



Étude de la connaissance dans le cadre d'observations partielles : La logique de l'observation

THÈSE

présentée et soutenue publiquement le 25 octobre 2002

pour l'obtention du

Doctorat de l'université Joseph Fourier – Grenoble I
(spécialité informatique)

par

Olivier Brunet

Composition du jury

Rapporteurs : Roland Ducournau
Patrice Enjalbert

Examineurs : Jean-Claude Fernandez
Jean Goubault-Larrecq

Directeur de thèse : Jérôme Euzenat

Mis en page avec la classe thloria.

Remerciements

Je tiens à remercier...

Monsieur Jean-Claude Fernandez, professeur à l'université Joseph Fourier de Grenoble, pour avoir accepté de présider le jury de ma thèse.

Monsieur Roland Ducournau, professeur à l'université de Montpellier, pour avoir accepté d'être rapporteur de ma thèse, et d'avoir eu par la même occasion à se plonger dans un travail utilisant des techniques et des idées assez éloignées de son domaine de prédilection.

Monsieur Patrice Enjalbert, professeur à l'université de Caen, pour avoir accepté d'être rapporteur de ma thèse, et d'avoir fait preuve d'un tel enthousiasme pour mon travail.

Monsieur Jean Goubault-Larrecq, professeur à l'École Normale Supérieure de Cachan, pour avoir accepté de faire partie de mon jury, ainsi que pour m'avoir fait profiter de son immense expertise en matière de logique.

Jérôme Euzenat, chargé de recherche à l'INRIA, qui m'a aidé, soutenu, supporté, encouragé tout au long de ces trois années de thèse, qui a respecté mon indépendance et ma liberté, qui m'a fait confiance et m'a ainsi permis de mener mes réflexions comme je le voulais.

Patrick Cousot, pour m'avoir fait découvrir et apprécier les ordres partiels, et Radhia Cousot pour son aptitude à remonter le moral en une phrase.

Ma famille pour ne pas m'avoir demandé trop souvent d'expliquer ce que je faisais, le groupe Mautrans (ainsi que Frédéric Bes, qui refuse d'intégrer les Mautrans uniquement pour être cité explicitement de ces remerciements) pour la même raison et pour le soutien inter-thésards si important, Nico de Cachan, surnommé "L^AT_EX-god" à juste titre.

Merci et pardon à toutes les personnes que je ne mentionne pas, de peur d'en citer trop ou au contraire d'en oublier.

Enfin et surtout, merci à Magali qui partage ma vie, ce qui n'est pas une mince affaire, et que j'aime.

À Julie et sa maman

Table des matières

I	Préliminaires	1
1	Introduction	3
2	Contexte	9
2.1	Approximation et abstraction	9
2.2	Logiques intuitionnistes et modales	14
2.3	Connaissances et observations	18
2.4	Approche retenue	19
3	Exemple : Des cartes et des enveloppes	23
3.1	Définition du système de représentations	23
3.2	Traduction logique	25
3.2.1	Définition du vocabulaire	25
3.2.2	Liens avec les descriptions	27
3.3	Quelques propriétés du système	29
3.4	Étude statique et dynamique	30
II	Formalisation	33
4	Une représentation	37
4.1	Présentation	37
4.1.1	Exemples	38
4.1.2	Enrichissement de la définition	39
4.2	Construction de représentations	43
4.3	Logique associée aux représentations	47

5	Deux représentations	53
5.1	Constructions à partir de deux représentations	54
5.1.1	Somme disjointe	54
5.1.2	Produit réduit	55
5.2	Cas extrêmes de relations entre représentations	56
5.2.1	Cas horizontal	56
5.2.2	Cas vertical	57
5.3	Approximation et abstraction	58
5.3.1	Approximation et représentations	59
5.3.2	Logique de l'approximation	60
6	Plusieurs représentations	63
6.1	Cas sans composition	64
6.1.1	Caractérisation de l'interprétation	65
6.1.2	Caractérisation de la logique	69
6.2	Cas avec composition	70
6.2.1	Renforcement de la logique	71
7	Quelques variations logiques autour de OL	75
7.1	Formulation en calcul des séquents	75
7.1.1	Localisation des propositions	75
7.1.2	Règles à la "Gentzen" pour OL	76
7.2	Propriétés et manipulations de preuves	80
7.3	Quelques propriétés sur les indices	83
7.4	Élimination de la règle d'universalité dans le cas infini	86
8	Quelques conséquences de la logique OL	89
8.1	La double négation intuitionniste	89
8.1.1	Le protocole de transfert de bit	90
8.1.2	Interprétation de la double négation	92
8.1.3	Lien avec la logique classique	92
8.2	L'axiome 5	93
8.2.1	Lien avec le tiers-exclu	96
8.3	La confiance réciproque entre agents	96
8.4	La consistance globale	98

III	Hiérarchies de représentations	101
9	Vers les hiérarchies de représentation	105
9.1	Un ordre sur les indices	105
9.1.1	Implications sur la logique	106
9.2	Hiérarchies de représentations	107
9.2.1	Définition	107
9.2.2	Approche équivalente	108
9.3	Comportement logique	111
9.3.1	Cas avec une infinité d'indices	111
9.3.2	Cas avec un ensemble fini d'indices	113
10	Localisation et consistance	117
10.1	Localisation des propositions	118
10.2	Hiérarchies consistantes	121
10.3	Logique des hiérarchies consistantes	125
10.4	Utilisation de hiérarchies consistantes	127
10.4.1	La localisation des propositions	127
10.4.2	La manipulation des hiérarchies consistantes	128
IV	Comparaison	131
11	Synthèse	135
11.1	Nouvelle approche de la localisation	136
11.2	Fermeture des propositions	139
11.3	Les observations atomiques	141
11.4	La connaissance accessible par observation	143
12	Vers une théorie de la connaissance observable	147
12.1	Objectivité d'une proposition	147
12.1.1	Quelques approches de l'objectivité	148
12.1.2	La notion de contexte	148
12.1.3	Définition de la notion d'objectivité	149
12.1.4	Quelques propriétés de $\text{Obj}(\Gamma)$	150
12.2	L'objectivité dans le cadre de la logique de l'observation	152

