent : il s’agit de définir des outils de raisonnement permettant d’inférer de nouvelles croyances à partir de croyances atomiques. Ce chapitre examine les opérateurs de conjonction, disjonction et d’implication, et considère le principe essentiel en logique doxastique de l’introspection qui repose sur des successions de modalités.

Formellement, en utilisant les notations introduites dans le chapitre précédent, la manipulation de croyances graduelles consiste à identifier les relations entre  $B(\varphi, \alpha)$ ,  $B(\psi, \beta)$  et  $B(\varphi \otimes \psi, \gamma)$ , pour  $\otimes \in \{\wedge, \vee, \rightarrow\}$ , ainsi que  $B(B(\varphi, \alpha), \gamma)$  pour  $\alpha, \beta, \gamma \in [0, 1]$ .

Pour chacune d’entre elles, nous examinons les définitions induites par le choix de représentation décrit dans le chapitre précédent et nous les comparons à la forme de l’axiome pondéré correspondant tel qu’établi dans la première partie du manuscrit, soit  $NA_\alpha$ ,  $K_\alpha$  ainsi que  $4_\alpha$  et  $5_\alpha$  : nous proposons ainsi de mettre en parallèle les définitions de l’approche sémantique mise en œuvre pour les logiques modales pondérées générales et l’approche floue dédiée à la logique doxastique graduelle.

Le chapitre est organisé de la façon suivante : les sections 6.1 à 6.3 sont respectivement consacrées à la conjonction et la disjonction, traitées simultanément, l’implication et l’introspection. Dans chacun des cas, nous discutons successivement la forme générale de la règle de manipulation, notre proposition, puis la comparaison avec l’axiome pondéré correspondant.

Les travaux présentés ont fait l’objet des publications (Legastelois et al., 2015a; Legastelois et al., 2017a).

## 6.1 Conjonction et disjonction

Les règles de manipulation de formules par conjonction et disjonction prennent des formes variées selon les systèmes logiques considérés, comme rappelé dans l'annexe B, p. 145. Elles soulèvent une double problématique, de factorisation et de distribution : la première consiste à combiner deux formules en une formule plus complexe, là où la distribution, au contraire, décompose une formule complexe en ses sous-formules.

Cette section étudie cette question, pour les formules  $B(\varphi, \alpha)$ ,  $B(\psi, \beta)$  et  $B(\varphi \otimes \psi, \gamma)$  avec  $\otimes \in \{\wedge, \vee\}$ , afin d'établir des relations entre les poids  $\alpha$ ,  $\beta$  et  $\gamma$ . Pour cela, elle s'inspire et exploite en particulier les formalismes des logiques pondérées et spécifiquement celui de la logique floue, en raison de son lien avec la théorie des sous-ensembles flous dans laquelle s'inscrit l'interprétation proposée de l'ensemble de croyances.

L'objectif de cette section est de déterminer des opérateurs flous utilisables dans le contexte de la combinaison de croyances conjointes ou disjointes.

Cette section présente le cas de la manipulation autour de la conjonction et de la disjonction. Après avoir présenté la forme générale des règles de manipulation proposées, cette section discute en détail les différentes propriétés des opérateurs d'agrégation, donnant des arguments logiques et sémantiques. Elle propose enfin une mise en correspondance avec les axiomes modaux pondérés établis dans les travaux présentés en première partie de ce manuscrit.

### 6.1.1 Forme générale

L'objectif des règles de manipulation pour la conjonction et la disjonction est de relier les formules suivantes :  $B(\varphi, \alpha) \otimes B(\psi, \beta)$  et  $B(\varphi \otimes \psi, \gamma)$ , avec  $\otimes$  défini dans l'ensemble  $\{\wedge, \vee\}$ . Alors que la factorisation de ces croyances graduelles, peut d'appuyer intuitivement sur une combinaison des valeurs de  $\alpha$  et  $\beta$  pour définir  $\gamma$ , leur distribution, passant réciproquement de  $\gamma$  à  $\alpha$  et  $\beta$  est moins évidente, surtout dans le cas de la disjonction : si l'on croit  $\varphi \vee \psi$ , il n'est pas facile de déduire des informations sur  $\varphi$  et sur  $\psi$ . Ainsi, pour les logiques modales, la double implication est établie dans le cas de la conjonction et constitue une propriété centrale ; alors qu'elle ne l'est pas pour le cas pour la disjonction.

La question des valeurs de ces degrés s'inscrit dans cette problématique : le lien à établir entre les poids se base sur le fait que  $\gamma$  corresponde à une combinaison entre  $\alpha$  et  $\beta$ .

Toutefois, pour la logique floue et la logique possibiliste, l'équivalence existe pour les deux cas, conjonctif et disjonctif, en raison de l'utilisation de degrés de vérité ou de possibilité. Elle préserve une certaine flexibilité en ce que les valeurs ne sont pas complètement spécifiques mais soumises à des contraintes sur la véracité des formules  $\varphi$  et  $\psi$  puisque ces formalismes permettent de manipuler des informations incomplètes. Cette observation

peut être également utilisée pour des croyances graduelles, prenant identiquement avantage de l'utilisation d'un degré.

Aussi, nous proposons de considérer les deux équivalences : en se basant avant tout sur les logiques pondérées, les règles de manipulation pour des croyances graduelles peuvent être écrites de la manière suivante.

$$B(\varphi, \alpha) \wedge B(\psi, \beta) \Leftrightarrow B(\varphi \wedge \psi, \gamma) \quad \text{pour } \gamma = \text{agg}_{\wedge}(\alpha, \beta)$$

$$B(\varphi, \alpha) \vee B(\psi, \beta) \Leftrightarrow B(\varphi \vee \psi, \gamma) \quad \text{pour } \gamma = \text{agg}_{\vee}(\alpha, \beta)$$

avec  $\text{agg}_{\wedge}$  et  $\text{agg}_{\vee}$  deux opérateurs d'agrégation qui sont à spécifier afin de différencier les deux opérateurs logiques  $\wedge$  et  $\vee$ .

En conséquence de l'équivalence imposée, la factorisation et la distribution sont définies par des opérateurs d'agrégation, choisis pour établir les relations entre les trois coefficients  $\alpha$ ,  $\beta$  et  $\gamma$ . La règle de manipulation est alors définie par le choix de l'opérateur d'agrégation.

Il existe un grand nombre de possibilités quant au choix de l'opérateur d'agrégation le plus pertinent : ceux-ci possèdent chacun des propriétés qui leur sont propres et ont été organisés selon leur nature (Detyniecki, 2000). La sélection d'un opérateur approprié nécessite plusieurs types d'arguments, discutés par la suite, chacun imposant des propriétés attendues : les arguments logiques sont basés sur les définitions précédentes et la compatibilité avec la logique classique, les arguments sémantiques sont basés sur la signification que l'on souhaite donner aux degrés de croyance, enfin les arguments formels, sont basés sur le paradigme formel choisi pour l'interprétation des degrés de croyance.

Les différentes propriétés sont donc successivement examinées par la suite en considérant la commutativité, l'associativité, la monotonie et l'élément neutre : les cas particuliers où l'un de ces poids vaut 0 ou 1 sont également considérés, permettant de mettre en relation les trois types de modalités, décrivant une croyance totale, partielle ou une incroyance, tels qu'introduits dans la section 5.1.3, p. 80. Pour chacune de ces propriétés, leur pertinence est discutée selon les trois aspects possibles : la justification logique, sémantique et formelle. L'opérateur d'agrégation est noté  $\text{agg}$  pour généraliser  $\text{agg}_{\wedge}$  et  $\text{agg}_{\vee}$ .

### 6.1.2 Commutativité

La commutativité est une propriété classique qui impose :

$$\text{agg}(\alpha, \beta) = \text{agg}(\beta, \alpha)$$

Nous discutons de la pertinence d'une telle propriété pour l'agrégation de degrés de croyance.

**Justification logique** Nous proposons un argument identique pour chacun des deux opérateurs. Nous ne donnons donc que la preuve pour le cas de la conjonction.

En appliquant la règle de factorisation donnée ci-dessus,  $F_1 = B(\varphi, \alpha) \wedge B(\psi, \beta)$  et  $F_2 = B(\psi, \beta) \wedge B(\varphi, \alpha)$  s'écrivent respectivement  $B(\varphi \wedge \psi, \text{agg}_\wedge(\alpha, \beta))$  et  $B(\psi \wedge \varphi, \text{agg}_\wedge(\beta, \alpha))$ .

Comme  $\varphi \wedge \psi \Leftrightarrow \psi \wedge \varphi$ , par commutativité du  $\wedge$  en logique classique et que, pour la même raison  $F_1 \Leftrightarrow F_2$ , les degrés associés obtenus par les deux applications successives des règles, en commençant respectivement par  $F_1$  et par  $F_2$ , sont les mêmes. L'opérateur  $\text{agg}_\wedge$  doit donc être commutatif, par cohérence avec la commutativité de  $\wedge$ .

Le même argument s'applique pour la disjonction, basé sur l'équivalence et la commutativité de  $\vee$ .

**Justification sémantique** En plus de l'argument lié à la manipulation logique donné ci-dessus, nous proposons de discuter la signification d'une telle propriété en terme de degrés de croyance. En effet, la propriété de commutativité peut être interprétée comme « associer une  $\alpha$ -croyance à une  $\beta$ -croyance est pareil que d'associer une  $\beta$ -croyance à une  $\alpha$ -croyance ». En d'autres termes, l'ordre entre les croyances à combiner n'a pas d'importance.

Nous considérons cette propriété comme raisonnable pour une raison majeure : les opérateurs  $\wedge$  et  $\vee$  sont utilisés ici comme une conjonction et une disjonction logique et n'ont pas un sens séquentiel. En effet, en langage naturel, «  $A$  et  $B$  » peut être compris comme un ordre et interprété comme « d'abord  $A$  et ensuite  $B$  ». Néanmoins, les croyances ne sont pas considérées ici dans un contexte temporel, l'interprétation dynamique étant un sujet à part entière qui dépasse celui des travaux présentés. De ce fait, la commutativité est une propriété attendue et souhaitable.

### 6.1.3 Associativité

L'associativité d'un opérateur d'agrégation impose :

$$\text{agg}(\text{agg}(\alpha, \beta), \gamma) = \text{agg}(\alpha, \text{agg}(\beta, \gamma))$$

À nouveau, nous étudions cette propriété afin d'en montrer la pertinence lorsqu'elle s'applique à l'agrégation de degrés de croyance.

**Justification logique** Comme pour l'associativité, la preuve est identique pour chacun des deux opérateurs  $\wedge$  et  $\vee$ . Nous ne donnons donc que celle pour le cas conjonctif.

Tout d'abord, en appliquant deux fois la règle de factorisation et en utilisant la propriété

d'équivalence, on obtient :

$$\begin{aligned} (B(\varphi_1, \alpha) \wedge B(\varphi_2, \beta)) \wedge B(\varphi_3, \gamma) &\Leftrightarrow B(\varphi_1 \wedge \varphi_2, \text{agg}_\wedge(\alpha, \beta)) \wedge B(\varphi_3, \gamma) \\ &\Leftrightarrow B((\varphi_1 \wedge \varphi_2) \wedge \varphi_3, \text{agg}_\wedge(\text{agg}_\wedge(\alpha, \beta), \gamma)) \\ &\Leftrightarrow B(\varphi_1 \wedge \varphi_2 \wedge \varphi_3, \text{agg}_\wedge(\text{agg}_\wedge(\alpha, \beta), \gamma)) \end{aligned}$$

La dernière ligne ici est obtenue par l'associativité de  $\wedge$ . De manière similaire, on obtient :

$$B(\varphi_1, \alpha) \wedge (B(\varphi_2, \beta) \wedge B(\varphi_3, \gamma)) \Leftrightarrow B(\varphi_1 \wedge \varphi_2 \wedge \varphi_3, \text{agg}_\wedge(\alpha, \text{agg}_\wedge(\beta, \gamma)))$$

Enfin, par associativité de  $\wedge$ , on sait que :

$$(B(\varphi_1, \alpha) \wedge B(\varphi_2, \beta)) \wedge B(\varphi_3, \gamma) \Leftrightarrow B(\varphi_1, \alpha) \wedge (B(\varphi_2, \beta) \wedge B(\varphi_3, \gamma))$$

Par conséquence les degrés associés à la formule  $\varphi_1 \wedge \varphi_2 \wedge \varphi_3$  sont identiques. Aussi, l'opérateur  $\text{agg}_\wedge$  doit être associatif.

**Justification sémantique** Considérant l'interprétation doxastique des formules, l'associativité peut être lue comme « agréger une  $\alpha$ -croyance avec une croyance de niveau  $\text{agg}(\beta, \gamma)$  revient à agréger une croyance de niveau  $\text{agg}(\alpha, \beta)$  avec une  $\gamma$ -croyance ». En d'autres termes, l'ordre d'inspection des croyances, dans un contexte non temporel, n'a pas d'influence sur le traitement ou la production des résultats. Ainsi, pour la sémantique des degrés de croyance, nous considérons que l'associativité est à la fois pertinente et souhaitable.

#### 6.1.4 Propriétés de l'agrégation imposée par l'interprétation floue

L'interprétation floue des degrés de croyance proposée et discutée dans le chapitre précédent impose des propriétés particulières aux opérateurs d'agrégation présentés. En effet, le modèle de manipulation des degrés d'appartenance en théorie des sous-ensembles flous peut s'appliquer au cas des croyances graduelles, si l'on considère celles-ci dans un cadre logique formel. De même, la logique floue permet de manipuler les degrés de vérité selon l'application de propriétés ensemblistes.

Nous discutons par la suite de la pertinence des propriétés floues dans le cadre doxastique.

Comme mentionné en section B.1, les opérateurs flous de conjonction et disjonction, classiquement la t-norme et la t-conorme sont commutatifs, associatifs, croissants et ont respectivement pour élément neutre 1 et 0. Les sections précédentes ont montré, d'un point de vue à la fois logique et sémantique, la pertinence de l'associativité et de la commutativité. Les sections suivantes discutent des propriétés restantes : la monotonie et la définition de l'élément neutre, d'un point de vue purement sémantique induit par l'interprétation floue des ensembles de croyance proposée.

#### 6.1.4.1 Monotonie

La propriété de monotonie pour un opérateur *agg* indique que :

$$\forall \alpha, \alpha', \beta \in [0, 1], \quad \text{si } \alpha \leq \alpha' \quad \text{agg}(\alpha, \beta) \leq \text{agg}(\alpha', \beta)$$

Nous proposons de justifier cette propriété par un argument d'ordre sémantique.

Dans sur l'interprétation floue des ensembles de croyance, cette propriété implique que le degré d'appartenance  $\text{agg}(\alpha, \beta)$  associé à la croyance combinée dérivé de  $B(\varphi, \alpha)$  et  $B(\psi, \beta)$ , est plus faible que celui associé à la formule dérivée de  $B(\varphi, \alpha')$  et  $B(\psi, \beta)$ , dont la valeur est égale à  $\text{agg}(\alpha', \beta)$ . Cela signifie que le degré de croyance, obtenu pour la factorisation de deux croyances graduelles, augmente avec le niveau de croyance des formules atomiques.

Cette propriété est attendue et souhaitable : si deux cryptozoologues ont le même degré de croyance dans l'existence des licornes mais que le premier croit plus à l'existence des dragons que le second, alors il croit aussi plus à l'éventualité de l'existence conjointe des licornes et dragons que le second.

#### 6.1.4.2 Élément neutre

La propriété de l'élément neutre qui établit que :

$$\forall \alpha \in [0, 1], \quad \text{agg}_{\wedge}(\alpha, 1) = \alpha \quad \text{et} \quad \text{agg}_{\vee}(\alpha, 0) = \alpha$$

Dans un contexte doxastique, ces éléments signifient que la conjonction entre une croyance totale avec un croyance partielle induit une croyance dont le niveau est le plus bas. Dans le cas d'une disjonction entre une croyance nulle et une croyance partielle, alors le niveau de croyance associé au résultat équivaut à la valeur maximale entre les deux valeurs possibles. Nous soulignons que l'interprétation proposée fait référence à la discussion sur les valeurs aux bornes des degrés de croyance dans l'interprétation floue (cf. section 5.3.2, page 82).

En particulier, le comportement du cas disjonctif est particulièrement cohérent avec l'interprétation du degré nul : en tant que marqueur d'une absence totale d'information sur l'appartenance de la formule à l'ensemble de croyance, toute croyance ayant un degré non nul qui est associée à une croyance à valeur nulle par une disjonction prévaut en terme de niveau de croyance.

De par cette propriété, le lien entre les modalités particulières des valeurs aux bornes et les modalités représentant les croyances partielles peut être établi dans le respect de l'interprétation floue et en toute cohérence avec la notion de degrés de croyance.