13 Conclusion	155
Bibliographie	161
Annexes	165
A Index des axiomes	165
B Preuves de correction et de complétude	167
B.1 Représentations	167
B.1.1 Correction	167
B.1.2 Complétude	168
B.2 Approximations de représentations	170
B.2.1 Correction	170
B.2.2 Complétude	171
B.3 Systèmes sans composition	172
B.3.1 Correction	172
B.3.2 Complétude	172
B.4 Systèmes compositionnels	175
B.4.1 Correction	175
B.4.2 Complétude	176
B.5 Hiérarchies consistantes	177
C Preuve d'élimination de coupures	179
C.1 La réduction de localisation	179
C.2 La réduction de coupures	186
C.2.1 Quelques définitions pratiques	186
C.2.2 La réduction proprement dite	187
D Manipulation des hiérarchies	197
D.1 Opérations de manipulation des indices	197
D.1.1 Hiérarchie atomique	197
D.1.2 Ajout de la représentation vide	197
D.1.3 Suppression d'indices	198
D.2 Produit de systèmes	199
D.3 Raffinage de systèmes	200

D.4	Autres constructions	205
D.4.1	Produit de deux représentations	205
D.4.2	Application de fonctions	205
D.4.3	Partition d'une représentation	207
D.4.4	Ensembles d'idéaux	208
D.4.5	Union de deux éléments	210

Première partie

Préliminaires

Chapitre 1

Introduction

Lorsque l'on s'intéresse à la modélisation d'un problème particulier, on se base sur les connaissances que l'on a sur celui-ci. Généralement, ces connaissances proviennent d'observations que l'on fait. Mais une observation ne fournit le plus souvent que des informations partielles sur le sujet étudié. L'approximation est donc partie intégrante de toute tentative de modélisation, puisque l'on ne se base pas sur des informations complètes. On utilise au contraire des descriptions partielles pour représenter un système. Il est possible de généraliser ce propos, et d'autres raisons veulent intervenir pour justifier cette partialité de l'information représentée. Nous en donnons quelques-unes, réparties en trois catégories.

Tout d'abord, on peut classer toute cause d'imprécision subie, que l'on ne peut éviter. On peut reprendre les arguments donnés dans le cadre d'observations imparfaites d'un système. On trouve cela en physique, où toutes les données proviennent de mesures qui comportent obligatoirement une part d'incertitude, d'imprécision. Mais il ne s'agit pas de l'unique origine. Ainsi, en informatique, il peut arriver que la quantité d'information que l'on a sur un résultat dépend du temps de calcul nécessaire pour l'obtenir. Alors, si l'on veut avoir un résultat à un moment donné, il faut interrompre le calcul et se satisfaire du résultat obtenu jusqu'alors (lorsque l'on calcule les décimales de π , que l'on s'intéresse aux traces de l'exécution d'un programme, etc.).

D'autres raisons sont à l'origine de la partialité de l'information dans un modèle. On peut citer l'utilisation d'approximations voulues. Il n'est en général pas nécessaire de représenter toutes les informations disponibles pour effectuer tel ou tel calcul. Ainsi, si l'on s'intéresse à la modélisation d'un dé à 6 faces, la seule partie intéressante de cette modélisation ne sera la plupart du temps que la valeur inscrite sur la face supérieure du dé. On se soucie peu de sa composition, sa couleur, sa densité, la profondeur des points, etc.

Enfin, même une fois que l'on a sélectionné la partie des informations que l'on juge pertinente, il est possible que l'on n'ait pas envie ni besoin de les représenter avec la meilleure précision disponible. Il existe de nombreux exemples où l'on choisit un niveau de description très grossier par rapport aux informations disponibles. Ainsi, on peut modéliser des circuits électriques de façon simplifiée, et en déduire néanmoins de nombreuses propriétés intéressantes sur leur comportement. En effet, les quantités manipulées sont définies macroscopiquement (l'intensité, la tension) et un grand nombre de propriétés de ces systèmes utilisent ces quantités (ainsi, une grandeur R et l'équation $U = RI$ permettent de caractériser une résistance). La thermodynamique est un exemple encore plus flagrant, puisque son existence repose sur la possibilité de définir de telles quantités et d'énoncer des propriétés les utilisant (un exemple célèbre étant l'équation des gaz parfaits $PV = nRT$).

Aussi, comme on le voit, des approximations sont quasiment toujours présentes dans un modèle. On peut alors s'interroger sur la place qui leur est donnée en général. Leur existence est-elle prise en compte ? Comment exprimer précisément les limitations qui en découlent ? Nous nous proposons dans ce travail d'étudier ce que l'utilisation de descriptions partielles implique sur la connaissance que l'on peut avoir d'un système que l'on modélise.

Pour cela, nous nous intéressons tout d'abord à ce que peut être un modèle, et ce que représentent les éléments descriptifs de ces modèles. Une première question à considérer est la suivante : Que choisit-on de modéliser ? En effet, on ne veut pas représenter des informations que l'on ne juge pas importantes pour un besoin particulier. On effectue une sélection des propriétés que l'on veut avoir explicitement dans le modèle. Un premier choix s'opère donc, que l'on peut qualifier d'élagage "horizontal". De plus, comme nous venons de l'évoquer, on doit choisir pour les informations que l'on a décidé de représenter une précision maximale, une résolution. On est obligé en général de décider d'un niveau d'approximation, puisque tout représenter est le plus souvent impossible. C'est l'élagage "vertical". Ces choix montrent que pour un même système à modéliser, il existe de nombreuses manières de procéder. Par la suite, nous qualifierons de *représentations* les différentes façons de modéliser un système donné, pour insister sur le fait qu'ils se rapportent à la description de ce système.

La présence d'une multitude de représentations pour un même système apparaît alors comme un élément important à prendre en compte. En effet, si tous les modèles que l'on considère se rapportent au même objet, il doit exister des moyens de relier les différents modèles. Nous arrivons à un point intéressant qui servira de base à notre étude : pour étudier la notion de connaissance basée sur des descriptions imparfaites, nous ferons abstraction du système étudié pour nous concentrer uniquement sur les représentations que l'on a définies. Ainsi, au lieu de mettre en relation une description partielle à la propriété qui lui correspond dans la réalité pour le système, nous mettrons en relation les descriptions partielles provenant de différentes mo-

délimitations, et nous nous intéresserons à la façon de traduire les informations d'un modèle à l'autre.

Un point important découlant de l'étude de ces relations entre représentations est la possibilité de définir une hiérarchie entre les différents modèles, en fonction de leur précision et de leur capacité à exprimer plus ou moins d'information. Cette notion d'ordre est sous-jacente au concept de modèle. En effet, un postulat que l'on peut énoncer à ce propos, et que l'on retrouve de façon tacite dans la quasi-totalité de la démarche scientifique, est que si une représentation (ou de façon similaire un modèle) est plus précise qu'une autre, toutes les propositions accessibles depuis la représentation la moins précise le sont aussi depuis la plus précise des deux. Une formalisation rigoureuse des relations existant entre diverses modélisations peut permettre d'étudier ce type de questions.

Ce que nous proposons

Nous commencerons par étudier la question de la définition de modélisations et des rapports qui existent entre elles. Nous analyserons ce problème d'une façon très générale, pour en isoler certains aspects qui nous semblent le caractériser.

À partir de cette discussion, nous allons définir un ensemble de structures algébriques permettant de formaliser ce sujet. Principalement, deux éléments entrent en jeu pour cela : un moyen de définir une représentation, version algébrique et abstraite de la notion de modèle, et la caractérisation de méthodes permettant de relier différentes représentations. Une *représentation* sera une structure très proche de celle d'ensemble partiellement ordonné. Ses éléments, appelés *descriptions* correspondent à diverses descriptions partielles que l'on peut faire de l'état du système, et l'ordre partiel définit une relation de précision sur les descriptions. Ainsi, une description d_1 sera dite plus précise qu'une description d_2 (ce que l'on note $d_1 \leq d_2$) si toutes les informations contenues dans d_2 sont aussi présentes dans d_1 .

Puisqu'il existe en général plusieurs manières d'observer et de représenter un système, nous devons considérer l'existence de plusieurs représentations pour le même système. Pour formaliser cela, nous introduirons l'usage de fonctions permettant d'effectuer des traductions de descriptions d'une représentations à une autre. Ces fonctions sont simplement des fonctions croissantes. Si l'on a deux représentation \mathcal{R}_1 et \mathcal{R}_2 et que l'on utilise $f_{2|1} : \mathcal{R}_1 \rightarrow \mathcal{R}_2$ pour traduire une description $d_1 \in \mathcal{R}_1$, $f_{2|1}(d_1)$ est la description de \mathcal{R}_2 qui contient le plus d'informations parmi celles fournies par d_1 , et qui n'en contient pas d'autres (il peut en particulier s'agir d'une description "vide" ne fournissant aucune information). À nouveau, à l'aide de considérations générales sur ce que peut être ce processus de traduction, nous identifierons des propriétés que l'on doit imposer aux fonctions de traduction. Ainsi, la traduction d'une représentation vers elle-même est l'identité, et si l'on a trois représentations \mathcal{R}_1 , \mathcal{R}_2 et \mathcal{R}_3 , et si $f_{i|j}$ est la fonction

de traduction de \mathcal{R}_j vers \mathcal{R}_i , on veut qu'elles vérifient :

$$\forall d_1 \in \mathcal{R}_1, f_{3|1}(d_1) \leq f_{3|2} \circ f_{2|1}(d_1)$$

En d'autres termes, cela signifie que lors d'une traduction d'une représentation vers une autre, on peut perdre de l'information si l'on passe par une représentation intermédiaire.

Nous intitulerons *système de représentations* la structure algébrique comportant un ensemble de représentations et un ensemble de fonctions de traduction vérifiant ces propriétés. Cette structure va servir de base pour effectuer une étude logique précise du comportement des connaissances dans le cadre d'observations. Pour cela, nous définirons un langage propositionnel dans lequel nous exprimerons les propriétés représentant le comportement de l'information. Ce langage comprend un ensemble d'opérateurs de modalité K_i , qui permettent de faire apparaître explicitement dans les propositions l'existence de plusieurs manières d'observer un système donné. Ainsi, un opérateur K_i est associé à chacune des représentations \mathcal{R}_i , et la signification d'un terme de la forme $K_i \varphi$ est qu'en utilisant les descriptions de \mathcal{R}_i , on est sûr que la proposition φ est vérifiée par le système.

Il existe plusieurs façons de relier les descriptions des diverses représentations d'une part et les propositions de notre langage logique d'autre part. Aussi, nous identifierons plusieurs logiques, correspondant à différentes approches de la modélisation. En particulier, nous nous intéresserons à l'étude des relations entre les connaissances associées à deux représentations dont l'une est plus précise que l'autre, c'est-à-dire dont l'une contient toutes les descriptions de l'autre. On peut supposer que dans de telles conditions, toute connaissance accessible grâce à la représentation la moins précise doit l'être aussi grâce à la plus précise.

La discussion sur les différentes logiques nous permettra d'identifier une logique particulière, que nous intitulerons **OL** comme *logique d'observation*. Cette logique, basée sur la logique intuitionniste et dont les axiomes modaux sont donnés dans la table 9 p.135, permet d'exprimer simplement de nombreuses propriétés concernant les rapports entre la connaissance et l'observation. Bien que cette logique ne permette pas de relier entre elles les connaissances associées à plusieurs représentations de façon très forte (ce qui est pourtant l'un des objectifs de cette étude), nous pensons qu'il s'agit d'une modélisation adéquate du sujet que l'on s'est proposé d'étudier.