### 6.1.5 Comparaison avec les axiomes normaux pondérés

En raison de l'analyse effectuée dans les sections précédentes, les règles de manipulation de conjonction et de disjonction données précédemment peuvent être spécifiées avec  $\top$  une t-norme et  $\perp$  une t-conorme :

$$B(\varphi, \alpha) \wedge B(\psi, \beta) \Leftrightarrow B(\varphi \wedge \psi, \gamma) \quad \text{pour } \gamma = \top(\alpha, \beta) \quad (6.1)$$

$$B(\varphi, \alpha) \vee B(\psi, \beta) \Leftrightarrow B(\varphi \vee \psi, \gamma) \quad \text{pour } \gamma = \perp(\alpha, \beta) \quad (6.2)$$

Les travaux sémantiques présentés en partie 1 ont permis de proposer dans la section 3.5, page 50, des versions pondérées des axiomes normaux à la sémantique proportionnelle proposée. Ils sont rappelés ici :

factorisation

$$(NA1_\alpha) \quad (\Box_\alpha \varphi \wedge \Box_\beta \psi) \rightarrow \Box_{\max(0, \alpha + \beta - 1)}(\varphi \wedge \psi)$$

$$(NA2_\alpha) \quad (\Box_\alpha \varphi \vee \Box_\beta \psi) \rightarrow \Box_{\min(\alpha, \beta)}(\varphi \vee \psi)$$

distribution

$$(NA1_\alpha) \quad \Box_\alpha(\varphi \wedge \psi) \rightarrow (\Box_\alpha \varphi \wedge \Box_\alpha \psi)$$

Plusieurs points sont à souligner lorsque l'on met en parallèle ces deux résultats. D'une part, concernant la factorisation, les implications gauche-droite des deux règles de manipulation (6.1) et (6.2) réalisent l'agrégation des degrés de croyance  $\alpha$  et  $\beta$  par une t-norme et une t-conorme. Les axiomes  $(NA1_\alpha)$  et  $(NA2_\alpha)$  proposent également une agrégation dont l'opérateur est imposé par la sémantique.

Cet opérateur est bien une t-norme pour la conjonction, plus précisément la t-norme de Łukasiewicz ; en revanche, pour la disjonction, c'est également une t-norme, la t-norme de Zadeh, et non une t-conorme comme induit par la sémantique floue. Le cas de la distribution montre également des différences entre les deux approches : un tel axiome n'existe pas pour la disjonction en logique modale, et donc en logique modale pondérée. De plus, pour la conjonction,  $(NA1_\alpha)$  associe le même poids  $\alpha$  aux deux modalités de la décomposition. Ce choix est en accord avec la règle de factorisation et il est compatible avec la règle (6.1) mais il réduit le degré de liberté, en excluant de différencier  $\alpha$  et  $\beta$ .

L'observation de cette différence entre les règles de manipulation basées sur le formalisme flou et les axiomes normaux pondérés ouvre des perspectives de validation, dans un contexte applicatif spécifique ou dans un contexte théorique, lié à des questions de correction et complétude.

## 6.2 Implication

La manipulation de croyances graduelles par rapport à l'implication est essentielle pour le raisonnement : elle permet d'effectuer des inférences dans l'ensemble des croyances. En

effet, si un l'agent croit  $\zeta$  un degré  $\alpha$  une règle sous forme d'implication, comme « si la vie existe dans une autre galaxie alors cette galaxie est nécessairement lointaine », alors il devrait être capable d'inférer, selon son degré de croyance  $\beta$  en la vie dans une autre galaxie, une croyance graduelle quant à la proximité de cette galaxie.

La formalisation de ce type d'inférence passe, en logique doxastique, par l'utilisation de l'axiome K :  $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ , qui se lit « si l'on croit une règle  $\varphi \rightarrow \psi$ , alors si l'on croit la prémisse  $\varphi$ , alors on croit la conclusion  $\psi$  ».

Cette section discute d'une adaptation de cet axiome au cas des croyances graduelles : il s'agit de déterminer à le degré auquel l'on croit  $\psi$  en fonction des degrés associés à la règle et à la prémisse dans le formalisme flou proposé, puis de le comparer à l'axiome  $(K)_\alpha$  établi dans le cadre des logiques modales pondérées de la partie 1.

### 6.2.1 Forme générale

Afin de proposer une variante pondérée de l'axiome K, nous le réécrivons sous la forme d'une unique implication en utilisant l'équivalence entre  $(A \rightarrow B) \rightarrow C$  et  $(A \wedge B) \rightarrow C$ , soit  $(\Box(\varphi \rightarrow \psi) \wedge \Box\varphi) \rightarrow \Box\psi$  : si l'on croit à la fois à une implication et sa prémisse, on croit à la conclusion de la règle.

On obtient ainsi comme candidat pondéré :  $(B(\varphi \rightarrow \psi, \alpha) \wedge B(\varphi, \beta)) \rightarrow B(\psi, \gamma)$ .

Ainsi, si l'on croit l'implication et sa prémisse à un niveau respectivement égal à  $\alpha$  et  $\beta$ , l'objectif ici est de combiner leur valeur afin de déterminer des conditions pour la valeur de  $\gamma$ .

En réécrivant la prémisse, on obtient :

$$\begin{aligned} B(\varphi \rightarrow \psi, \alpha) \wedge B(\varphi, \beta) &\leftrightarrow B((\varphi \rightarrow \psi) \wedge \varphi, \top(\alpha, \beta)) \\ &\leftrightarrow B(\varphi \wedge \psi, \top(\alpha, \beta)) \end{aligned}$$

La première ligne est obtenue en appliquant la factorisation de la modalité pondérée sur la conjonction, donnée par la règle de manipulation 6.1, s'écrit ainsi :

$$B(\varphi \wedge \psi, \top(\alpha, \beta)) \rightarrow B(\psi, \gamma)$$

En remplaçant dans le candidat pondéré, on obtient donc :

$$B(\varphi \rightarrow \psi, \alpha) \rightarrow (B(\varphi, \beta) \rightarrow B(\psi, \gamma)) \quad \text{pour } \gamma \geq \top(\alpha, \beta) \quad (6.3)$$

### 6.2.2 Comparaison avec l'axiome pondéré

Nous rappelons l'axiome pondéré établi dans les travaux sémantiques de la première partie :

$$K_\alpha \quad \Box(\varphi \rightarrow \psi, \alpha) \rightarrow (\Box(\varphi, \beta) \rightarrow \Box(\psi, \max(0, \alpha + \beta - 1)))$$

Si l'on met ce résultat en perspective de la règle obtenue précédemment, dans l'équation 6.3, on constate qu'elles proposent le même type de manipulation des degrés : combinant  $\alpha$  et  $\beta$  afin de déterminer le degré issu du mécanisme de modus ponens propre à cet axiome. (K) $\alpha$  utilisant pour cela la t-norme de Łukaziewicz, la proposition d'agrégation de l'équation 6.3 est compatible avec ce résultat.

## 6.3 Introspection

La logique doxastique comporte, enfin, les axiomes dits d'introspection positive et négative (cf. section 4.2, p. 71). Ceux-ci établissent respectivement les formules :  $\Box\varphi \rightarrow \Box\Box\varphi$  et  $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$ .

Cette section étudie l'extension de ces axiomes au cas des croyances graduelles : en soulignant les limites de l'interprétation floue, nous proposons des arguments sémantiques pour les contrebalancer.

### 6.3.1 Forme générale

Les formes graduelles des axiomes d'introspection, que l'on écrit respectivement  $B(\varphi, \alpha) \rightarrow B(B(\varphi, \gamma), \beta)$  et  $\neg B(\varphi, \alpha) \rightarrow B(\neg B(\varphi, \gamma), \beta)$ , la question est de déterminer des valeurs pertinentes pour les poids  $\alpha$ ,  $\beta$  et  $\gamma$ .

L'interprétation de telles formules lorsque les degrés de croyance sont considérés comme des degrés d'appartenance à l'ensemble flou des croyances est problématique car les formules comportent deux modalités successives : en effet, la lecture floue de la formule  $B(B(\varphi, \gamma), \beta)$  implique que la formule  $B(\varphi, \gamma)$   $\beta$ -appartient à l'ensemble de croyances, soit que  $\mu_{\mathcal{B}}(B(\varphi, \gamma)) = \beta$ .

Néanmoins, aucun opérateur flou n'autorise la manipulation de ce type d'objets récursifs. D'une part, il n'existe pas d'outil dans le formalisme flou qui permette de combiner ce degré d'appartenance par translation et combinaison avec le degré interne de la formule  $B(\varphi, \gamma)$  qui établit que  $\mu_{\mathcal{B}}(\varphi) = \gamma$ . D'autre part, nous ne pouvons pas non plus établir de relation entre cette appartenance et  $\mu_{\mathcal{B}}(\varphi) = \alpha$  qui découle de la prémisse de l'axiome d'inspection positive.

### 6.3.2 Arguments sémantiques

Une approche intuitive consiste à poser que  $\gamma = \alpha$ . Cela exprime un mécanisme intuitif d'introspection : l'information concernant les propres croyances de l'agent est conservée. En effet, cette formule indique que l'agent est capable d'inférer de nouvelles croyances au sujet de celles qui existent déjà. Il faut souligner que les croyances inférées seraient alors

affaiblies si  $\gamma < \alpha$ , ou enrichies si  $\gamma > \alpha$ . Le maintien de la même valeur pour  $\gamma$  et  $\alpha$  permet de stabiliser le niveau de la croyance considérée.

Nous proposons d'examiner le cas où  $\beta < 1$ . Une telle valeur pour  $\beta$  signifie que l'agent est partiellement convaincu de ses propres croyances : elle décrit le fait que l'agent n'a pas une croyance totale en ses croyances existantes, induisant de fait qu'il peut avoir des doutes à leur sujet. En d'autres termes, avec cette extension graduelle de l'introspection, l'agent transmet une certaine méfiance concernant ses propres croyances.

Ainsi, nous proposons de fixer  $\beta = 1$  : de cette façon, l'agent possède une croyance totale en ses propres croyances. Ce choix nécessite d'utiliser la modalité de croyance classique et ouvre alors une discussion sur la limite entre le plus haut niveau de croyance possible et la notion de savoir. Cette idée a déjà été proposée par Laverny et Lang (2004). Néanmoins le modèle présenté ici, qui est basé sur le degré d'appartenance à un ensemble de croyance, est placé dans un contexte différent : nous proposons de modéliser le raisonnement sur des croyances graduelles dans une version adaptée de KD45.

En ce qui concerne l'axiome d'introspection négative, en raison de l'ajout d'une négation à sa prémisse de l'axiome d'introspection négative, celui-ci fait également appel à l'application de la négation aux formules de croyances graduelles, discutée dans le chapitre suivant. Néanmoins, en considérant la combinaison de modalités induite, son fonctionnement est identique au cas de l'introspection positive.

### 6.3.3 Comparaison avec les axiomes pondérés

La règle de manipulation obtenue après la discussion de la section précédente est donc :

$$B(\varphi, \alpha) \rightarrow B(B(\varphi, \alpha), 1)$$

Dans les travaux sémantiques effectués dans la première partie de ce manuscrit, nous avons proposé une version pondérée de l'axiome 4, d'introspection positive, et une version impondérable de l'axiome 5, d'introspection négative. Par conséquent, nous ne pouvons discuter que de l'axiome 4 car il est le seul dont nous donnons une extension pondérée.

La version pondérée de cet axiome, obtenue par la sémantique proportionnelle, est la suivante :  $\Box_\alpha \varphi \rightarrow \Box_1 \Box_\alpha \varphi$ . Il est notable que cette formule est strictement identique à celle proposée ci-dessus. Les travaux sémantiques et flous se rejoignent sur l'introspection positive.

## 6.4 Bilan

Ce chapitre a proposé une discussion de certains axiomes de la logique doxastique afin d'en donner une version pondérée. Pour cela, nous nous basons à la fois sur des arguments

---

logiques, visant à exploiter les propriétés des connecteurs logiques et des opérateurs de combinaison des degrés ; et sur des arguments sémantiques, qui s'intéressent au sens donné aux degrés ainsi qu'à leur interprétation dans le contexte d'un ensemble de croyances flou.

Pour chacun des axiomes présentés, nous proposons une extension pondérée qui s'inscrit dans l'utilisation du modèle flou proposé dans le chapitre précédent. Les propriétés de ce modèle sont exploitées afin de justifier des choix de pondération des axiomes.



## Chapitre 7

# Négation de croyances graduelles

La question de la négation de croyances graduelles fait ici l'objet d'un chapitre dédié, en raison de la richesse des interprétations qu'on peut lui associer : comme discuté par la suite, on peut en effet distinguer les cas différents, selon que la négation porte sur la formule, le degré ou la modalité elle-même. Il conduit à considérer des cadres formels différents, au-delà du seul cadre flou, pour modéliser cette diversité d'interprétation.

La section 7.1 formalise la tâche de négation de croyances graduelles et les trois interprétations sont étudiées tour à tour dans les sections 7.2 à 7.4. Dans chacun des cas, le principe en est d'abord détaillé, le choix d'un cadre formel approprié est discuté, les outils qu'il offre pour la modélisation de la négation sont rappelés, en soulignant leur expressivité et leur richesse ; enfin, leur transposition au cas des croyances graduelles est examinée.

Les travaux présentés dans ce chapitre ont fait l'objet des publications (Legastelois et al., 2016b; Legastelois et al., 2016c).

### 7.1 Formalisation de la négation

La négation d'une croyance graduelle  $B(\varphi, \alpha)$  s'exprime linguistiquement comme « il n'est pas vérifié que  $\varphi$  est crue à un degré  $\alpha$  ». Nous proposons de la noter  $\mathcal{N}(B(\varphi, \alpha))$ , où  $\mathcal{N}$  est un opérateur de négation général dont l'interprétation et la proposition de modélisation adéquate fait l'objet de ce chapitre.

Cette notation spécifique souligne la différence avec la négation classique  $\neg$  : celle-ci conduirait à écrire  $\neg B(\varphi, \alpha)$  et pourrait amener à transposer des résultats classiques associés à l'opérateur  $\neg$  au cas des croyances graduelles. Nous choisissons de traiter ce problème de négation au-delà de la négation classique afin de prendre en compte les particularités sémantiques de la notion de croyance graduelle.

En effet, la négation d'une croyance graduelle  $B(\varphi, \alpha)$  peut être interprétée de diverses manières, selon la composante sur laquelle porte la négation. Formellement, nous proposons

d'examiner trois formules :

	transfert de $\mathcal{N}$ sur
$B(\mathcal{N}_1(\varphi), \beta)$	la formule
$B(\varphi, \mathcal{N}_2(\alpha))$	le degré
$\mathcal{N}_3(B)(\varphi, \gamma)$	la modalité

$\mathcal{N}_1, \mathcal{N}_2$  et  $\mathcal{N}_3$  sont trois déclinaisons de l'opérateur global de négation  $\mathcal{N}$ , qui sont discutées dans les sections suivantes. Les degrés, possiblement distincts,  $\alpha, \beta, \gamma$  et  $\mathcal{N}_2(\alpha)$  sont utilisés pour ne pas imposer de contraintes a priori sur leurs relations ou leurs égalités, également discutées ci-dessous.

À titre illustratif, dans ce chapitre, nous considérons l'exemple introductif du chapitre et l'assertion « l'agent croit fermement que les licornes existent » réécrite  $B(\varphi, 0.8)$ .

## 7.2 Négation de la formule

Nous nous intéressons tout d'abord au cas où la négation s'applique à la formule que l'on croit. Après avoir donné le principe de cette interprétation, nous montrons qu'elle s'inscrit naturellement dans le cadre formel des logiques modales pondérées.

### 7.2.1 Principe

Lorsque l'on considère le transfert de la négation à la formule crue,  $\varphi$ , c'est-à-dire  $B(\mathcal{N}_1(\varphi), \gamma)$ , la négation s'applique à une formule logique, elle est donc naturellement interprétée comme la négation logique :  $\mathcal{N}_1 \equiv \neg$ .

La question est alors de savoir si l'assertion  $\mathcal{N}(B(\varphi, \alpha))$ , lue « il n'est pas vrai que  $\varphi$  est crue à un degré  $\alpha$  », permet de tirer des conclusions sur la croyance en  $\neg\varphi$ , formellement notée  $B(\neg\varphi, \beta)$  : alors que le tiers exclu ne permet pas à  $\varphi$  et  $\neg\varphi$  d'être simultanément vraies, des croyances partielles en  $\varphi$  et  $\neg\varphi$  sont possibles simultanément, conduisant à examiner des contraintes éventuelles sur leurs degrés respectifs. Cette question fait donc appel aux axiomes C et CD des logiques modales.

En considérant l'exemple illustratif, si « il n'est pas vrai qu'un agent croit fermement que les licornes existent », croit-il que les licornes *n'existent pas* et, si c'est le cas, dans quelle mesure ?

### 7.2.2 Négation en logiques modales

**Logiques modales classiques** La combinaison de l'opérateur modal  $\Box$  avec la négation soulève la question des relations entre les formules  $F_1 = \Box\varphi$ ,  $F_2 = \neg\Box\varphi$ ,  $F_3 = \Box\neg\varphi$  et, en appliquant à la fois les négations interne et externe, c'est-à-dire avant et après l'opérateur

modal,  $F_4 = \neg\Box\neg\varphi$ , soit  $F_4 = \Diamond\varphi$ . Le fait qu'il existe ces quatre variantes souligne la richesse de sens que la négation peut exprimer en logique modale.

L'application à la formule  $\Box\varphi$  de la loi du tiers exclu, par compatibilité avec la logique classique, implique que  $\Box\varphi \vee \neg\Box\varphi$  est une tautologie, ce qui établit une relation entre  $F_1$  et  $F_2$ , et, de façon similaire, entre  $F_3$  et  $F_4$ .

Les relations entre  $F_1$  et  $F_4$  sont établies par les axiomes  $(D)$  et  $(CD)$ , réciproques l'un de l'autre :

$$(D) \quad \vdash \Box\varphi \rightarrow \neg\Box\neg\varphi \quad (7.1)$$

$$(CD) \quad \vdash \neg\Box\neg\varphi \rightarrow \Box\varphi \quad (7.2)$$

En les appliquant à  $\neg\varphi$  et en utilisant la propriété de double négation  $\neg\neg\varphi = \varphi$ , on établit des relations semblables entre  $F_2$  et  $F_3$ .

Nous rappelons qu'en termes doxastiques, l'axiome  $(D)$  exprime le fait que, si  $\varphi$  est crue, alors son contraire ne l'est pas : il est impossible de croire à la fois une formule et son contraire.  $(D)$  est donc considéré comme modélisant une des facettes de la rationalité supposée de l'agent dont les croyances sont représentées. Il est appelé axiome de cohérence et inclus dans la définition axiomatique classique de la logique doxastique, KD45 (Hintikka, 1962).

Sa réciproque,  $(CD)$ , exprime une cohérence similaire, mais complémentaire, en exprimant que si une formule n'est pas crue, son contraire l'est : sa prémisse s'applique à une absence de croyance et sa conclusion à une croyance. En réécrivant l'implication avec une négation et une disjonction,  $(CD)$  impose donc que l'agent croie soit une formule, soit son contraire, conduisant à une forme modale du tiers exclu selon laquelle  $\Box\neg\varphi \vee \Box\varphi$  est une tautologie. Cette contrainte ne permet donc pas de modéliser un état d'esprit neutre, perplexe ou indéterminé, correspondant à une absence de croyance.  $(CD)$  est, en général, considéré comme trop restrictif et exclu de la définition axiomatique de la logique doxastique.

**Logiques modales pondérées** Dans le cadre des logiques modales pondérées examinées dans la première partie de nos travaux (voir sections 3.5.3, page 53 et 3.7.1, page 59), nous avons établi les variantes pondérées de D et CD rappelées ci-dessous :

$$(D)_\alpha : \Box_\alpha\varphi \rightarrow \neg\Box_\beta\neg\varphi$$

$$(CD)_\alpha : \Diamond_\alpha\varphi \rightarrow \Box_{1-\alpha}\varphi$$

### 7.2.3 Transposition aux croyances graduelles

L'extension pondérée de l'axiome ( $D$ ) constitue un candidat pertinent pour la relation désirée entre la négation de croyance globale  $\mathcal{N}(B(\varphi, \alpha))$  et la croyance en la formule niée  $B(\neg\varphi, \beta)$  : en appliquant ( $D_\alpha$ ) à  $\neg\varphi$  et en considérant  $\mathcal{N}$  (resp.  $\neg$ ) comme la négation externe (resp. interne) par rapport à l'opérateur modal, on obtient

$$B(\neg\varphi, \alpha) \rightarrow \mathcal{N}(B(\varphi, \beta))$$

pour tous  $\alpha, \beta \in [0, 1]$  tels que  $\beta > 1 - \alpha$ .

Cette formule peut être lue comme 'si  $\neg\varphi$  est crue à un degré  $\alpha$ , alors il n'est pas vrai que  $\varphi$  est crue à un degré  $\beta$ '. Ainsi,  $\varphi$  et  $\neg\varphi$  peuvent être partiellement crues simultanément, à condition que leurs degrés respectifs vérifient la contrainte d'inégalité  $\beta > 1 - \alpha$ .

## 7.3 Négation du degré

Cette section présente le cas de l'application de la négation au degré de croyance. Après en avoir donné les principes généraux, nous proposons d'interpréter cette négation dans le cadre de la logique floue. Cette section présente également la transposition de la négation floue à la négation de degrés de croyance.

### 7.3.1 Principe

Lorsque la négation de croyances graduelles est transférée sur le degré, « *il n'est pas vrai que  $\varphi$  est crue à un degré  $\alpha$*  » signifie que 'l'agent croit  $\varphi$ ' peut être vérifié, mais à un degré différent de  $\alpha$ . Ainsi, en reprenant l'exemple illustratif précédent, s'il n'est pas vrai que l'agent croit que les licornes existent avec un degré 0.8, avec quel degré le croit-il? Aussi, la négation,  $\mathcal{N}_2$ , s'applique aux degrés de croyance et désigne un opérateur de négation numérique à définir.

Une interprétation littérale suggère que toute valeur autre que  $\alpha$  convient. Elle pourrait conduire à définir  $\mathcal{N}_2(\alpha)$  comme un ensemble et non une valeur, par exemple  $\mathcal{N}_2(\alpha) = [0, 1] \setminus \{\alpha\}$ . Cette interprétation conduit à une compréhension très imprécise et non informative de l'assertion. Nous ne la considérons pas plus avant et proposons de définir  $\mathcal{N}_2$  comme une fonction de  $[0, 1]$  dans  $[0, 1]$ , de telle manière que nier une croyance graduelle fournisse une information sur cette croyance –et non sur la croyance dans la négation de formule, discutée dans la section précédente.

### 7.3.2 Négation en logique floue

Les opérateurs de négation numériques ont été en particulier étudiés dans le cadre de la logique floue et de la théorie des sous-ensembles flous (Zadeh, 1965), pour représenter la négation logique et l'opération ensembliste de complémentation. De façon générale, un tel opérateur est défini par trois propriétés (Weber, 1983) : c'est une fonction (i)  $n : [0, 1] \rightarrow [0, 1]$ , (ii) décroissante et (iii) qui vérifie les conditions aux limites  $n(0) = 1$  et  $n(1) = 0$ . Ce principe de négation peut alors être écrit, pour un degré associé  $d$  :  $d(\neg\varphi) = n(d(\varphi))$  (cf. section 7.4.2, page 103).

On peut noter que la définition générale des opérateurs de négation, rappelée ci-dessus, n'impose pas la contrainte d'involutivité, ce qui constitue une différence importante avec le cas classique pour le traitement des doubles négations logiques : le degré de vérité de  $\neg\neg\varphi$  peut être différent de celui de  $\varphi$ .

La définition la plus classique (Zadeh, 1965),  $n(x) = 1 - x$  est, toutefois, à la fois involutive et stricte (strictement décroissante et continue). On peut également citer (pour une liste complète, voir Weber (1983))  $n(x) = \sqrt{1 - x^2}$  ou  $n(x) = 1 - x^2$  qui est stricte mais non involutive ou l'opérateur de Gödel, défini par  $n(x) = 0$  si  $x > 0$  et  $n(0) = 1$ , et son dual, défini par  $n(x) = 1$  si  $x < 1$  et  $n(1) = 0$ . Ces derniers ne sont ni continus ni involutifs, on peut noter qu'ils constituent des choix drastiques : leur utilisation restreint les valeurs de vérité aux booléens  $\{0, 1\}$ . Ainsi, la variété des fonctions vérifiant la définition générale des négations floues permet de définir des postures de négation complexes, plus riches que les comportements du cas classique.

On peut noter aussi que la logique floue offre deux types de négation pour les prédicats flous définis à partir d'une partition floue, distinguant compléments et antonymes (voir par exemple de Soto & Trillas, 1999; Moysse et al., 2015). Formellement, étant donné un prédicat flou défini sur un univers numérique de la forme  $[a^-, a^+]$ ,  $A : [a^-, a^+] \rightarrow [0, 1]$ , et un opérateur  $n$  de négation floue, le complément et l'antonyme de  $A$  sont respectivement définis par  $\bar{A}(x) = n(A(x))$  et  $\hat{A}(x) = A(a^+ - a^- - x)$ .

### 7.3.3 Transposition aux croyances graduelles

Le choix d'un cadre flou pour l'interprétation des degrés de croyance, comme examiné dans les chapitres 5 et 6, conduit ici à poser  $\mathcal{N}_2 \equiv n$  et à examiner la pertinence des propriétés des opérateurs de négation flous : les conditions aux limites et la monotonie, ainsi qu'optionnellement l'involutivité.

**Conditions aux limites** L'interprétation doxastique des conditions aux limites des opérateurs de négation floue peut être illustrée par la mise en correspondance des assertions

« l'agent croit à un degré  $n(1)$  que les licornes existent » et « l'agent croit à un degré 0 que les licornes existent » (et vice-versa) et dépend donc de l'interprétation des degrés. Si l'on considère  $B(\varphi, 1)$  comme équivalent à  $\Box\varphi$ , les conditions aux limites peuvent être interprétées comme préservant la compatibilité avec le cas classique de l'expression modale du tiers exclu, selon laquelle  $\Box\varphi \vee \neg\Box\varphi$  est une tautologie.

Nous proposons de réécrire ce tiers exclu pour tenir compte des formules de croyances graduelles :  $B(\varphi, 1) \vee \mathcal{N}(B(\varphi, 1))$ , soit, en appliquant la négation au degré,  $B(\varphi, 1) \vee B(\varphi, \mathcal{N}(1))$ . En proposant  $\mathcal{N}$  comme une négation floue, alors  $\mathcal{N}(1) = 0$ . Ainsi, conformément au tiers-exclu, on a soit  $B(\varphi, 1)$ , soit  $B(\varphi, 0)$ .

**Monotonie** D'un point de vue doxastique, la monotonie des opérateurs de négation floue correspond à la comparaison des degrés de croyance et leur négation : considérons les deux croyances graduelles  $B_1$  : 'l'agent croit à un degré 0.8 que les licornes existent' et  $B_2$  : 'l'agent croit à un degré 0.6 que les farfadets existent'. La négation de ces deux croyances avec un opérateur flou, et donc décroissant, implique que  $\mathcal{N}(B_1)$  est cru, au plus, autant que  $\mathcal{N}(B_2)$ . Aussi, la propriété de monotonie semble compatible avec les comportements souhaités.

**Involutivité** Enfin, une propriété non imposée mais offerte par certains opérateurs de négation floue est l'involutivité : choisir de croire  $\mathcal{N}(\mathcal{N}(B_1))$  autant que  $B_1$  peut être compris comme le refus de gagner ou de perdre en croyance lorsqu'il est faux qu'on n'accepte pas le degré de croyance en  $B_1$ . Même si cette expression n'est pas naturelle, la notion sous-jacente semble correspondre à un comportement souhaitable. Une négation floue involutive est donc en accord avec notre interprétation de la négation des degrés de croyance.

De plus, cette proposition est cohérente avec la représentation floue présentée en chapitre 5 : la négation floue utilisée pour les degrés d'appartenance est involutive.

## 7.4 Négation de la modalité

Cette section s'intéresse au dernier cas d'application de la négation à une formule doxastique graduelle : la négation appliquée à la modalité, ce qui est relatif à la notion d'incroyance introduite en chapitre 5, page 82. Après avoir exposé les principes de cette négation, nous proposons de nous placer dans le cadre de la logique multi-valente. Cette section présente enfin la transposition de la négation de la logique multi-valente au cas doxastique graduel.

### 7.4.1 Principe

La troisième interprétation de la négation des croyances graduelles transfère la négation à la modalité : l'assertion négative « *il n'est pas vrai que  $\varphi$  est crue à un degré  $\alpha$*  » est vue comme fournissant une information sur une non-croyance, requérant une modalité additionnelle, en plus de la modalité de croyance. Notons que la langue anglaise dispose, pour un tel cas, d'un verbe spécifique, *disbelieve*, sans autre équivalent français que le désuet *décroire*.

Afin d'augmenter l'expressivité, dans le cas où les degrés de croyance sont exprimés linguistiquement par des adverbes, nous proposons de considérer plusieurs modalités, combinant intrinsèquement les modalités avec le niveau, en distinguant par exemple des modalités « croire fermement » et « croire modérément ». Ainsi la croyance graduelle est représentée par des modalités distinctes, bien que comparables, plutôt qu'une unique modalité nuancée par des degrés continus : cette interprétation consiste à considérer un ensemble fini de modalités, ayant plus des 2 modalités, croyance et non-croyance (ou incroyance), déjà évoquées.

### 7.4.2 Cadre formel considéré : logique multi-valente

**Rappels** La logique multi-valente (Rescher, 1968) considère comme valeurs de vérité admissibles l'ensemble totalement ordonné  $\mathcal{L}_M = \{\tau_0, \dots, \tau_{M-1}\}$ , où  $M$  est un entier positif prédéfini et  $\forall \alpha, \beta \in \{0, \dots, M-1\}, (\alpha \leq \beta \Leftrightarrow \tau_\alpha \leq \tau_\beta)$ .  $\mathcal{L}_M$  représente des valeurs de vérité intermédiaires entre « faux »,  $\tau_0$ , et « vrai »,  $\tau_{M-1}$ , à différents niveaux de granularité, dépendant du nombre total de niveaux,  $M$ .

L'une des motivations de la logique multi-valente est sa capacité à fournir une représentation linguistique des degrés de vérité, reposant sur une mise en correspondance entre l'échelle discrète ordonnée  $\mathcal{L}_M$  et un ensemble d'étiquettes linguistiques, par exemple basé sur des adverbes qualifiant le degré de vérité. Ainsi,  $\mathcal{L}_5$  est en bijection avec l'ensemble {« faux », « plutôt faux », « ni faux ni vrai », « plutôt vrai », « vrai »}.

**Négation en logique multi-valente** Comme dans le cas de la logique floue, le principe de négation peut être écrit, pour un degré associé  $d$  :  $d(\neg\varphi) = n(d(\varphi))$ , où l'opérateur de négation multi-valent est classiquement défini comme :

$$\begin{aligned} n : \mathcal{L}_M &\rightarrow \mathcal{L}_M \\ \tau_i &\mapsto \tau_{M-1-i} \end{aligned}$$

Cette extension préserve la compatibilité avec la logique classique : la négation de « faux »,  $n(\tau_0) = \tau_{M-1}$ , est « vrai », et vice-versa. Pour les degrés intermédiaires, l'opérateur de négation peut être vu comme calculant la valeur symétrique par rapport au milieu de

l'échelle  $\tau_{\frac{M-1}{2}}$  : il peut être interprété selon la notion linguistique d'antonymie, évoquée précédemment (Muehleisen, 1997).

Dans le cas où le nombre de niveaux,  $M$ , est impair, c'est-à-dire quand cette valeur médiane fait partie de l'échelle, elle est égale à sa propre négation  $n(\tau_{\frac{M-1}{2}}) = \tau_{\frac{M-1}{2}}$ . Aussi, une formule ni fausse ni vraie a le même degré de vérité que sa négation.

Ce comportement spécifique de la négation, qui autorise une formule et sa négation à être simultanément vraies dans une certaine mesure, peut être interprété comme une généralisation de la loi de non-contradiction.

### 7.4.3 Transposition aux croyances graduelles

**Approche bivalente : croyance et non-croyance** L'introduction d'une modalité de non-croyance correspond à l'interprétation de « *il n'est pas vrai que  $\varphi$*  est crue à un degré  $\alpha$  » comme ' $\varphi$  est non-crue à un degré  $\beta$ ', transférant la négation globale à l'opérateur modal et, simultanément, permettant un effet sur les degrés associés, afin de ne pas limiter a priori l'expressivité de l'interprétation de la négation. Formellement, on peut introduire une modalité  $Db$  (pour *disbelieve*), poser  $\mathcal{N}_3(B) \equiv Db$ , c'est-à-dire écrire la négation  $\mathcal{N}(B(\varphi, \alpha)) = Db(\varphi, \beta)$ .

Il faut noter que le coefficient  $\beta$  décrit alors un degré de non-croyance, et non un degré de croyance, ce qui induit une différence importante par rapport aux cas précédents, où tous les degrés ont la même signification et sont mesurés sur une même échelle. Comme, néanmoins, croyance et non-croyance sont liées, les degrés associés conduisent à considérer une échelle signée plutôt que deux échelles indépendantes. Cette interprétation a été mentionnée dans la section 7.3.3, où l'interprétation d'une 0-croyance comme non-croyance est suggérée. Elle conduit à la question de savoir à quel degré, autour de quel  $\alpha$ , une croyance se transforme en une non-croyance.

L'interprétation floue discutée dans la section précédente peut être vue comme un cas extrême, où l'intervalle ouvert  $]0, 1]$  représente différents degrés de croyance et une unique valeur, 0, est considérée pour la non-croyance.

On peut néanmoins attribuer une plage plus importante pour les degrés de non-croyance. Il peut sembler naturel de considérer comme valeur-pivot 0.5, située au milieu de l'échelle :  $]0.5, 1]$  représente alors des degrés de croyance intermédiaires,  $[0, 0.5[$  des degrés de non-croyance et 0.5 une valeur neutre. On peut toutefois aussi choisir une autre valeur, pour modéliser différentes attitudes vis-à-vis des croyances, en donnant plus d'importance à l'une ou l'autre des modalités.

L'expression de  $\beta$  en fonction de celle de la valeur du degré  $\alpha$  peut alors se faire par le biais d'une fonction vérifiant la contrainte d'être croissante. Ainsi organisée autour d'une

TABLEAU 7.1 – Interprétations de la négation de croyances graduelles

Cadre formel	Expression de la négation	Contrainte sur les degrés
logique modale pondérée	$B(\neg\varphi, \alpha) \rightarrow \mathcal{N}(B(\varphi, \beta))$	$\alpha, \beta \in [0, 1]$ et $\beta > 1 - \alpha$
logique floue	$\mathcal{N}(B(\varphi, \alpha)) = B(\varphi, n(\alpha))$	p. ex. $n(\alpha) = 1 - \alpha$
logique bimodale	$\mathcal{N}(B(\varphi, \alpha)) = Db(\varphi, f(\alpha))$	$f$ fonction croissante
logique multimodale	$\mathcal{N}(B_t(\varphi)) = B_{n(t)}(\varphi)$	$t \in \mathcal{L}_M, n(\tau_i) = \tau_{M-1-i}$

valeur-pivot, cette fonction croissante permet, de par la valeur du degré, de décrire une attitude de croyance ou d'incroyance. Un exemple trivial est la fonction  $f$  définie sur  $[0, 1] \times [0, 1]$  telle que  $f : x \mapsto 1 - x$ .