Les résultats obtenus

On peut regrouper les résultats que nous avons obtenus et que nous présentons dans ce travail en deux catégories. Tout d'abord, nous avons défini un ensemble de structures algébriques pour la formalisation de systèmes. Parmi elles, il y a les *systèmes de représentations*, évoqués

plus haut, et nous avons de plus défini les *hiérarchies de représentations*, qui sont une variante de ces systèmes de représentations. Elles diffèrent des premières en ce sens que l'on impose des conditions sur les représentations utilisées – elles doivent contenir une description “vide” ne contenant aucune information, ainsi qu’une opération de conjonction entre les descriptions – ainsi que sur les relations entre les représentations – puisqu’en utilisant la comparaison de précision entre représentations évoquée plus haut, nous imposons qu’il existe toujours une représentation plus précise que deux représentations données.

Nous présentons ces structures avec un ensemble d’outils et de constructions de base permettant de les manipuler. En particulier, ces constructions que l’on propose sont très générales et flexibles, ce qui fait de la structure de hiérarchie de représentations un formalisme adéquat pour la modélisation de systèmes.

Parallèlement à cela, nous proposons une discussion de nombreux aspects concernant le comportement logique de la connaissance basée sur l’observation, et la logique OL introduite plus haut. L’un des principaux résultats de cette discussion est la nécessité de se baser sur la logique intuitionniste dans une telle approche, ainsi que la non-vérification de l’axiome modal 5. L’origine de ces deux faits est que dans nos postulats, l’information provient d’observations effectives, et que l’on considère que ne rien observer ne fournit aucune information puisque l’on ne peut être sûr d’effectuer une observation avec une finesse, une résolution suffisantes.

Comme annoncé, nous nous intéressons de plus à la relation qu’il y a entre l’information accessible depuis des représentations dont l’une est plus précise que l’autre. Nous montrons que l’idée intuitive, selon laquelle ce qui est accessible pour la moins précise l’est aussi pour la plus précise, n’est pas vérifiée, en particulier à cause de la présence d’implication dans le langage propositionnel. À partir de ce résultat, nous nous penchons sur d’autres manières de relier les connaissances relatives à plusieurs représentations, ainsi que sur la possibilité de relier une connaissance obtenue par observation au système lui-même, au système réel. Un élément important que l’on identifie pour cela est l’importance de faire figurer, pour une observation donnée, la méthode utilisée pour obtenir ce résultat, ce qui au niveau logique se traduit par la nécessité, pour les observations, les connaissances de base que l’on manipule, d’indiquer explicitement la représentation à laquelle elles se rapportent. On se penche de plus sur une classe particulière de propositions que l’on qualifie d’*objectives*, et telles qu’elles peuvent être utilisées dans toutes les représentations sans avoir à tenir compte de leur représentation de départ.

Le résultat principal de cette discussion peut être interprété comme le fait que dans notre approche, il n’est pas possible en général d’énoncer des propriétés se rapportant au système. À la place, on doit se restreindre à des propositions qui relient entre elles les observations que l’on peut faire. Ainsi, cela donne à ces observations le statut d’éléments *réels*, puisque le système lui-même peut être oublié totalement, au profit de ces observations qui deviennent les éléments

tangibles que l'on manipule. Enfin, la discussion sur les propositions objectives montre dans quelle mesure ces informations peuvent être manipulées et regroupées, avec le fait particulier que les implications autres que celles déduites directement de la structure du système ne sont généralement pas objectives.

Structure du document

Le mémoire que nous proposons s'articule en trois parties. Tout d'abord, nous étudions de façon théorique différents aspects liés à l'observation partielle de systèmes et de définition de modèles. C'est l'occasion de définir les structures algébriques de base que nous utiliserons par la suite, comme les *représentations* et les *systèmes de représentations*. De plus, nous définissons à l'aide des systèmes de représentations la logique observationnelle **OL**, et nous en donnons quelques propriétés.

Dans la partie suivante, nous présentons une discussion sur l'existence d'un ordre entre les représentations relatives à un même système, ce qui nous conduit à définir une variante de la structure de systèmes de représentations, que nous appelons *hiérarchies de représentations*. Ces nouvelles structures nous permettent de définir une deuxième façon de modéliser la connaissance sous la forme de la logique **IS4 + KV**, et de mettre en valeur deux notions particulières s'y rattachant : la localisation de propositions et la consistance de hiérarchies.

Enfin, dans la dernière partie, nous revenons sur les deux logiques que l'on a défini dans les parties précédentes, et nous les comparons en mettant particulièrement l'accent sur l'interprétation que nous en donnons du point de vue de la modélisation de connaissances. Ainsi, nous suggérons que **OL**, bien qu'apparemment moins puissante et expressive que **IS4 + KV**, constitue dans notre approche la "bonne" logique, celle la plus adaptée, en utilisant les faiblesses apparentes de cette logique comme autant d'indications sur ce qu'il est loisible ou non d'énoncer comme propriétés concernant l'observation d'un système.

Chapitre 2

Contexte

Nous allons dans un premier temps donner un aperçu de divers travaux existant sur des éléments que nous utiliserons par la suite. Il s'agit tout d'abord de la notion d'approximation qui apparaît sous de nombreuses formes dans divers domaines de l'informatique théorique, puis nous présenterons certains aspects concernant les logiques intuitionnistes et modales ainsi que la modélisation de la connaissance en relation avec les flots d'informations et les observations partielles.

Enfin, nous précisons l'approche retenue, celle que nous développerons dans la suite de ce travail.

2.1 Approximation et abstraction

Systemes formels en intelligence artificielle

Il existe en intelligence artificielle une longue tradition d'utilisation de processus d'abstraction dans le cadre du raisonnement, qui se justifie par le fait qu'abstraire un problème permet de supprimer des éléments qui sont jugés inutiles, pour avoir ensuite moins de données à considérer et rendre le problème plus facile à traiter.