**Approche multi-valente : multiples modalités de croyance** La représentation de la croyance graduelle combinant intrinsèquement les modalités avec le niveau conduit naturellement à exploiter une formalisation dans un cadre de logique multi-valente, en définissant formellement des modalités  $B_t, t \in \mathcal{L}_M = \{\tau_0, \dots, \tau_{M-1}\}$ . On peut par exemple considérer 5 niveaux de croyance, représentées par des étiquettes linguistiques, par exemple {« faible », « modérée », « moyenne », « ferme », « forte »}.

La négation consiste alors à définir, pour tout  $t \in \mathcal{L}_M, \mathcal{N}_3(B_t) = B_{n(t)}$  avec l'opérateur de négation multi-valente : la négation est appliquée à l'adverbe qui représente une combinaison de la modalité avec le degré et peut donc être interprétée en terme d'antonymes linguistiques.

Il faut noter que ce cas est équivalent au cadre flou discuté dans la section 7.3, après une discrétisation des degrés permettant leur manipulation linguistique. La différence essentielle avec l'approche bivalente précédente est liée l'interprétation du degré 0 et de la non-croyance.

## 7.5 Bilan

Les différentes interprétations proposées pour la négation appliquée à une formule de croyance graduelle font appel à plusieurs formalismes présentés : logique modale, logique floue, logique multi-valente. Trois interprétations de cette négation ont été données, selon son objet : la formule à propos de laquelle une assertion modale est exprimée, la modalité qui exprime la croyance ou le degré représentant le degré de croyance. Le tableau 7.1 récapitule les interprétations proposées et les formalismes associés.

Bien que les objets de la négation soient considérés séparément, ils apparaissent étroitement liés, en particulier le degré de croyance qui subit des modifications dans tous les

cas : la manipulation de croyances graduelles apparaît comme étant principalement liée à la définition de règles de transformation des degrés. De la même façon, nous observons que la transformation proposée est intimement liée à la négation floue discutée en section 7.3. Ces résultats sont cohérent avec l'interprétation floue des degrés de croyance, présentée en chapitre 5.

Troisième partie

Applications



# Introduction

Cette dernière partie est consacrée aux applications effectuées dans le cadre de la thèse, qui, comme les parties théoriques, sont de deux types : la première se place dans le cadre des logiques modales pondérées générales, la seconde dans le cadre des croyances graduelles.

Le chapitre 8 présente la mise en œuvre de la sémantique proportionnelle proposée, dans le chapitre 2 : il décrit un outil de vérification des formules modales pondérées dans un modèle de Kripke, qui est également muni d'un système de génération de modèles sous contraintes.

Le chapitre 9 décrit l'exploitation de croyances graduelles dans le cadre d'un jeu coopératif à information incomplète, nommé Hanabi : les joueurs s'échangent des informations à l'aide d'annonces publiques qui font évoluer leurs croyances concernant les cartes dont ils disposent. Le chapitre décrit une IA permettant de jouer, basée sur une formalisation de l'état courant du jeu en logique doxastique graduelle.



## Chapitre 8

# Outil de vérification de formules modales pondérées

Ce chapitre présente un outil implémentant la sémantique modale pondérée que nous avons proposée, décrite dans le chapitre 2, page 27 : il en offre la mise en pratique en permettant de calculer la satisfiabilité d'une formule modale pondérée dans un modèle de Kripke donné. Il permet, de plus, de générer aléatoirement des modèles de Kripke respectant des propriétés souhaitées de la relation d'accessibilité. Pour des modèles générés sans contrainte, l'outil développé permet de caractériser la relation et d'identifier les propriétés qu'elle satisfait. L'implémentation de cet outil a fait l'objet du stage d'été d'Étienne Sahel, étudiant de L3 informatique à l'UPMC.

Ce chapitre est organisé comme suit : il donne tout d'abord une présentation générale du logiciel ; la deuxième section présente l'outil de génération et de caractérisation de modèles de Kripke. Enfin, la troisième section est consacrée à l'outil de vérification de formules modales pondérées.

### 8.1 Présentation générale

Le logiciel est composé d'une interface graphique, d'un moteur de vérification des formules et d'un générateur de graphes, ces deux derniers éléments étant détaillés dans les prochaines sections. La figure 8.1 présente une vue d'ensemble de l'interface graphique du logiciel, partagée en deux parties successivement décrites ci-dessous : une zone d'affichage et une zone d'interaction.

**Zone d'affichage** La zone supérieure de l'interface graphique permet de représenter et de réorganiser à la souris un modèle de Kripke  $\langle W, R, s \rangle$ . Celui-ci est naturellement affiché

sous la forme d'un graphe dont les sommets représentent les mondes possibles et les arcs la relation d'accessibilité.

Un monde  $w \in W$  est représenté par un rectangle que l'utilisateur peut sélectionner et déplacer à la souris. Le coin supérieur gauche indique l'identifiant du monde, sous la forme d'un entier. Comme illustré sur la figure 8.2, sur demande de l'utilisateur, le coin inférieur droit peut donner les variables propositionnelles valuées à vrai (les variables non affichées étant implicitement à faux). La fonction de valuation  $s$  peut ainsi être représentée dans l'interface.

Deux mondes  $u$  et  $v$  en relation  $uRv$  sont liés par un segment dont une extrémité comporte un cercle pour préciser le monde destinataire  $v$ . Si les deux extrémités comportent un cercle, comme c'est le cas entre les mondes 3 et 4 de la figure 8.1 par exemple, alors la relation est telle que  $uRv$  et  $vRu$ . Un monde en relation avec lui-même est représenté par un cercle au centre du rectangle, comme c'est le cas pour les mondes 0, 1, 3 et 4, sur la figure 8.1.

De plus, comme illustré sur la figure 8.2 pour le cas du monde 3, si un monde  $w$  est sélectionné, un affichage dynamique indique en bleu les liens entrants, de la forme  $uRw$ , en rouge les liens sortants ( $wRu$ ) et en violet les liens symétriques (tels que  $uRw$  et  $wRu$ ).

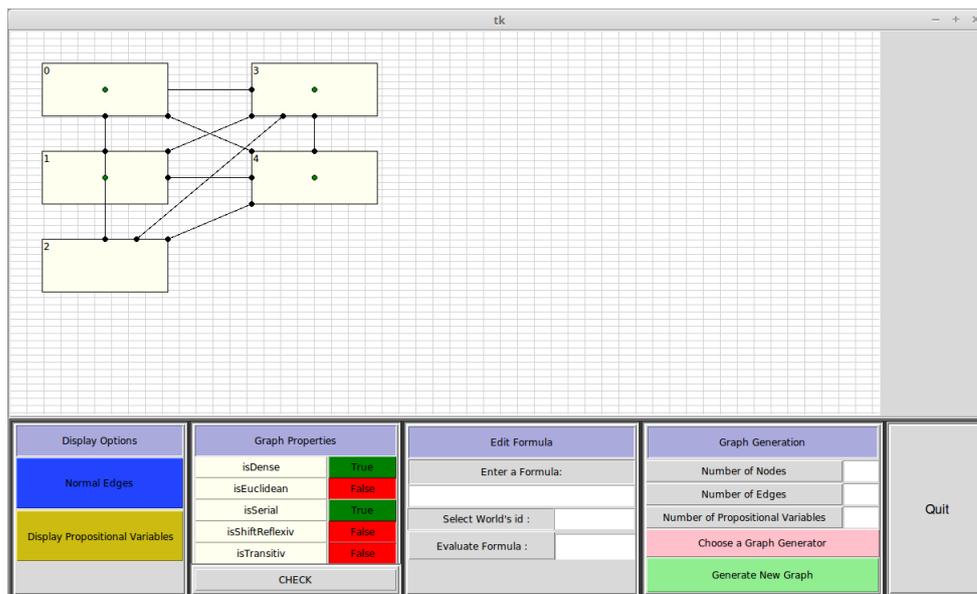


FIGURE 8.1 – Vue d'ensemble de l'interface graphique.

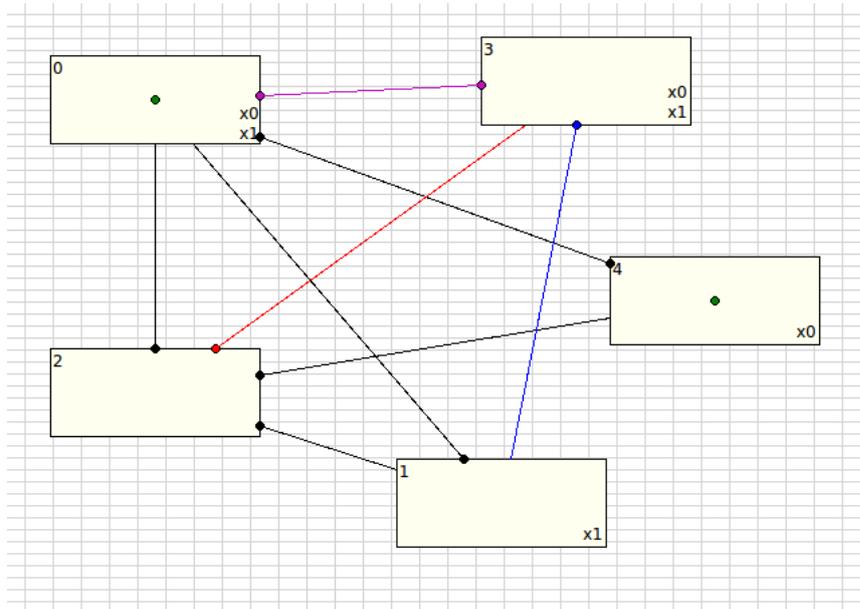


FIGURE 8.2 – Représentation d'un cadre de Kripke après sélection d'un monde, ici le monde 3 : les rectangles sont les mondes, les liens bleus (resp. rouges) sont les entrants (resp. les sortants) et les liens violets sont les liens symétriques.

**Zone d'interaction** Dans la zone sous l'aire de dessin, cinq cadres sont affichés et contiennent différents outils, de gauche à droite :

1. le cadre « Display Options » offre un outil de remise à zéro de l'affichage après les modifications provoquées par un clic sur un monde, concernant la couleur des relations (remises en noir avec le bouton « Normal Edges ») et la valuation des formules propositionnelles (effacées avec le bouton « Display Propositional Variables »).
2. le cadre « Graph Properties » permet de vérifier les propriétés sur la relation du modèle : un clic sur le bouton « CHECK » indique si la relation est dense, euclidienne, sérielle, shift-réflexive ou transitive.
3. le cadre « Edit Formula » permet de faire appel au vérificateur de formules, en précisant un monde et une formule, un clic sur le bouton « Evaluate formula » donne le résultat de la vérification. Cet outil est détaillé dans la section 8.3, page 115.
4. le cadre « Graph Generation » permet d'utiliser le générateur de graphes implémenté, décrit, ainsi que les champs associés, dans la section 8.2.
5. le cadre « Quit » contient un bouton qui permet de quitter le programme.

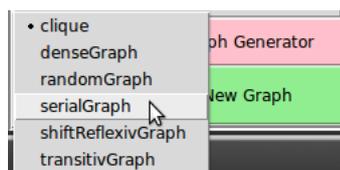


FIGURE 8.3 – Menu déroulant qui permet de choisir le type de relation voulue.

## 8.2 Génération et caractérisation de modèles

La zone « Graph Generation » de l'interface du logiciel proposé permet de générer des modèles de Kripke : elle permet, d'une part, de créer dynamiquement des cadres de Kripke, qui respectent des propriétés particulières sur la relation, et d'autre part, d'attribuer une valuation des variables propositionnelles aux mondes.

Cette section décrit les principes de génération des modèles ainsi que les algorithmes employés. Elle présente également l'outil de caractérisation de la relation, accessible dans la zone « Graph Properties » de l'interface.

### 8.2.1 Génération

Trois modes de génération sont proposés, accessibles par le bouton rose « choose a graph generator » de l'interface (voir figure 8.1, p. 112), qui ouvre un menu déroulant proposant une liste de générateurs : d'abord, un mode contraint permet à l'utilisateur de sélectionner la propriété que la relation d'accessibilité doit vérifier. Comme illustré sur la figure 8.3, un menu déroulant permet de choisir si la relation doit être dense, euclidienne, sérielle, shift-réflexive ou transitive.

Le second mode, accessible dans la liste des générateurs, correspond à une génération aléatoire des liens entre les mondes, sans forcer de propriété en particulier.

Enfin, le troisième mode correspond au générateur de clique, également accessible dans la liste des générateurs, qui construit un cadre de Kripke possédant un lien entre toutes les paires de mondes.

Au-dessus du bouton de sélection du mode de génération, trois champs sont disponibles pour permettre à l'utilisateur de configurer le modèle de Kripke souhaité : le premier indique le nombre de mondes désiré, le second le nombre de liens. Cette valeur n'est exploitée que pour le mode de génération aléatoire, qui tire au hasard des paires différentes de mondes source et destination pour chaque lien requis. Si le nombre entré par l'utilisateur est supérieur au nombre maximum de liens possibles, alors le graphe généré est une clique, avec un nombre de liens égal à ce maximum. Pour le générateur contraint et le générateur

de cliques, le nombre d'arcs est déterminé par les propriétés de la relation.

Le troisième champ permet de spécifier le nombre total de variables propositionnelles à considérer dans le cadre de Kripke construit : comme mentionné précédemment, la valuation des variables dans chaque monde est tirée aléatoirement.

La figure 8.4 représente, à titre illustratif, un modèle aléatoire généré avec 8 mondes, 16 liens et 4 variables propositionnelles.

### 8.2.2 Caractérisation

Pour le cas où le modèle est généré aléatoirement, nous avons ajouté un outil de caractérisation de la relation d'accessibilité : grâce à cet outil de vérification des relations, il est possible de lister les propriétés du modèle.

L'implémentation de cet outil combine deux approches, qui énumèrent respectivement les liens et les mondes du cadre de Kripke considéré. La première effectue une lecture exhaustive du graphe afin de vérifier, pour chaque propriété, si elle est, ou non, valide. L'outil étudie chaque lien dans ce but : par exemple, pour vérifier la sérialité, il sélectionne successivement chaque lien afin d'extraire son monde source puis compare l'ensemble ainsi construit à l'ensemble des mondes du cadre. La seconde approche repose sur les propriétés du modèle liées à la relation d'accessibilité : par exemple, comme démontré dans le lemme 1 du chapitre 3, p. 48, pour une relation euclidienne, tous les successeurs d'un monde donné ont le même ensemble de mondes accessibles. Ainsi, à l'aide d'un opérateur d'égalité ensembliste du langage Python, l'outil peut aisément effectuer cette comparaison pour tous les mondes.

## 8.3 Vérification de formules

Après avoir généré un modèle de Kripke, il est possible de vérifier si une formule est vraie dans un monde, pour tous les types de formules : les formules classiques, modales et modales pondérées. La formule à vérifier et le monde de référence sont à entrer dans les champs correspondants.

Cet outil comporte deux étapes : la réécriture d'une formule syntaxiquement correcte sous forme d'arbre binaire puis le traitement de l'arbre pour évaluer la formule.

**Syntaxe des formules** Le logiciel implémenté considère une notation infixe des formules utilisant le langage décrit ci-dessous.

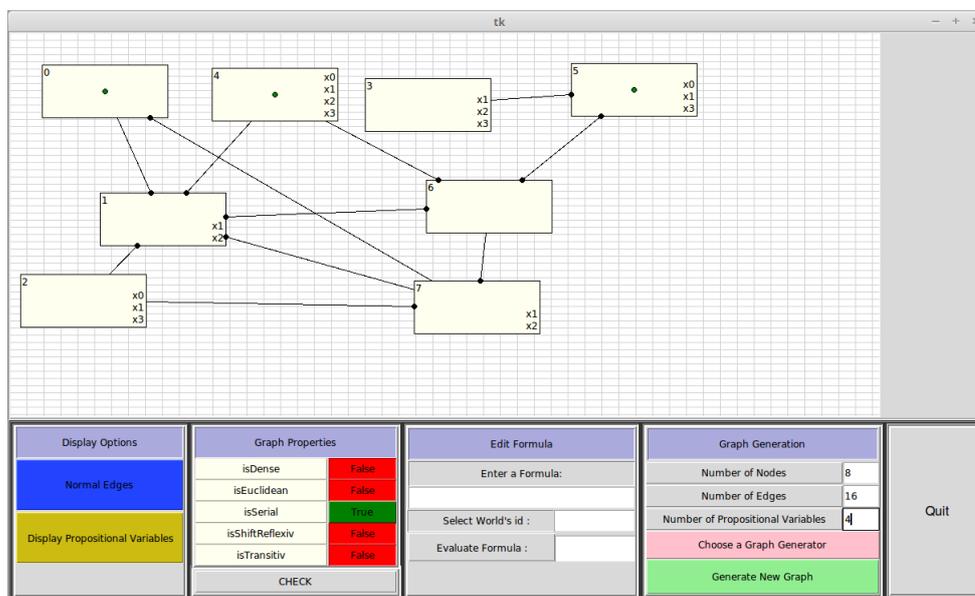


FIGURE 8.4 – Génération d'un modèle aléatoire à 8 mondes, 16 liens et 4 variables.

$F_1$ and $F_2$	pour	$F_1 \wedge F_2$
$F_1$ or $F_2$	pour	$F_1 \vee F_2$
$F_1$ imp $F_2$	pour	$F_1 \rightarrow F_2$
not $F_1$	pour	$\neg F_1$
$[\alpha] F_1$	pour	$\Box_\alpha F_1$
$\langle \alpha \rangle F_1$	pour	$\Diamond_\alpha F_1$

Les formules atomiques sont des variables propositionnelle, notées  $x_0$  à  $x_n$  si le nombre de variables indiquées dans le champ correspondant (cf. section précédente) est  $n+1$ . Ainsi, la formule  $\Box_{0.2}x_1 \wedge \neg\Diamond_{0.4}x_0$  s'écrit :  $[0.2]x_1$ andnot  $\langle 0.4 \rangle x_0$ .

**Conversion en arbre binaire** Lorsqu'une formule syntaxiquement correcte est entrée dans le champ prévu, elle est transmise sous forme infixe à une fonction qui la convertit en arbre binaire. Si l'opérateur est unaire, il est considéré comme un opérateur binaire ayant un argument de valeur **None**, permettant au nœud correspondant dans l'arbre de n'avoir qu'un seul fils. Chaque variable est également considérée comme un opérateur binaire, ayant deux arguments de valeur **None**, pour pouvoir être traitée comme une feuille dans l'arbre.

Ainsi, la régularité de cette notation infixe, ne considérant que des opérateurs binaires dont l'un des arguments, ou les deux, peut être de valeur nulle, permet une conversion rapide en arbre binaire.

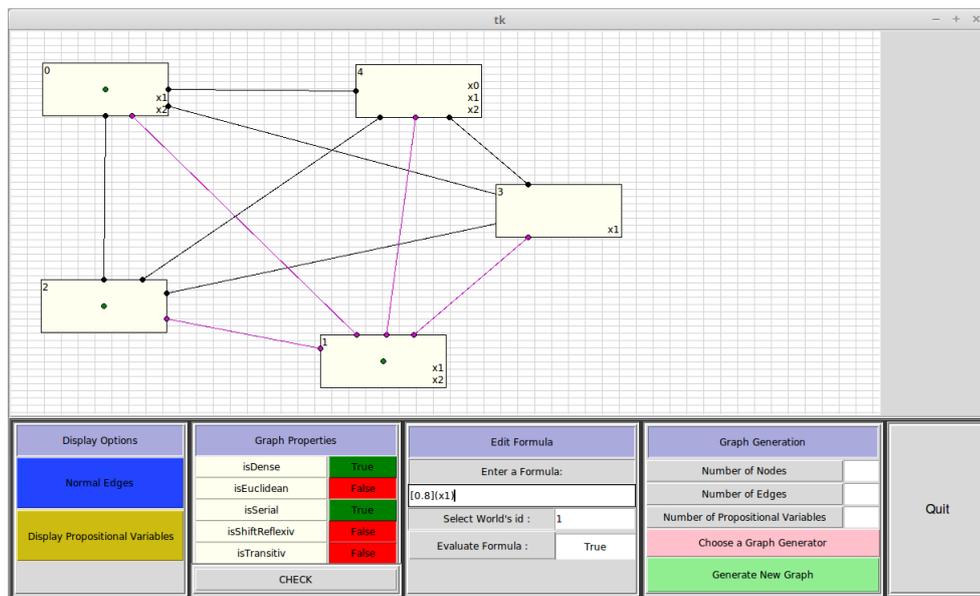


FIGURE 8.5 – Vérification d’une formule modale pondérée dans le monde numéro 1.

**Traitement de l’arbre binaire** À partir de cet arbre, l’évaluation de la formule est opérée dans l’ordre de lecture usuel et, pour chaque opérateur, une vérification est effectuée pour le monde de référence. Par conséquent, deux cas sont possibles : si la formule ne contient pas d’opérateur modal, alors elle est évaluée selon la sémantique propositionnelle classique dans le monde de référence. Dans le cas où elle est modale, les opérateurs modaux sont évalués en considérant chaque monde accessible. Dans le cas d’un carré, ils sont tous évalués, tandis que l’algorithme s’arrête au premier monde validant trouvé dans le cas du losange.

Enfin, dans le cas d’opérateurs modaux pondérés, les mondes accessibles sont explorés successivement jusqu’à atteindre la proportion donnée en paramètre, auquel cas la valeur **True** est renvoyée. Si la proportion n’est jamais atteinte, alors l’évaluateur renvoie **False**.

**Exemples** La figure 8.5 illustre l’exemple de la formule  $\Box_{0.8}x_1$  évaluée dans le monde 1 : le champ permettant d’entrer la formule comporte  $[0.8](x_1)$  et le monde 1 a été indiqué dans le champ suivant.

Un clic sur le bouton « Evaluate Formula » donne alors **True**. En effet, on a bien 4 mondes accessibles dans lesquels  $x_1$  est vraie parmi 5 mondes accessibles au total, soit une proportion de 0.8.

## 8.4 Bilan

Ce chapitre a présenté l'implémentation de la sémantique des formules modales pondérées que nous avons proposée, qui permet de calculer automatiquement la valeur de vérité d'une formule dans un modèle de Kripke donné. De plus, la fonctionnalité de génération automatique de modèles de Kripke contraints permet de réaliser des expérimentations en faisant varier les configurations.

L'outil implémenté ouvre des perspectives ergonomiques, concernant par exemple une disposition qui sépare les outils d'affichage et d'analyse, la mise en valeur des liens d'un monde sélectionné ou l'ajout et la suppression interactifs de mondes ou de liens.

En outre, au-delà de la vérification des formules, l'outil pourrait proposer plusieurs fonctionnalités supplémentaires. D'une part, nous envisageons d'y intégrer une fonctionnalité de calcul de poids maximal : pour un modèle et une formule donnée, en examinant les mondes de manière exhaustive, le logiciel pourrait calculer la proportion de mondes validant pour chaque monde et ainsi donner le poids maximum pouvant être attribué à une modalité de manière à ce que la formule modale pondérée soit vraie dans chacun des mondes.

D'autre part, à plus long terme, nous souhaitons ajouter une fonctionnalité à cet outil afin de proposer un modèle ne validant pas une formule donnée. De cette manière, le logiciel contiendrait un outil de génération de contre-exemple, pour un degré donné, afin d'analyser la validité de certaines formules comme des axiomes modaux pondérés candidats.

## Chapitre 9

# De l'utilisation d'une modalité pondérée dans le cadre d'un jeu coopératif

Dans la lignée des travaux sémantiques, et afin de réaliser une application du raisonnement sur des croyances graduelles, nous proposons une autre application : il s'agit d'utiliser une forme de raisonnement doxastique dans le cadre d'un jeu coopératif entre agents artificiels.

Nous avons co-encadré quatre étudiants dans le cadre d'un projet logiciel de M1 Androïde UPMC, qui a porté sur l'implémentation d'une intelligence artificielle capable de jouer à un jeu coopératif à information incomplète, qui met en œuvre des annonces publiques et du raisonnement sur les croyances.

Le jeu, nommé Hanabi, est d'abord présenté en section 9.1. Nous détaillons ensuite les différentes fonctionnalités, l'interface graphique du logiciel développée par les étudiants étant présentée dans l'annexe A.

La section 9.2 présente ensuite le modèle théorique sur lequel se base le logiciel pour procéder au raisonnement sur les connaissances partielles des joueurs. Ce modèle repose sur la sémantique de Kripke et met en œuvre la sémantique des logiques modales pondérées que nous proposons pour représenter les configurations possibles du jeu. Ainsi, nous procédons à une forme de raisonnement modal sur les informations que le joueur acquiert par les annonces publiques. La section 9.3 détaille le processus de prise de décision basé sur le modèle adapté à un cas pondéré.

Nous proposons également une étude expérimentale, décrite dans la section 9.4 dans laquelle nous montrons que l'utilisation d'une extension pondérée pour le raisonnement modal permet d'améliorer les performances des joueurs par rapport à la version classique.

## 9.1 Présentation du jeu Hanabi

Hanabi est un jeu de cartes coopératif à annonces, dans lequel les joueurs collaborent en échangeant des informations sur leurs mains respectives afin de construire collectivement un résultat commun.

**Matériel** Hanabi comporte 30 cartes, chacune portant une valeur numérique (un entier de 1 à 5) et une couleur (bleu, jaune, rouge, vert, blanc et arc-en-ciel).

La répartition par nombre, pour chaque couleur, est la suivante : trois cartes numérotées par un 1, un 2 ou un 3, deux cartes numérotées 4 et une seule carte numérotée 5.

**Objectif** Le but de la partie est que les joueurs construisent ensemble des séries ordonnées de cartes de même couleur, en les posant successivement sur la table de jeu.

**Déroulement** Au début de la partie, chaque joueur, selon leur nombre, reçoit quatre ou cinq cartes, les cartes restantes constituent la pioche. Un joueur ne peut pas regarder ses propres cartes, en revanche, il voit les cartes des autres joueurs ainsi que, au cours du jeu, les cartes qui ont été posées sur la table.

À chaque tour du jeu, un joueur peut :

- faire une annonce, comme décrit ci-dessous,
- sélectionner une carte de sa main pour la défausser, ce qui écarte définitivement de la partie cette carte
- sélectionner une carte de sa main pour la poser sur la table.

Dans le cas où le joueur choisit de jouer une carte, celle-ci doit impérativement, pour une couleur donnée : soit continuer la suite de chiffres de la couleur correspondante, soit commencer une nouvelle suite de chiffres et donc comporter la valeur 1. Dans le cas contraire, les joueurs reçoivent un jeton rouge en guise de sanction, ils n'ont droit qu'à deux erreurs.

Après avoir joué ou écarté une carte, les joueurs doivent piocher afin de toujours avoir quatre, ou cinq, cartes dans leur main.

La partie se termine de deux façons possibles : soit les joueurs reçoivent collectivement un troisième jeton rouge et ils perdent la partie, soit la pioche est vide et ils terminent le tour en cours avant de compter leur score final. Celui-ci est égal au nombre de cartes posées sur la table.

**Annonces** Un joueur peut, à son tour de jeu, choisir de faire une annonce à une coéquipier, deux formes sont possibles : s'il choisit de fournir une information concernant la couleur, il doit indiquer toutes les cartes de la main du joueur visé par l'annonce qui sont de cette couleur. Ainsi, il peut dire « ta première et ta troisième carte sont jaunes ».

La second type d'annonce est similaire mais porte sur la valeur numérique : il peut prendre la forme « ta première carte est le seul 1 » .

Le nombre d'annonces est également géré par un système à base de jetons : une annonce n'est possible qu'à condition d'utiliser un jeton bleu, initialement au nombre de huit. Les jetons consommés peuvent être récupérés en écartant une carte du jeu, celle-ci est alors retirée de la partie et ne pourra plus être utilisée pour construire une suite de cartes.

Dans le cadre de leur projet, les étudiants ont développé une interface graphique présentée en annexe, p. 141.

## 9.2 Formalisation en logique modale et logique modale pondérée

Hanabi est un jeu à information incomplète puisqu'un joueur ne connaît pas sa propre main. Il se prête donc bien à une formalisation en logique modale (pondérée ou non) qui permet de représenter ses croyances sur les cartes dont il dispose, qui évoluent selon les observations qu'il fait des mains des autres joueurs ainsi que les annonces dont il fait l'objet.

Nous présentons ici le choix effectué pour la formalisation : le cadre de Kripke considéré la mise à jour des mondes possibles selon l'évolution du jeu. La section suivante décrit la prise de décision et les stratégies de jeu.

**Modèles de Kripke** Chacune des quatre (ou cinq) cartes de la main du joueur est représentée par un modèle de Kripke, dans lequel un monde possible représente une valeur possible de la carte, définie par le couple (couleur, nombre). Toutes les cartes possibles sont représentées dans chaque modèle. La relation entre les mondes est une relation d'indistinguabilité, connectant deux cartes si elles ont une couleur ou un nombre en commun, puis par transitivité tous les mondes sont en relation en début de partie donc le graphe est complet : la figure 9.1 représente ce modèle mais, par souci de lisibilité, elle ne comporte pas toutes les relations. Cette représentation est générée de manière exhaustive au début de partie mais des configurations sont rapidement éliminées en fonction des cartes apparaissant dans les mains des autres joueurs.

Prenons l'exemple suivant : en considérant le modèle qui représente la première carte d'un joueur, on a initialement l'ensemble de mondes représenté sur la figure 9.1. Chaque monde comporte une carte possible, celle-ci pouvant être l'un des trois 1, deux 2, 3 et 4, ou le 5, de chaque couleur. Le nombre de mondes possibles est égal au nombre de valeurs possibles multiplié par le nombre de couleurs :  $(3 + 2 + 2 + 2 + 1) \times 5 = 60$ .

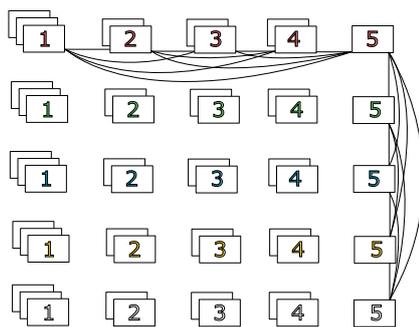


FIGURE 9.1 – Modèle de Kripke représentant l'une des cartes d'un joueur au début de la partie : chaque rectangle représente un monde possible, la relation d'accessibilité, qui relie tous les couples de mondes, n'est pas représentée intégralement

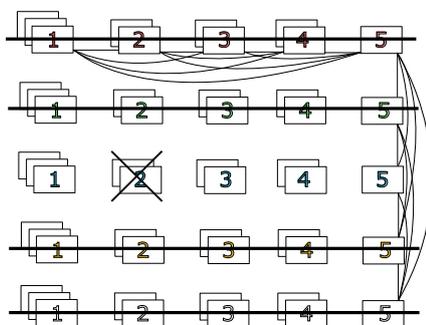


FIGURE 9.2 – Représentation des valeurs possibles après annonce attribuant la couleur bleue à la carte considérée

**Mise à jour d'un modèle** Un joueur met à jour les modèles de Kripke représentant ses cartes selon les annonces dont elles font l'objet et selon l'observation des cartes des autres joueurs. En effet, chaque annonce qui est faite à un joueur lui permet d'éliminer des mondes : par exemple, dans le cas de la carte associée au modèle donné sur la figure 9.1, si l'on annonce au joueur qu'elle est de couleur bleue, alors il élimine tous les mondes dans lesquels celle-ci n'est pas bleue, laissant pour seuls mondes possibles ceux représentés sur la troisième ligne. Notons qu'il peut également tirer des informations sur les cartes non-désignées : s'il lui est indiqué qu'une carte est rouge, alors il est garanti que ses autres cartes sont d'une autre couleur, permettant d'éliminer les mondes correspondants pour celles-ci.

D'autre part, le contenu des mains des autres joueurs ainsi que les cartes écartées ou déjà jouées apportent des indices implicites sur l'état de la main du joueur : par exemple, si deux cartes comportant un 2 bleu apparaissent dans les mains d'autres joueurs, il peut éliminer les mondes représentant cette carte de son modèle de Kripke.

La figure 9.2 illustre par exemple le modèle de Kripke obtenu en combinant cette information avec l'annonce précédente selon laquelle la carte est bleue.

**Interprétation doxastique pondérée** L'état doxastique des joueurs décrit l'ensemble de leurs connaissances concernant leur main : ils peuvent ainsi exprimer des formules du type « je crois que ma première carte est un deux rouge ». Dans le cadre pondéré, leurs croyances sont plus riches et comparables : « je crois à un degré 0.4 que ma carte  $C1$  est un deux rouge et je crois à un degré 0.7 que ma carte  $C2$  est un quatre bleu ».

Formellement, on peut considérer que chaque variable propositionnelle décrit une hypothèse quant à la valeur de chaque carte : par exemple, «  $C1$  est un 1 rouge », «  $C1$  est un 2 rouge », «  $C2$  est un 1 jaune », «  $C3$  est un 5 bleu », etc. Ainsi, la formule doxastique « je crois à un degré 0.6 que ma première carte est un 1 rouge » est écrite  $B_{0.6}C_{rouge}^1$ .

Formellement, la formule  $B_{\alpha}C_c^i$  modélise le fait que l'agent croit à un degré  $\alpha$  que la carte a le numéro  $i$  et est de couleur  $c$ . Le calcul du degré de croyance  $\alpha$  correspond au rapport entre le nombre de mondes accessibles validant  $C_c^i$  et le nombre total de mondes accessibles. Comme le graphe est complet, cela correspond au nombre de mondes du modèle validant  $C_c^i$  sur le nombre de mondes possibles total.

De même, le calcul du degré de croyance en la jouabilité d'une carte correspond au nombre de mondes portant la valeur d'une carte jouable sur le nombre total de mondes possibles. Ainsi la formule  $B_{\alpha}jouable(C)$  représente l'assertion « je crois que ma carte  $C$  est jouable à un degré  $\alpha$  ».