Historiquement, l'abstraction a d'abord été utilisée dans les domaines de la résolution de problèmes et de la planification [56, 46, 47] et la preuve de théorèmes [52] avant de la voir appliquée à de nombreux autres domaines (voir [32] pour un survol). Ces abstractions sont définies comme des fonctions de traduction de systèmes formels préservant certains aspects algébriques [52, 18, 32, 31], un système formel pouvant être défini comme un langage logique dont on a identifié un sous-ensemble (les axiomes), éventuellement muni d'un mécanisme de déduction ou d'inférence. Ainsi, dans [32], on définit un système formel comme un triplet $\Sigma = \langle \Lambda, \Omega, \Delta \rangle$ où Ω est un ensemble de formules du langage Λ et Δ est un mécanisme de déduction (*Deductive Machinery*) pour Σ . Dans ce cas, une abstraction écrite $f : \Sigma_1 \Rightarrow \Sigma_2$ est la donnée

de deux systèmes formels Σ_1 et Σ_2 , ainsi qu'une fonction $f : \Lambda_1 \rightarrow \Lambda_2$ de traduction du langage. Dans [31], on utilise justement la préservation ou non des trois parties d'un système formel pour classer les différentes sortes d'approximations.

Plus récemment, la notion d'abstraction a été étendue et raffinée en déplaçant la transformation du langage logique vers les modèles. Ainsi, [51] définit une abstraction comme étant une fonction transformant les interprétations de langages. On compare alors les langages en fonction de leurs modèles respectifs et de la façon dont ceux-ci sont transformés. De même, notre approche se rapproche de celle présentée dans [57]. En effet, les différents modèles que l'on considère dans l'étude d'un système correspondent à différentes façons de l'observer. Ainsi, on définit une relation d'approximation entre différentes observations d'un système (qui là sont appelés *systèmes de perception*) à l'aide d'une fonction transformant un système de perception en un autre système de perception plus simple (la simplicité étant ici définie comme le fait que les observations d'une perception peuvent être traduites en observations de l'autre système, qui est donc considéré comme moins précis ou plus simple).

Théories de Galois

Une autre approche générale de l'approximation est celle, très simple, où l'on a deux ensembles A et B , et une relation entre ces deux ensembles : $\mathcal{R} \in \wp(A \times B)$, et l'on associe à un élément $x \in A$ l'ensemble $\{y \mid x\mathcal{R}y\}$. Les éléments de l'ensemble B peuvent être vus comme autant d'attributs, ou de propriétés pouvant être vérifiées par les éléments de A . Ce genre de structure ou des structures équivalentes se retrouvent dans de nombreux domaines de la littérature, souvent sous des noms différents. L'une des dénominations les plus générales pour cela est celle d'*espace de Chu*, nommée ainsi du nom d'un étudiant de M. Barr [2, 37]. G. Birkhoff l'appelle *polarité* [6] et elle apparaît sous le nom de *contexte* chez G. Hardegree [40, 19].

On retrouve cette même structure chez J. Barwise et J. Seligman sous le nom de classification [3]. Dans ce cas, le premier ensemble (celui dont les éléments sont "étudiés" par l'intermédiaire du second ensemble) est celui des *tokens*, correspondant au système, le second ensemble étant celui des *types*, servant à *classifier* les tokens. On note alors $a \models \alpha$ pour indiquer que le token a est de type α .

On retrouve ce genre de construction sous forme ensembliste chez B. Ganter et R. Wille [28], avec la notion de treillis de concepts. Ces structures sont définies à partir de ce qui est nommé un *contexte formel*, et qui est un triplet $\langle O, A, I \rangle$ où O est un ensemble d'objets, A un ensemble d'attributs et $I \in \wp(O \times A)$ une relation d'incidence entre objets et attributs. Nous retrouvons là exactement les éléments constitutifs d'un espace de Chu. À partir d'un contexte formel, on peut associer à un ensemble d'objets X des informations symbolisées par l'ensemble $\triangleright X$ des attributs vérifiés par tous les objets, soit $\triangleright X = \{y \mid \forall x \in X, xIy\}$. De même, à un

ensemble Y d'attributs, on peut associer l'ensemble $\triangleleft Y$ d'objets correspondant à cet ensemble d'attributs. Avec ces définitions, il est possible d'étudier les différentes propriétés associées à la composition de ces fonctions (les combinaisons $\triangleleft\triangleright$, $\triangleright\triangleleft$, $\triangleleft\triangleleft$, etc.). En particulier, on a la relation importante $A \subseteq \triangleleft B \Leftrightarrow B \subseteq \triangleright A$. Si l'on ordonne les ensembles de parties $\wp(A)$ et $\wp(B)$ par l'inclusion, cette relation montre que la paire $\langle \triangleright, \triangleleft \rangle$ est une correspondance de Galois entre A et B (la définition étant la donnée de deux fonctions croissantes $f : A \rightarrow B$ et $g : B \rightarrow A$ vérifiant $\forall a \in A, \forall b \in B, a \leq g(b) \Leftrightarrow b \leq f(a)$, voir [6] pour plus de développements).

En utilisant les fonctions \triangleleft et \triangleright , on peut de plus définir la notion de concept formel, qui est une paire $\langle X, Y \rangle$ telle que $Y = \triangleright X$ et $X = \triangleleft Y$. On considère donc des ensembles d'objets et d'attributs tels que l'un définit exactement l'autre. L'ensemble des objets est alors l'*extension* du concept, alors que l'ensemble des attributs constitue son *intension*. D'après la définition, on voit que l'extension autant que l'intension permettent seuls de définir un concept. L'ensemble des concepts formels, ordonnés par inclusion des objets (et de façon équivalente par inclusion inverse sur les attributs) est un treillis complet que l'on nomme habituellement treillis de concepts (ou encore treillis de Galois).