### 9.3 Prise de décision en logique modale

Cette proposition, dans laquelle on a un modèle par carte, permet au joueur de maximiser ses connaissances concernant sa propre main. Il peut alors jouer une carte si ses doutes concernant sa valeur sont suffisamment faibles : nous établissons la tolérance des joueurs vis-à-vis du risque en fixant une valeur de seuil en début de partie. Dans le cas contraire, il a toujours la possibilité de donner un indice à un autre joueur ou d'écarter une carte.

Pour déterminer si une carte est jouable, la fonction de décision associée considère tous les emplacements possibles parmi les cartes déjà jouées : la valeur sommet de chaque suite de couleur détermine la valeur numérique à poser. Par exemple, si pour la couleur rouge, il y a déjà un 1, un 2 et un 3, alors le 4 rouge est une carte requise. Si une couleur n'apparaît pas encore, alors la carte requise est un 1 de cette couleur. Ainsi, la fonction de décision détermine si une carte est jouable si elle correspond à une carte requise.

À tout moment, l'agent peut déterminer son degré de croyance en la jouabilité d'une carte, modélisé formellement par  $B_{\alpha}jouable(C)$ . Il décide alors de jouer une carte si le

degré de croyance associé à sa jouabilité est supérieur au seuil de tolérance au risque fixé en début de partie.

Ainsi, pour un risque de valeur  $x \in [0,1]$ , alors la carte maximisant  $\alpha$  est jouée à condition que  $\alpha \geq 1-x$ . Par exemple, pour un risque de 25%, la carte jouée doit obligatoirement avoir plus de 75% de ses valeurs possibles qui sont jouables.

Si aucune carte jouable ne peut être jouée sans dépasser le seuil de risque, alors le joueur écarte une carte dont il sait qu'elle n'est plus nécessaire à la partie. Si ce n'est pas possible, il donne un indice à un autre joueur, en privilégiant les indices donnant une information sur un maximum de cartes. Si aucun jeton d'indice n'est disponible, le joueur écarte une carte au hasard.

## 9.4 Expérimentation

Pour étudier l'influence de l'assouplissement du quantificateur des mondes devant comporter une carte jouable, nous avons simulé 1000 parties entre 2, 3, 4 et 5 joueurs épistémiques ayant tous le même seuil de risque. Nous proposons d'étudier les conséquences de l'utilisation d'une croyance pondérée dans la prise de décision sur deux éléments : le score moyen par partie et le nombre d'erreurs commises, représentées par les jetons rouges obtenus par les joueurs.

Les graphiques de la figure 9.3 représentent l'évolution du score moyen, et son écart-type, en fonction du seuil de risque des joueurs et utilisant les cartes multicolores.

Plusieurs observations sont à souligner : tout d'abord, la tendance de la courbe est identique pour toutes les figures : d'une part, la valeur moyenne est maximisée autour de 0.6 bien que celle-ci ne varie pas de manière particulièrement notable.

L'écart-type, resté quasiment constant, ne donne pas d'indication particulière quant à une amélioration significative du score dans le cas d'un risque modulé : il est même légèrement plus élevé lorsque la moyenne est, elle aussi, en augmentation. Néanmoins, la marge de manœuvre est limitée, car les scores possibles sont dans un ordre de grandeur limité.

Ainsi, la faible variabilité des scores s'explique par le fait qu'une différence d'un point correspond à une carte supplémentaire correctement posée, ce qui, dans le cadre d'une partie à 2, 3 ou même 4 joueurs, nécessite un effort collectif considérable. Dans le cas d'une partie à 5 joueurs, la moyenne augmente de deux points, notamment car l'opportunité de poser une carte augmente.

Enfin, les histogrammes de la figure 9.4 décrivent l'évolution du nombre de jetons rouges en fonction du seuil de risque, pour chaque valeur comprise entre 0.0 et 1.0, en simulant 1000 parties.

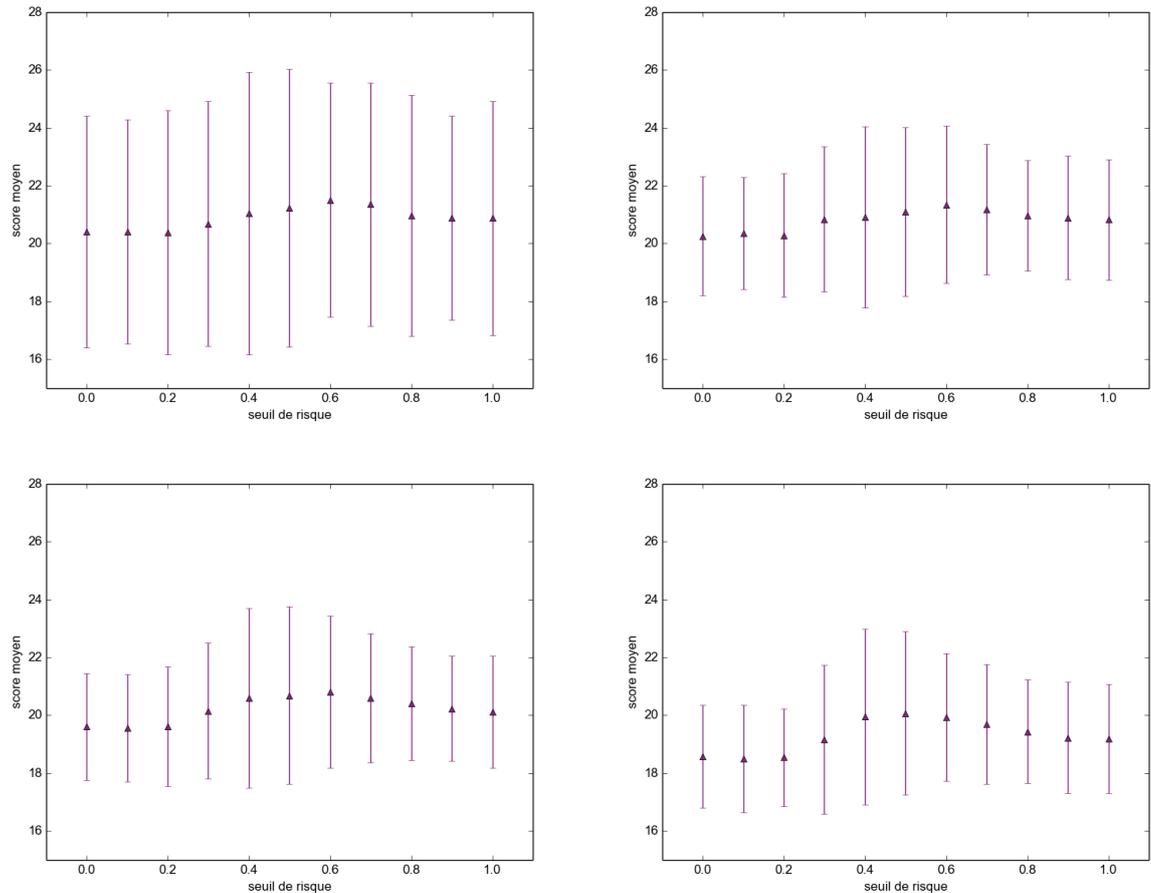


FIGURE 9.3 – Score moyen et son écart-type en fonction du seuil de risque, pour une partie à deux joueurs en haut à gauche, trois joueurs en haut à droite, quatre joueurs en bas à gauche et cinq joueurs en bas à droite

On peut observer deux tendances dans l'évolution des résultats : dans les cas à 2 et 3 joueurs, pour un seuil trop faible, toutes les parties finissent à deux jetons rouges tandis que le nombre de parties à 1 ou 0 jeton augmente progressivement jusqu'au seuil maximal où les parties à 0 jeton sont majoritaires.

Dans les cas de parties à 4 et 5 joueurs, le résultat est identique pour un seuil trop faible mais le nombre de parties à 2 jetons ne diminue pas drastiquement au fil de l'augmentation du seuil. On constate cependant une amélioration du nombre d'erreurs pour les seuils les plus élevés : le nombre cumulé de parties à 0 et 1 jetons devient tout de même majoritaire. Nous ajoutons que les parties à 0 jeton deviennent également moins fréquentes que les parties à 1 jeton : ce phénomène est accentué pour les parties à 5 joueurs.

Bien que contre-intuitive, cette tendance peut s'expliquer par la stratégie de jeu des joueurs : à seuil de risque très faible, les joueurs ont tendance à favoriser le don d'indices

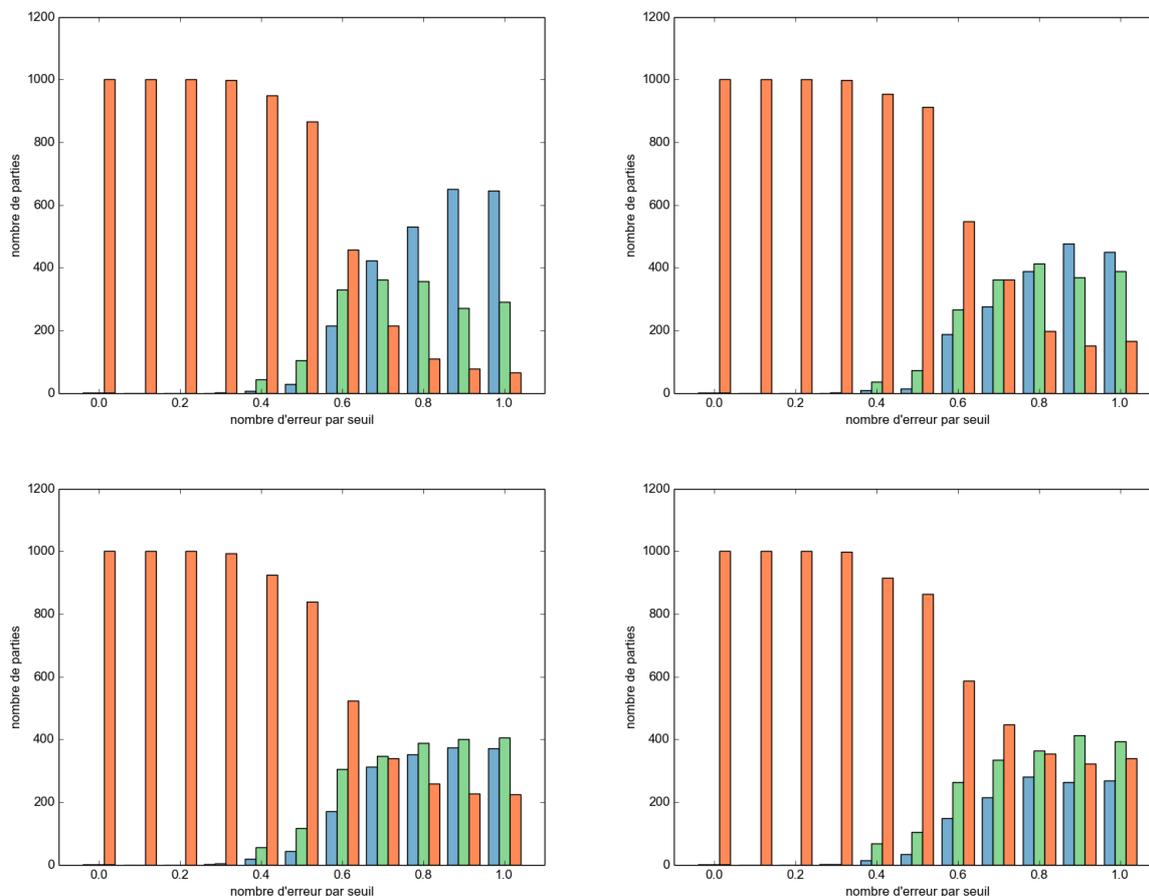


FIGURE 9.4 – Nombre de jetons rouges pour 1000 parties : 0 en bleu, 1 en vert et 2 en rouge ; de gauche à droite et de haut en bas : pour 2, 3, 4 et 5 joueurs.

mais risquent de se retrouver dans une impasse à devoir jouer au hasard par manque de jetons bleus. Cela s'inverse avec une prise en compte d'un plus grand risque : les joueurs osent jouer un carte malgré un certain risque, et ainsi évitent de poser une carte au hasard. On constate toutefois qu'à nombre de joueurs élevé (4 ou 5), le risque pris par tous les joueurs réduit l'effet d'augmentation du seuil en augmentant les chances de poser une mauvaise carte.

Cependant, il est nécessaire de vérifier cette hypothèse par plus d'expérience : notamment en faisant varier le seuil de risque au cours de la partie (élevé en début de partie mais qui diminue ensuite, par exemple), ou encore en faisant varier le seuil des joueurs afin de faire coopérer des joueurs averses au risque avec des joueurs qui en prennent plus.

## 9.5 Bilan

Nous avons mis en œuvre les travaux sémantiques pour les logiques modales pondérées, dans le cadre du jeu Hanabi. Ce jeu est coopératif à information incomplète et fonctionne sur la base d’annonces publiques.

Le logiciel implémenté propose, en plus d’une interface graphique, une intelligence artificielle capable de jouer de manière performante à Hanabi. Ce logiciel est complet et permet à l’utilisateur de jouer avec des joueurs artificiels convaincants. En effet, leur attitude vis-à-vis du risque et leur modèle calculatoire pour déterminer les cartes à jouer sont efficaces et crédibles. Les performances de l’intelligence artificielle sous-jacente permettent d’imiter des parties entre joueurs humains en reproduisant les scores obtenus dans ces cas.

L’intelligence artificielle implémentée repose sur une représentation modale des configurations du jeu, permettant au joueur artificiel d’établir des degrés de croyance sur les cartes de sa main. Sa stratégie de jeu tient compte de cette croyance graduelle sur la jouabilité des cartes. Enfin, ses prises de décision quant aux cartes à sélectionner pour les jouer sont basées sur un calcul de proportion de mondes validant, ce qui permet une mise en pratique des travaux sémantiques présentés en partie 1 de ce manuscrit.

Plusieurs pistes d’améliorations sont envisagées afin de permettre au logiciel d’évoluer et de gagner en fonctionnalité et ergonomie.

Tout d’abord, en terme d’optimisation du score, certaines améliorations sont possibles. En particulier, comme mentionné ci-dessus, nous envisageons de tester des méthodes heuristiques, pour faire varier le seuil de risque tout au long de la partie. De cette façon, le joueur peut adapter sa stratégie selon l’évolution de la partie.

Nous souhaitons également nous concentrer sur un autre aspect du jeu pour lequel le modèle épistémique peut être amélioré : les connaissances de tous les joueurs pourraient être mutualisées afin d’augmenter leur connaissance générale sur l’état du jeu. En d’autres termes, chaque joueur pourrait faire des hypothèses sur l’état épistémique de ses coéquipiers et les utiliser afin de faire des déductions sur les valeurs de ses propres cartes. Un tel modèle se base sur une extension doxastique de la logique épistémique multi-agents, qui permet de représenter et utiliser des connaissances communes et partagées pour le raisonnement des agents.

Il est ainsi possible de se pencher sur la notion de raisonnement méta-épistémique. L’objectif d’une telle extension du joueur artificiel est de mettre en œuvre ce type de raisonnement et d’enrichir les capacités de ces joueurs à anticiper les coups de leurs coéquipiers.

Dans la continuité de ces travaux, nous souhaitons également nous intéresser à d’autres jeux du même type mais avec des variantes, par exemple, en proposant la prise en compte

d'annonces publiques mensongères. C'est le cas d'un grand nombre de jeux non-coopératifs basés sur le bluff. Pour mettre en œuvre les travaux en logiques modales pondérées, nous souhaitons nous placer dans le cadre de la gestion d'annonces mensongères. Il s'agirait de déterminer si le formalisme modal pondéré se prête particulièrement à ce type de problématiques.

# Conclusion et perspectives

Cette conclusion présente les différentes contributions de cette thèse ainsi que les perspectives qu'elle ouvre.

## Contributions

Si la question considérée initialement porte sur le raisonnement sur des croyances graduelles, plusieurs problématiques émergent : en nous basant sur les logiques modales comme formalisme logique pour les croyances, et en considérant les degrés de croyance comme objets à manipuler pour le raisonnement sur les croyances graduelles, deux thématiques se dessinent.

D'une part, se pose la question des conséquences de la pondération sur le formalisme modal, que nous avons étudiées en examinant l'extension de la logique modale classique au cas pondéré.

D'autre part, nous avons proposé un formalisme logique pour représenter et manipuler des croyances graduelles, basé sur une interprétation floue de leurs degrés.

## Logiques modales pondérées

En logiques modales pondérées, nous avons tout d'abord établi une étude comparative et organisée des modèles existants. En regard de cette étude, nous avons proposé une sémantique proportionnelle pour interpréter les formules modales, basée sur un comptage relatif des mondes accessibles validants qui permet de lier la valeur des poids à la connexité du monde considéré. Celle-ci possède en particulier une propriété essentielle : la gradualité décroissante, laquelle permet de déduire à partir d'une formule modale pondérée  $\Box_\alpha\varphi$ , toute formule modale pondérée  $\Box_\beta\varphi$  pour tout  $\beta$  inférieur à  $\alpha$ .

En conséquence de cette sémantique et de sa propriété de gradualité décroissante, nous avons discuté et établi ce qui est attendu, en terme d'objectif de maximisation et minimisation des poids dans une implication. Ces attentes posent une base pour l'étude détaillée des conséquences de cette sémantique sur les axiomes : il s'agit de substituer les modalités classiques des axiomes modaux par des modalités pondérées et d'ébaucher une extension

de la théorie de la correspondance. En effet, si chaque axiome est classiquement mis en correspondance avec une propriété particulière de la relation d'accessibilité des modèles, il est indispensable de tenir compte de la sémantique proposée dans cette correspondance car elle modifie nécessairement le lien entre axiome et modèle. Ces travaux nous ont permis d'établir une typologie des axiomes pondérés selon l'enrichissement qu'ils permettent vis-à-vis de leur version classique et de leur correspondance sémantique.