Exemple

Pour illustrer cette approche, nous allons développer un exemple inspiré de [3], où l'on regarde une façon de classer les noms de la langue anglaise, en s'inspirant du dictionnaire *Webster*. On y trouve quatre catégories pour classer les mots : il y a les noms (*NOUN*), les adjectifs (*ADJ*), les verbes transitifs (*TR VB*) et intransitifs (*INT VB*). Voici un extrait d'une telle classification :

Webster	NOUN	INT VB	TR VB	ADJ
bet	1	1	1	0
eat	0	1	1	0
fit	1	1	1	1
friend	1	0	1	0
square	1	0	1	1
swallow	1	1	1	0

Si l'on se limite à ces seuls mots, on peut explorer la notion de concept. Par exemple, on a $\triangleright \{\text{eat}\} = \{\text{INT VB}, \text{TR VB}\}$. Dans l'autre sens, $\triangleleft \{\text{INT VB}, \text{ADJ}\} = \{\text{fit}\}$. Pour la notion de concept formel, on exige que les ensembles d'objets et d'attributs s'envoient l'un sur l'autre avec les opérations \triangleleft et \triangleright . Le premier exemple nous fournit un premier contexte formel : $\langle \{\text{eat}\}, \{\text{INT VB}, \text{TR VB}\} \rangle$. On a en effet vu la première égalité, et la seconde se vérifie facilement. Le deuxième exemple cependant ne nous fournit pas un concept formel, puisque l'on a $\triangleright \{\text{fit}\} = \{\text{NOUN}, \text{INT VB}, \text{TR VB}, \text{ADJ}\}$. On peut

vérifier que l'on a alors obtenu un concept :

$$\langle \{\text{fit}\}, \{\text{NOUN, INT VB, TR VB, ADJ}\} \rangle$$

Il est possible de représenter l'ensemble des concepts formels associé à un contexte formel sous la forme d'un graphe où chaque point est étiqueté par un ensemble (eventuellement vide) d'attributs et d'objets. Dans cette représentation, chaque point p correspond à un concept donc les objets sont ceux apparaissant dans les étiquettes des points sous p tandis que les attributs sont ceux dans les étiquettes des points au-dessus de p . La figure 2.1 donne une représentation du contexte tiré du dictionnaire *Webster*.

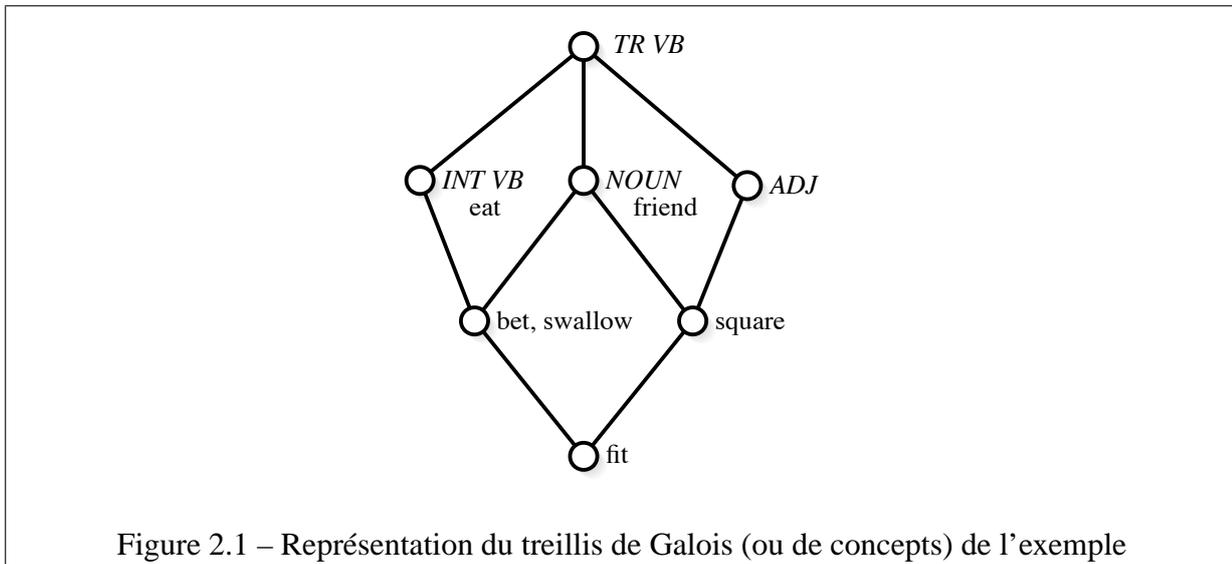


Figure 2.1 – Représentation du treillis de Galois (ou de concepts) de l'exemple

Une autre variante importante est celle utilisée dans le cadre de l'interprétation abstraite [16, 17]. Dans ce cas, le premier ensemble désigne l'espace *concret*, par opposition à l'autre ensemble qui est dit *abstrait*. Traditionnellement, on parle alors de fonction d'abstraction pour f (et on la note α) et de concrétisation pour g (on la note γ). De plus, on considère que l'ordre est inversé pour l'espace abstrait, et on obtient alors ce que rigoureusement on nomme une correspondance de Galois semi-duale (ce qui se traduit par rapport à la définition donnée ci-dessus par le renversement de l'ordre à gauche de l'implication). Elle se caractérise par l'équivalence suivante :

$$\forall a \in A, \forall b \in B, \alpha(a) \leq b \Leftrightarrow a \leq \gamma(b)$$

Ce formalisme peut alors être étendu de façon générale à des fonctions opérant sur les ensembles concrets et abstraits. Ainsi, si l'on a une fonction "concrète" $f : A \rightarrow A$, on peut définir la fonction "abstraite" associée $\alpha \circ f \circ \gamma$, qui est en général plus simple que la fonction de départ, et l'étude de propriétés sur la fonction abstraite permet d'obtenir un certain nombre d'information sur son équivalent concret.