### **Représentation et manipulation ensembliste des croyances graduelles**

Dans le système de raisonnement sur des croyances graduelle, que nous avons établi, nous avons proposé un modèle théorique pour les représenter et les manipuler. Il repose sur des travaux épistémologiques qui permettent de définir le concept de croyance graduelle et d'en définir les principes attendus : nous avons proposé un modèle ensembliste flou comme structure arithmétique pour les degrés de croyance.

De cette façon, les croyances d'un agent sont définies comme les éléments de son ensemble de croyance et dont le degré correspond au degré d'appartenance à cet ensemble. Le degré est défini selon le formalisme de la théorie des sous-ensembles flous, qui fournit des règles permettant de combiner les degrés. Nous avons étudié les conséquences et propriétés sous-jacentes, qui découlent de cette interprétation floue des degrés, pour les opérateurs logiques de conjonction et disjonction, ainsi que pour les axiomes doxastiques  $K$ , qui distribue l'implication sur la modalité, et d'introspection. Nous avons de plus détaillé l'étude de la négation des croyances graduelles.

L'ensemble des règles de manipulation proposées constitue un système d'axiomes adaptés à l'interprétation floue des degrés de croyance. À partir de celui-ci, nous pouvons ébaucher un modèle pour le raisonnement en logique doxastique graduelle.

### **Mises en œuvre**

Deux applications ont été proposées, pour faire suite à ces travaux théoriques. D'une part, nous avons proposé un outil de génération de modèles et vérification de formules modales pondérées : le logiciel implémenté met en œuvre la sémantique proportionnelle et permet, après avoir généré et affiché un modèle, d'entrer une formule modale pondérée au clavier pour en connaître la véracité dans ce modèle.

D'autre part, nous avons mis en pratique la sémantique proportionnelle que nous avons proposée dans le cadre d'un jeu de cartes coopératif à information incomplète et dont la résolution est basée sur des annonces publiques. Le jeu est modélisé selon une sémantique modale : la main des joueurs est représentée par des modèles de Kripke dont les mondes comportent les valeurs d'une carte. Le joueur artificiel implémenté possède un modèle de

raisonnement basé sur un seuil de risque lorsqu'il s'agit de jouer une carte.

Ces deux applications constituent un premier résultat de mise en pratique des modèles logiques proposés. Nous nous sommes basé en particulier sur la seconde pour proposer une étude expérimentale des conséquences de la pondération des modalités sur un cas pratique de raisonnement sur des croyances graduelles.

## Perspectives

Outre les perspectives spécifiques mentionnées dans les conclusions de chapitre, les contributions de cette thèse ouvrent un ensemble de perspectives détaillées ci-dessous et organisées en trois thématiques : la première porte sur une validation expérimentale de nos propositions sémantiques, la seconde regroupe des perspectives logiques et la troisième concerne l'exploitation de nos propositions pour une mise en œuvre dans d'autres domaines.

### Validation expérimentale

Dans le cadre de la thèse, nous avons mis en œuvre le modèle proposé pour raisonner sur des croyances graduelles en l'appliquant au jeu Hanabi, afin de réaliser une étude expérimentale de nos propositions sémantiques de logique doxastique pondérée. Ce jeu a la particularité de conduire à un cadre de Kripke spécifique dans lequel le graphe des mondes accessibles est complet.

Nous souhaitons réaliser d'autres expérimentations de la logique doxastique pondérée que nous avons proposée et la mettre en œuvre dans d'autres applications afin de considérer des cadres de Kripke plus généraux : l'objectif est d'étendre, et de valider, les interprétations de l'axiomatique et son pouvoir d'inférence, en le testant sur des données réelles.

Il pourra, par exemple, s'agir d'établir des modèles de Kripke à partir de la formalisation logique d'assertions exprimées dans un corpus de documents ; de type articles de journaux. Ainsi, un modèle peut représenter les différentes versions d'un même sujet, reliés selon une relation d'indistinguabilité. La validation du modèle sémantique va notamment nécessiter d'interpréter les propriétés, ainsi qu'une lecture pratique des axiomes, en tenant compte du contexte d'application.

### Perspectives logiques

**Correction et complétude** Nous avons comparé de manière informelle la définition du système  $(KD45)_\alpha$  issu de la sémantique proportionnelle et les expressions des axiomes basés sur le formalisme flou : l'établissement d'une preuve montrant que tout ce qui est valide dans la sémantique est prouvable dans le système de preuves établis à partir de la représentation ensembliste, et réciproquement, est une étape indispensable pour proposer

un système abouti pour raisonner sur des croyances graduelles et basé sur une logique modale pondérée.

Comme dans la preuve classiquement construite pour la logique modale (Fagin et al., 2004), nous devons d’abord définir formellement la conséquence sémantique  $\models$  ainsi que les tautologies impliquant une validation sémantique pour les modèles. Il nous faudra également définir précisément l’ensemble des axiomes considérés pour la preuve, ainsi que les règles d’inférence qui, associées aux axiomes, forment un système de preuve. La dernière étape consistera à montrer l’équivalence entre les modèles et le système de preuve.

**Étude de systèmes modaux pondérés** Dans l’étude typologique des axiomes pondérés que nous avons réalisée, nous avons considéré les axiomes individuellement, indépendamment les uns des autres. Une perspective vise à étudier leur prise en compte simultanée et les conséquences qu’elle provoque sur les versions pondérées des systèmes d’axiomes modaux classiques, définis par des sous-ensembles d’axiomes : la correspondance classique étant modifiée, les systèmes modaux le sont également

On peut en particulier s’intéresser à l’extension graduelle de systèmes connus comme par exemple KD45, S4 ou S5 : en tenant compte du fait que l’axiome pondéré  $(4)_\alpha$  nécessite une relation euclidienne ou que tout système modal pondéré contient d’office l’axiome  $(CD)_\alpha$ , comme nous l’avons montré dans le chapitre 3, une nouvelle problématique émerge, liée à la composition de systèmes modaux pondérés et à leurs particularités.

**Méthode des tableaux** Un autre perspective logique vise à proposer et implémenter une extension pondérée des outils d’inférence modale, en particulier pour la méthode des tableaux : la question est celle de la définition d’un algorithme de construction itérative de structure de Kripke lors de la décomposition récursive des formules et de la réécriture des modalités pondérées. Il s’agit de généraliser les approches proposées (voir Goré (2013) pour une présentation détaillée) pour la gestion de contraintes sur les proportions de mondes accessibles.

D’un point de vue pratique, une implémentation des règles tableaux qui seront définies, dans des outils tels que LoTREC (Gasquet et al., 2005), est envisagée.

### Exploitations, mises en œuvre et généralisations

**Jeux à annonces publiques mensongères** Dans la suite de l’implémentation proposée pour le jeu Hanabi, nous souhaitons étendre l’exploitation du modèle sémantique proposé pour résoudre des jeux à annonces publiques. En particulier, nous nous intéressons au cas d’annonces mensongères, pour lequel il faut gérer cette falsification de la vérité

---

de la part des joueurs. Nous envisageons de baser une modélisation de ces annonces publiques mensongères par un système modal pondéré : ainsi, chaque annonce peut pousser un agent à moduler sa croyance en tenant compte des différentes définitions possibles du mensonge (Capet, 2012). Cette utilisation de croyances graduelles est suffisamment souple pour aboutir à un modèle pertinent. Enfin, cela pourrait nous permettre de faire le lien entre les formalisations possibles du mensonge et de ses nuances (voir Capet (2012) et Sakama et al. (2014) pour des approches philosophiques et modales du mensonge) et les croyances graduelles.

**Confiance et cotation d'information** Nous souhaitons également nous intéresser au cadre concret de l'évaluation de la qualité d'information : avec l'émergence de multiples sources d'information, notamment liée à l'évolution des supports de diffusion et la démocratisation du web, il devient indispensable de traiter les informations afin d'en déterminer la qualité intrinsèque.

La cotation d'information (Capet & Revault d'Allonnes, 2014) est une tâche de mesure de la qualité d'une information et, en particulier, la confiance qu'on peut lui accorder. Elle est importante, car elle permet de valoriser une information objectivement meilleure, et délicate en raison de la complexité du problème soulevé. Cette dernière peut être exprimée comme la question de savoir à quel point on peut croire l'information. Ainsi, la croyance peut être prise en compte pour évaluer la qualité d'une information. La confiance peut également être considérée comme une notion combinant plusieurs couches de croyances basiques (Falcone et al., 2002).

La gradualisation des croyances permettrait d'établir un modèle riche et expressif pour valuer la qualité d'une information ou en estimer la confiance accordée. En utilisant les travaux présentés dans ce manuscrit, il est possible d'inscrire le modèle pour les croyances graduelles au cas de la cotation d'une information ou de l'évaluation de la confiance.



# Bibliographie

- Armstrong, D. M. (1968). *A materialist theory of the mind*. Routledge.
- Askounis, D., Koutras, C. D., & Zikos, Y. (2012). Knowledge means ‘all’, belief means ‘most’. In *Logics in artificial intelligence*, 41–53. Springer.
- Bacchus, F., Grove, A. J., Halpern, J. Y., & Koller, D. (1996). From statistical knowledge bases to degrees of belief. *Artificial Intelligence*, 87, 75–143.
- Banerjee, M., & Dubois, D. (2009). A simple modal logic for reasoning about revealed beliefs. In *Proc. of european conf. on symbolic and quantitative approaches to reasoning with uncertainty, ecsqaru*, 805–816. Springer.
- Blackburn, P., De Rijke, M., & Venema, Y. (2001). *Modal logic*. Cambridge University Press.
- Bou, F., Esteva, F., Godo, L., & Rodriguez, R. (2009). Characterizing fuzzy modal semantics by fuzzy multimodal systems with crisp accessibility relations. *Proc. of the Joint 2009 IFSA World Congress and EUSFLAT Conference* (pp. 1541–1546).
- Bouchon-Meunier, B. (1993). *La logique floue*, vol. 2702. Que sais-je.
- Boutilier, C. (1994). Modal logics for qualitative possibility theory. *International Journal of Approximate Reasoning*, 10, 173–201.
- Capet, P. (2012). Qu’est-ce que mentir ?
- Capet, P., & Revault d’Allonnes, A. (2014). Information evaluation in the military domain : doctrines, practices and shortcomings in Information Evaluation (eds P. Capet and T. Delavallade). *Information Evaluation*, 103–127.
- Castelfranchi, C., & Falcone, R. (2000). Trust is much more than subjective probability : Mental components and sources of trust. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences* (pp. 10–pp).
- Chellas, B. F. (1980). *Modal logic : an introduction*. Cambridge University Press.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78, 67–90.

- Davidson, D. (2001). *Inquiries into truth and interpretation : Philosophical essays*, vol. 2. Oxford University Press.
- De Caro, F. (1988). Graded modalities II. *Studia Logica*, 47, 1–10.
- De Finetti, B. (1974). *Theory of probability : A critical introductory treatment*. John Wiley and Sons.
- de Soto, A. R., & Trillas, E. (1999). On antonym and negate in fuzzy logic. *Int. Journal on Intelligent Systems*, 14, 295–303.
- Demolombe, R. (2004). Reasoning about trust : A formal logical framework. In *Trust management*, 291–303. Springer.
- Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, 88, 27–51.
- Detyniecki, M. (2000). *Mathematical aggregation operators and their application to video querying*. Thèse de doctorat, Université Pierre et Marie Curie, Paris.
- Dretske, F. I. (1988). *Explaining behavior : Reasons in a world of causes*. Cambridge University Press.
- Dubois, D., Lang, J., & Prade, H. (1994a). Possibilistic logic 1.
- Dubois, D., Lang, J., & Prade, H. (1994b). Possibilistic logic 1. *Hand-Book of Logic in Artificial Intelligence and Logic Programming*, 3, 439–513.
- Dubois, D., & Prade, H. (2011). Generalized possibilistic logic. *Proc. of the 5th Int. Conf. on Scalable Uncertainty Management, SUM* (pp. 428–432). Springer.
- Dubois, D., Prade, H., & Schockaert, S. (2011). Rules and meta-rules in the framework of possibility theory and possibilistic logic. *Scientia Iranica*, 18, 566–573.
- Dubois, D., Prade, H., & Schockaert, S. (2012). Stable models in generalized possibilistic logic. *13th International Conference on the Principles of Knowledge Representation and Reasoning, KR 2012* (pp. 519–529). AAAI Press.
- Dubois, D., Prade, H., & Schockaert, S. (2014). Reasoning about uncertainty and explicit ignorance in generalized possibilistic logic. *Proc. of European Conference on Artificial Intelligence, ECAI 2014* (pp. 261–266). IOS Press.
- Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. (2004). *Reasoning about knowledge*. MIT press.
- Falcone, R., Pezzulo, G., & Castelfranchi, C. (2002). A fuzzy approach to a belief-based trust computation. *Workshop on Deception, Fraud and Trust in Agent Societies* (pp. 73–86).
- Fariñas del Cerro, L., & Herzig, A. (1991). A modal analysis of possibility theory. *Fundamentals of Artificial Intelligence Research* (pp. 11–18). Springer.

- 
- Fattorosi-Barnaba, M., & De Caro, F. (1985). Graded modalities. I. *Studia Logica*, 44, 197–221.
- Festinger, L. (1962). *A theory of cognitive dissonance*, vol. 2. Stanford University Press.
- Fine, K. (1972). In so many possible worlds. *Notre Dame Journal of Formal Logic*, 13, 516–520.
- Fishbein, M., & Ajzen, I. (1976). *Belief, attitude, intention, and behavior : An introduction to theory and research*. Longman Higher Education.
- Fodor, J. A. (1981). *Representations*. The MIT press.
- Fodor, J. A. (1990). *A theory of content and other essays*. The MIT press.
- Gärdenfors, P. (2003). *Belief revision*, vol. 29. Cambridge University Press.
- Garson, J. (2016). Modal logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Spring 2016 édition.
- Gasquet, O., Herzig, A., Longin, D., & Sahade, M. (2005). Lotrec : Logical tableaux research engineering companion. *TABLEAUX* (pp. 318–322).
- Gentzen, G. (1934). Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 39, 176–210.
- Godo, L., Hájek, P., & Esteve, F. (2003). A fuzzy modal logic for belief functions. *Fundamenta Informaticae*, 57, 127–146.
- Goré, R. (2013). *Handbook of tableau methods*, chapter Tableau methods for modal and temporal logic, 297–396. Springer Science & Business Media.
- Hackl, M. (2009). On the grammar and processing of proportional quantifiers : most versus more than half. *Natural Language Semantics*, 17, 63–98.
- Hähnle, R. (2001). Tableaux and related methods. *Handbook of automated reasoning*, 1, 4.
- Harman, G. (2008). *Change in view : Principles of reasoning*. Cambridge University Press.
- Hintikka, J. (1962). *Knowledge and belief : an introduction to the logic of the two notions*. Cornell University Press Ithaca.
- Huber, F. (2016). Formal representations of belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Spring 2016 édition.
- Keenan, E. (2004). Standard and non-standard quantifiers in natural language. *Proc. of the York conference on Types of Quantification*. Oxford University Press.

- Klement, E. P., Mesiar, R., & Pap, E. (2013). *Triangular norms*. Springer Science & Business Media.
- Kripke, S. (1963). Semantical considerations on modal logic. *Acta philosophica fennica*, 16, 83–94.
- Laverny, N., & Lang, J. (2004). From knowledge-based programs to graded belief-based programs. Part I : On-line reasoning (regular paper). *European Conference on Artificial Intelligence, ECAI 2004* (pp. 368–372). IOS Press.
- Laverny, N., & Lang, J. (2005). From knowledge-based programs to graded belief-based programs, Part II : off-line reasoning. *Proc. of the 9th Int. Joint Conference on Artificial Intelligence, IJCAI 2005* (pp. 497–502). Gallus.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2015a). Aggregation of belief degrees in graded doxastic logic for conjunction and disjunction. *Non-classical logics. Theory and applications*.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2015b). Typologie d’axiomes pour une logique modale pondétée. *Intelligence Artificielle Fondamentale @PFIA2015*.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2015c). Typology of axioms for a weighted modal logic. *Proc IJCAI 2015 Workshop on Weighted Logics for Artificial Intelligence* (pp. 40–47).
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2016a). Belief, semi-belief and unbelief. *Non-classical logics. Theory and applications*.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2016b). Négation de croyances graduelles. *Proc. Logique Floue et ses Applications*.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2016c). Negation of graded beliefs. *Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 16* (pp. 148–160). Springer.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2017a). A fuzzy take on graded beliefs. In *Proc. of the Eusflat 2017 Conference*, 392–404. Springer.
- Legastelois, B., Lesot, M.-J., & Revault d’Allonnes, A. (2017b). Typology of axioms for a weighted modal logic. *International Journal of Approximate Reasoning*.
- Marcus, R. B. (1990). Some revisionary proposals about belief and believing. In *Causality, method, and modality*, 143–173. Springer.
- Moon, A. (2017). Beliefs do not come in degrees. *Canadian Journal of Philosophy*, 1–19.
- Moyse, G., Lesot, M.-J., & Bouchon-Meunier, B. (2015). Oppositions in fuzzy linguistic summaries. *Proc. of the IEEE Int. Conf. on Fuzzy Systems, FUZZ-IEEE 2015* (pp. 1–8).