Exemple

Un exemple simple d'un tel processus d'approximation est donné par la règle des signes lors de calculs sur les entiers. On considère que l'ensemble concret A est l'ensemble des parties de \mathbf{Z} ordonné par l'inclusion. L'ensemble abstrait B est lui constitué des différentes façons de comparer un ensemble d'entiers à zéro. Il est représenté dans la figure 2.2.

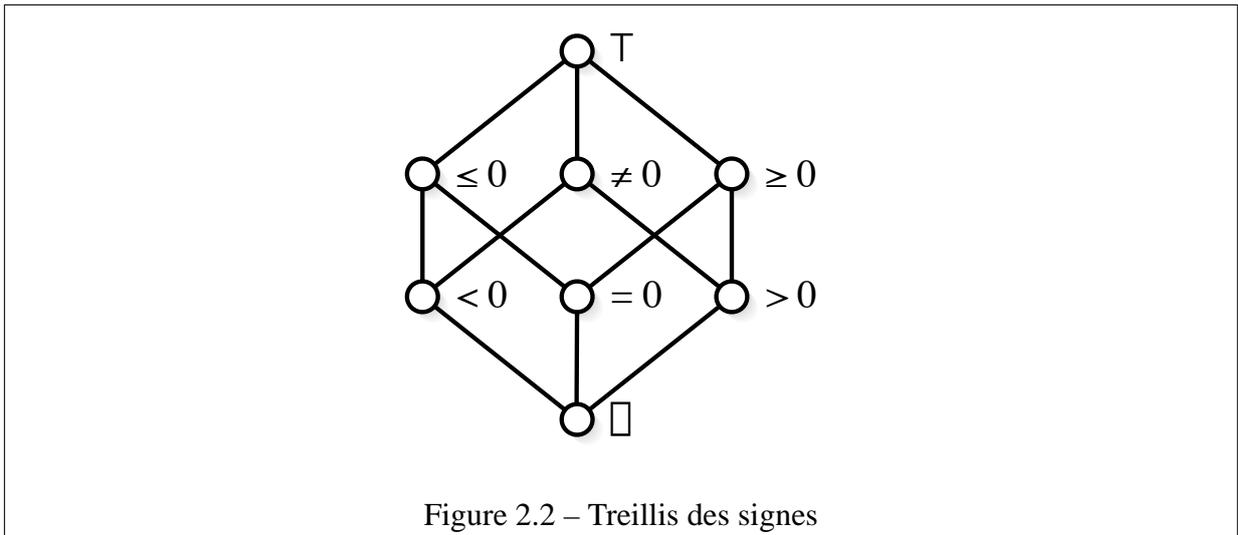


Figure 2.2 – Treillis des signes

Les fonctions α et γ se définissent alors naturellement. Ainsi, par exemple, $\gamma(\neq 0) = \mathbf{Z}^*$, ou encore $\alpha(\{0, 1, 2\}) = \geq 0$. On peut alors par exemple définir une version abstraite notée $\times^\#$ de la multiplication \times définie sur des ensembles de nombres (soit $X \times Y = \{xy \mid x \in X, y \in Y\}$). On définit alors $x \times^\# y = \alpha(\gamma(x) \times \gamma(y))$. Des calculs simples permettent de vérifier que, entre autres :

$$\begin{aligned} >0 \times^\# >0 &= >0 & >0 \times^\# <0 &= <0 \\ 0 \times^\# \top &= 0 & \neq 0 \times^\# \neq 0 &= \neq 0 \end{aligned}$$

On le voit, ce formalisme permet de manipuler la notion d'approximation ou d'abstraction de façon simple et puissante, puisque l'on peut étendre celles-ci à des fonctions et donc à envisager l'approximation dans un cadre non seulement statique (où l'on ne considère que les descriptions de façon intrinsèque) mais aussi dynamique (où l'on peut en plus avoir des relations entre descriptions, ces relations provenant de l'usage de fonctions entre elles).

Toutes ces constructions ont effectivement la même origine et les versions les plus abstraites peuvent être vues comme des généralisations et la construction de départ, exprimée par des espaces de Chu. Aussi, le point de départ est très simple, et la structure de base, du fait du peu de contraintes qu'elle impose, permet de l'appliquer à une très grande variété de domaines et

de sujets. Ainsi, on la trouve dans la littérature utilisée dans le cadre de la concurrence [37, 54], la modélisation de la logique linéaire [48, 55], l'analyse de concepts [28], l'étude de flux d'information (*information flow analysis*) [3], l'analyse statique de programmes [17, 15], et de nombreux autres encore [53, 42].

2.2 Logiques intuitionnistes et modales

Pour effectuer l'étude logique du comportement de l'information de notre approche, nous allons utiliser un langage propositionnel, et mettre en relation une modélisation sémantique de ce comportement (en utilisant des modèles, et plus particulièrement des structures algébriques inspirées de celles que l'on utilise avec les correspondances de Galois) avec une modélisation syntaxique, où nous définissons des règles de manipulation des propositions pour effectuer des déductions. Nous utiliserons pour cela deux variantes de la logique classique, qui sont la logique intuitionniste d'une part, et la logique modale d'autre part.

Logique intuitionniste

Nous allons tout d'abord présenter la logique intuitionniste. Celle-ci est issue de réflexions faites par le mathématicien hollandais L. E. J. Brouwer au début du XX^{ème} siècle à propos des mathématiques en tant qu'activité intellectuelles ([7, 8], tous deux repris et traduits dans [9]). Pour lui, les mathématiques ne sont rien d'autre qu'une construction intellectuelle, et les faits que l'on manipule ne peuvent être considérés comme vrais que si on les a prouvés. De même, si l'on affirme qu'il existe un objet vérifiant une certaine propriété P , c'est que l'on a exhibé, construit un tel objet.