- 
- Muehleisen, V. L. (1997). *Antonymy and semantic range in english*. Thèse de doctorat, Northwestern University.
- Pacuit, E., & Salame, S. (2004). Majority logic. *Proc. of Int. Conf. Principles of Knowledge Representation and Reasoning, KR 2004* (pp. 598–605). AAAI Press.
- Pacuit, E., & Salame, S. (2006). Majority logic : Axiomatization and completeness. *Studia Logica*, 1–36.
- Partee, B. H. (1989). Many quantifiers. *Proc. of the Eastern States Conference on Linguistics*, 383–402.
- Poggiolesi, F. (2008). A cut-free simple sequent calculus for modal logic S5. *The Review of Symbolic Logic*, 1, 3–15.
- Rescher, N. (1968). *Many-valued logic*. Springer.
- Sahlqvist, H. (1975). Completeness and correspondence in the first and second order semantics for modal logic. *Studies in Logic and the Foundations of Mathematics*, 82, 110–143.
- Sakama, C., Caminada, M., & Herzig, A. (2014). A formal account of dishonesty. *Logic Journal of the IGPL*, 23, 259–294.
- Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.
- Schwitzgebel, E. (2015). Belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Summer 2015 édition.
- Shafer, G. (1976). *A mathematical theory of evidence*, vol. 1. Princeton University Press.
- Shirazi, A., & Amir, E. (2007). Probabilistic modal logic. *Proc. of the Int. Conf. on Artificial Intelligence, AAAI* (pp. 489–495).
- Smith, G. F., Benson, P. G., & Curley, S. P. (1991). Belief, knowledge, and uncertainty : A cognitive perspective on subjective probability. *Organizational Behavior and Human Decision Processes*, 48, 291–321.
- Staffel, J. (2013). Can there be reasoning with degrees of belief ? *Synthese*, 190, 3535–3551.
- Straccia, U. (2001). Reasoning within fuzzy description logics. *J. Artif. Intell. Res.(JAIR)*, 14, 137–166.
- Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42, 139–165.
- Thiele, H., & Kalenka, S. (1993). On fuzzy temporal logic. *Proc. of the Second IEEE Int. Conf. on Fuzzy Systems, 1993*. (pp. 1027–1032).
- Van Benthem, J. (1984). *Correspondence theory*. Springer.
- van der Hoek, W., & De Rijke, M. (1995). Counting objects. *Journal of Logic and Computation*, 5, 325–345.

- van der Hoek, W., & Meyer, J. (1992). Graded modalities in epistemic logic. *Logical Foundations of Computer Science—Tver'92*, 503–514.
- Weber, S. (1983). A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. *Fuzzy Sets and Systems*, 11, 103–113.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.
- Zadeh, L. (1978). Fuzzy sets as the basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3–28.

## Annexe A

# Implémentation du jeu Hanabi

Cette annexe présente l'interface graphique développée pour le jeu Hanabi, par des étudiants dans le cadre de leur projet de M1, dont le contexte et la composante théorique sont présentés dans le chapitre 9.

### Présentation du logiciel

Au lancement du logiciel, une fenêtre d'accueil est ouverte, présentée en figure A.1. L'utilisateur peut lancer une partie, ce qui l'amène dans une nouvelle fenêtre avec laquelle il peut modifier les paramètres du jeu, voir figure A.2 : cela lui permet de choisir le nombre de joueurs artificiels et de moduler leur niveau.

Après l'avoir paramétrée, le joueur lance la partie : il se retrouve sur la fenêtre de jeu représentant les mains de ses coéquipiers autour d'une table, illustrée par la figure A.3. La main du joueur est représentée en bas au centre, et montre le dos des cartes, les indices donnés au joueur apparaissent au-dessus durant la partie. Au-dessous de sa main, le joueur a trois boutons qui lui permettent de choisir un coup à jouer : donner un indice, jouer une carte ou défausser une carte.

Dans le cas où le joueur souhaite donner un indice, il doit d'abord choisir son coéquipier auquel s'adresse l'indice, puis décider si celui-ci porte sur un chiffre ou sur une couleur, en cliquant sur les cercles correspondants en bas à gauche.

En fin de partie, le score est affiché et l'on propose au joueur de recommencer une partie.

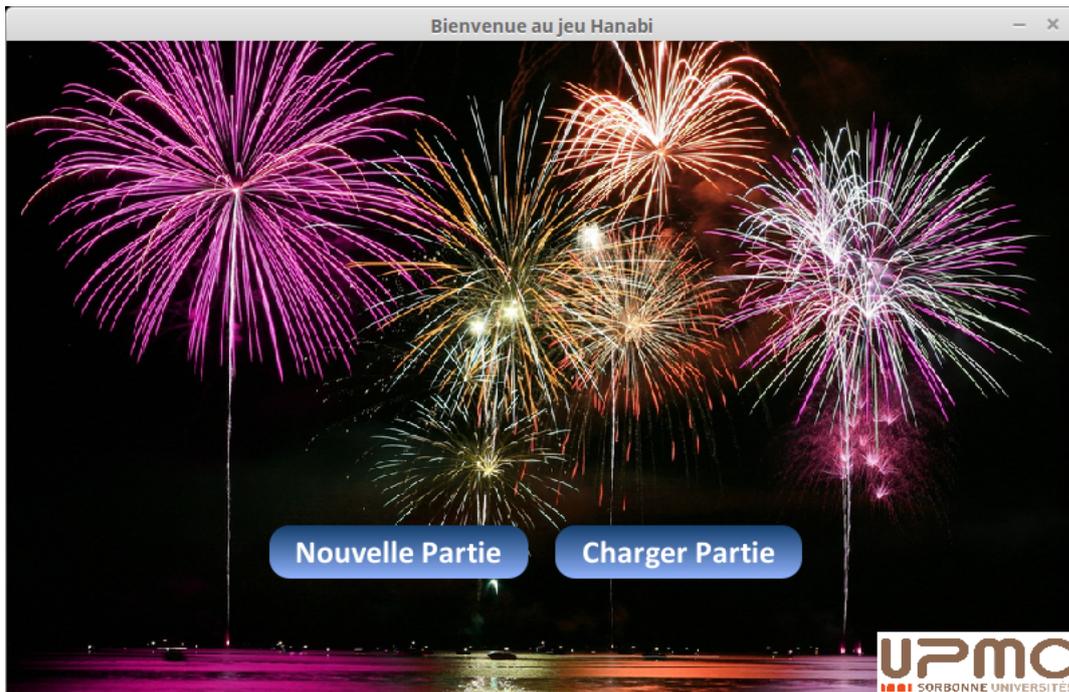


FIGURE A.1 – Fenêtre d'accueil du logiciel Hanabi

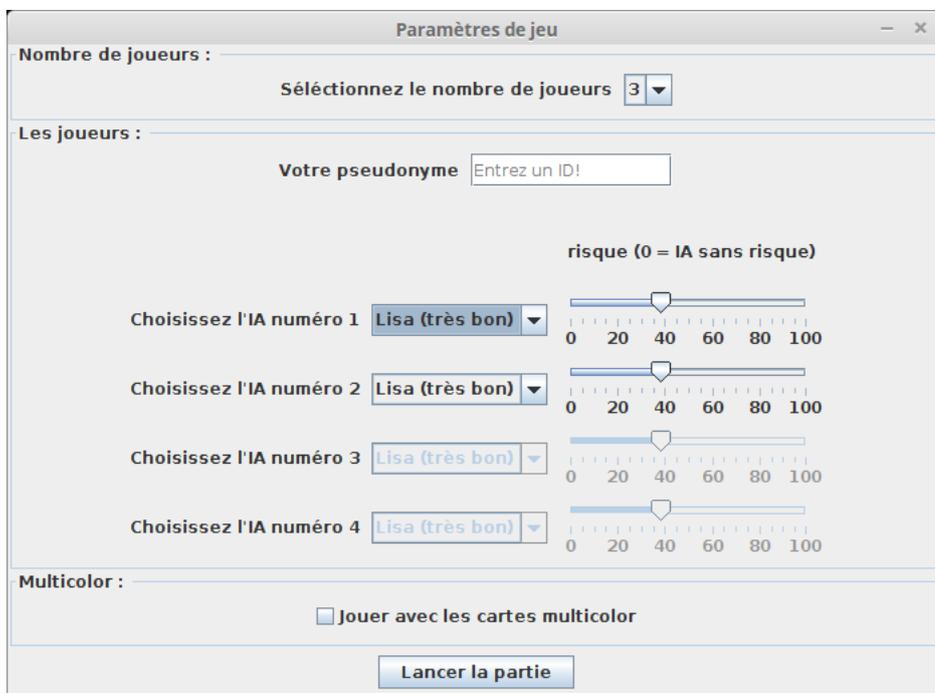


FIGURE A.2 – Fenêtre de paramétrage d'une partie



FIGURE A.3 – Fenêtre de jeu



## Annexe B

# Rappels de manipulation conjonctive et disjonctive en logiques classique et non classiques

Cette annexe rappelle brièvement les règles de transformation des formules logiques autour de la conjonction et de la disjonction, dans les différents systèmes desquels se rapproche le modèle que nous proposons dans les chapitres 5 et 6 : elle donne d'abord dans la section B.1 les principes généraux de ces manipulations puis décrit les cas de la logique classique (section B.2), modale (section B.3), floue (section B.4) et possibiliste (B.5).

On peut noter que ces formalismes s'appliquent à différents types d'objet et diffèrent quant aux concepts qu'ils représentent. En effet, la logique classique manipule des formules propositionnelles tandis que la logique modale les enrichit par des modalités. Les logiques floue et possibiliste étant des extensions pondérées de la logique classique, elles permettent de manipuler des coefficients numériques, respectivement interprétés comme degrés de vérité et de possibilité.

### B.1 Principes : factorisation et distribution

Nous considérons dans cette annexe les liens entre les formules  $\varphi$ ,  $\psi$ ,  $\varphi \wedge \psi$  et  $\varphi \vee \psi$ .

Il faut alors distinguer deux types de manipulation : la factorisation, qui combine deux formules  $\varphi$  et  $\psi$  afin d'en produire une plus complexe, comprenant un opérateur additionnel. La distribution, quant à elle, effectue une transformation contraire : elle décompose une formule complexe contenant un de ces opérateurs en un couple de formules plus simples. Il existe une réciprocity entre le processus de factorisation et celui de distribution.

On peut noter que la distribution est liée à une opération donnée par la méthode

des tableaux (voir (Hähnle, 2001) pour une revue détaillée dans les principaux formalismes logiques). Cet algorithme permet classiquement de calculer la valeur de vérité d'une formule complexe, en la décomposant récursivement en formules atomiques.

## B.2 Logique classique

En logique classique, la manipulation de formules contenant un opérateur de conjonction ou de disjonction est définie par deux règles : l'introduction et la suppression de l'opérateur. Dans le système de déduction naturelle de Gentzen (Gentzen, 1934), par exemple, ces règles pour la conjonction sont :

$$\frac{\varphi \quad \psi}{\varphi \wedge \psi} \wedge_I \quad \frac{\varphi \wedge \psi}{\varphi} \wedge_{E1} \quad \frac{\varphi \wedge \psi}{\psi} \wedge_{E2}$$

La règle d'introduction du  $\wedge$ , qui correspond à l'opération de factorisation, nommée  $\wedge_I$ , établit que si les formules  $\varphi$  et  $\psi$  sont séparément démontrées, alors leur conjonction peut être dérivée. De manière symétrique, si la conjonction entre  $\varphi$  et  $\psi$  est démontrée, alors chacun de ces termes peut être séparément dérivé par les règles d'élimination  $\wedge_{E1}$  et  $\wedge_{E2}$ , lesquelles portent une propriété de distribution.

Le cas des formules disjonctives est différent : les règles de manipulation sont :

$$\frac{\varphi}{\varphi \vee \psi} \vee_{I1} \quad \frac{\psi}{\varphi \vee \psi} \vee_{I2} \quad \frac{\varphi \vee \psi \quad \begin{array}{c} [\varphi] \\ \vdots \\ \rho \end{array} \quad \begin{array}{c} [\psi] \\ \vdots \\ \rho \end{array}}{\rho} \vee_E$$

Les règles d'introduction offrent deux règles de factorisation. Néanmoins, la règle d'élimination ne fonctionne pas de manière réciproque et se base plutôt sur une succession de preuves par épuisement des hypothèses. Ainsi, aucune connaissance spécifique au sujet de  $\varphi$  et  $\psi$  ne peut en être dérivée, n'autorisant pas de distribution du  $\vee$ .

## B.3 Logique modale

Les règles de manipulation de conjonction et disjonction en logique modale sont exprimées par les axiomes normaux (Chellas, 1980) : pour la conjonction, une équivalence existe entre la factorisation et la distribution :

$$(\Box\varphi \wedge \Box\psi) \Leftrightarrow \Box(\varphi \wedge \psi)$$

La disjonction ne peut qu'être factorisée dans ce système normal :

$$(\Box\varphi \vee \Box\psi) \Rightarrow \Box(\varphi \vee \psi)$$

Ainsi, la différence entre les deux opérateurs, déjà soulignée en logique classique, est également observée dans le cas de la manipulation de modalités.

## B.4 Logique floue

La logique floue (Zadeh, 1965) formalise la manipulation de degrés de vérité, nous utilisons ici la notation  $(\varphi, \alpha)$  pour représenter le fait que  $\varphi$  est vraie au degré  $\alpha$ .

Les combinaisons de formules consistent à agréger des degrés de vérité, utilisant des opérateurs spécifiques, appelés t-normes et t-conormes (Klement et al., 2013). Une t-norme est une fonction  $\top : [0, 1] \times [0, 1] \rightarrow [0, 1]$  associative, commutative, croissante en ses deux arguments et qui a pour éléments neutre et absorbant les valeurs 1 et 0 respectivement. De manière similaire, une t-conorme est une fonction  $\perp : [0, 1] \times [0, 1] \rightarrow [0, 1]$  qui est associative, commutative, croissante en ses deux arguments et qui a pour éléments neutre et absorbant les valeurs 0 et 1 respectivement. Une t-norme et une t-conorme sont duales si  $\top(1 - x, 1 - y) = 1 - \perp(x, y)$ . Par exemple, les fonctions maximum et minimum, nommées respectivement t-norme et t-conorme de Zadeh, forment un couple dual.

Les règles de manipulation pour les formules conjonctives et disjonctives établissent que :

$$(\varphi, \alpha) \wedge (\psi, \beta) \Rightarrow (\varphi \wedge \psi, \top(\alpha, \beta))$$

$$(\varphi, \alpha) \vee (\psi, \beta) \Rightarrow (\varphi \vee \psi, \perp(\alpha, \beta))$$

Ces règles formalisent la factorisation des formules et de leur degré associé, permettant de combiner des formules afin de constituer des formules plus complexes.

De plus, par compatibilité avec la factorisation, il est possible d'imposer des contraintes sur le processus inverse de décomposition. L'algorithme des tableaux étendu à la logique floue propose une telle règle, par exemple établi dans le cas de la logique de description floue (Straccia, 2001) : si un ensemble de formules  $\mathcal{S}$ , dont la véracité doit être vérifiée, contient  $(\varphi \wedge \psi, \gamma)$ , alors deux coefficients  $\alpha, \beta \in [0, 1]$  tels que  $\top(\alpha, \beta) = \gamma$  sont introduits, ainsi que les formules  $(\varphi, \alpha)$  et  $(\psi, \beta)$  qui sont ajoutées à  $\mathcal{S}$ .

De la même façon, si  $\mathcal{S}$  contient  $(\varphi \vee \psi, \gamma)$ , alors les mêmes formules simples sont ajoutées, mais avec la contrainte  $\perp(\alpha, \beta) = \gamma$ . Le processus de vérification consiste alors à déterminer si les contraintes numériques introduites peuvent toutes être satisfaites simultanément.

Ces propriétés donnent à la logique floue des propriétés spécifiques à la manipulation des degrés : contrairement aux logiques classique et modale, il n'y a pas de différence essentielle entre factorisation et distribution, ni entre conjonction et disjonction. En effet, la distribution est définie comme établissant l'existence de valeurs pour les degrés qui garantissent que, si la factorisation s'applique, alors on retrouve les valeurs initiales. La conjonction et la disjonction utilisent un processus identique, la seule différence résidant dans la fonction d'agrégation considérée, qui est soit une t-norme, soit une t-conorme.

## B.5 Logique possibiliste

Le formalisme de la logique possibiliste (Zadeh, 1978; Dubois et al., 1994b) est utilisé pour représenter et manipuler des degrés de possibilité ou nécessité associées aux formules logiques : pour une formule évaluée que l'on écrit  $(\varphi, \alpha)$ , le coefficient numérique  $\alpha$ , classiquement considéré comme le degré de nécessité de  $\varphi$  et noté  $N(\varphi)$ , indique que  $\varphi$  est certaine à un degré au moins égal à  $\alpha$ . La manipulation de degrés de nécessité est alors exprimée par les règles :

$$N(\varphi \wedge \psi) = \min(N(\varphi), N(\psi)) \quad N(\varphi \vee \psi) \geq \max(N(\varphi), N(\psi))$$

À nouveau, il faut noter que la règle du cas conjonctif est plus informative que celle du cas disjonctif, elle est définie comme une égalité alors que la disjonction conduit à une inégalité.