Or, cela revient à ne pas accepter le principe du tiers-exclu, qui dit que pour une propriété P donnée, soit P est vraie, soit P est fausse ou dit autrement sa négation $\neg P$ est vraie (ce que l'on l'écrit $P \vee \neg P$). Par exemple, considérons la conjecture de Goldbach, qui dit que tout nombre pair est somme de deux nombres premiers (notons-la G). Les recherches actuelles ont permis de vérifier cette conjecture pour de nombreux entiers (jusqu'à environ 10^{14} [21]), et aucun contre-exemple n'a encore été trouvé. Cependant, si l'on a testé empiriquement cette conjecture, aucune preuve de cette conjecture n'existe à l'heure actuelle. Aussi, pouvons-nous écrire $G \vee \neg G$ et que cela voudrait-il signifier ?

La logique intuitionniste insiste sur l'aspect constructif d'une preuve. Ainsi, la preuve de la proposition $\varphi \rightarrow \psi$ est une construction qui transforme une preuve de φ en une preuve de ψ . La preuve de $\varphi \rightarrow \varphi$ ne dépend pas de la prouvabilité de φ . Un autre point intéressant, la preuve d'une disjonction $\varphi \vee \psi$ est donnée par soit la preuve de φ , soit la preuve de ψ . On voit donc dans ce cas que le principe du tiers-exclu n'est pas vérifié.

Nous donnons une formalisation de cette logique sous forme axiomatique à la Hilbert dans la table 1. On pourra de plus trouver une version en calcul des séquents dans la table 4 p.50. On pourra se référer à [41, 62, 26] pour plus d'informations sur la logique intuitionniste.

<p style="text-align: center;">Axiomes :</p> $\varphi \rightarrow (\psi \rightarrow \varphi)$ $(\varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow (\psi \rightarrow \vartheta)) \rightarrow (\varphi \rightarrow \vartheta))$ $\varphi \rightarrow (\psi \rightarrow \varphi \wedge \psi)$ $\varphi \wedge \psi \rightarrow \varphi \quad \varphi \wedge \psi \rightarrow \psi$ $\varphi \rightarrow \varphi \vee \psi \quad \psi \rightarrow \varphi \vee \psi$ $(\varphi \rightarrow \vartheta) \rightarrow ((\psi \rightarrow \vartheta) \rightarrow (\varphi \vee \psi \rightarrow \vartheta))$ $\perp \rightarrow \varphi$ <p style="text-align: center;">Règle :</p> $\frac{\varphi \quad \varphi \rightarrow \psi}{\psi} \text{ Modus Ponens}$ <p style="text-align: center;">Table 1 – Axiomes et règles de la logique intuitionniste</p>

Historiquement, le modèle de la logique intuitionniste est fourni par la classe des algèbres de Heyting complètes. Une algèbre de Heyting est un treillis $\langle \mathcal{L}, \leq, \vee, \wedge \rangle$ pour lequel il est possible de définir une opération de pseudo-complémentation \rightarrow telle que :

$$\forall x, y \in \mathcal{L}, \forall z \in \mathcal{L}, z \leq x \rightarrow y \Leftrightarrow x \wedge z \leq y$$

On parle d'algèbre de Heyting complète quand le treillis sous-jacent est lui-même complet, c'est-à-dire que l'on peut définir l'union et l'intersection d'une famille quelconque d'éléments de \mathcal{L} . Un exemple emblématique d'algèbre de Heyting complète est constitué, pour un ensemble E donné, de son ensemble de parties ordonné par l'inclusion. Dans ce cas, le pseudo-complément $A \rightarrow B$ est l'ensemble $\{x \mid x \in A \Rightarrow x \in B\} = \{x \mid x \notin A \text{ ou } x \in B\}$.

Il est intéressant de remarquer que les opérations de réunion, d'intersection et de pseudo-complément d'une algèbre de Heyting complète correspondent exactement aux opérations de disjonction, conjonction et implication de la logique intuitionniste. On pourra se reporter à [6] pour de plus amples développements.

Logique modale et épistémique

Nous allons maintenant présenter la logique modale, qui comme la logique intuitionniste, a été inventée au début du XX^{ème} siècle dans le cadre d’une réflexion sur divers aspects de la logique classique. Elle a été pour la première fois proposée par C. Lewis [50] qui explorait divers aspects de l’implication en logique classique. Dans cette nouvelle logique, on introduit un nouvel opérateur \Box , tel qu’une proposition $\Box\varphi$ signifie “ φ est nécessairement vraie”. L’ajout de cet opérateur permet alors de distinguer des propositions qui sont forcément vraies et qui le sont dans tous les contextes de celles qui sont vraies “par hasard” dans certains contextes. Une présentation approfondie de la logique modale peut être trouvée dans [44, 45, 14].

Il est possible de donner plusieurs interprétations du rôle de cet opérateur de modalité. Nous allons nous intéresser plus particulièrement à la logique épistémique. Dans cette approche, on utilise la logique pour modéliser le raisonnement que l’on peut faire sur la connaissance. Le premier traitement en profondeur de ce sujet a été fait au début des années 1960 par Hintikka [43]. On peut trouver un survol du domaine dans [38]. En logique épistémique, on considère un ensemble d’agents $\{\mathcal{A}_i\}_{i \in \mathcal{I}}$ qui peuvent avoir des connaissances sur le monde qui les entoure, et on associe à chacun un opérateur de modalité noté K_i (K pour *knows*). Ainsi, un terme $K_i\varphi$ signifie que l’agent \mathcal{A}_i sait l’information φ . En formalisant de différentes manières le comportement de ces opérateurs, on peut exprimer autant de façons de relier la connaissance entre les différents agents. Nous allons nous attarder sur quelques-uns de ces comportements. Pour cela, introduisons dans la table 2 quelques axiomes usuels de la logique modale.

$K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$	K
$K_i\varphi \rightarrow \neg K_i\neg\varphi$	D
$K_i\varphi \rightarrow \varphi$	T
$K_i\varphi \rightarrow K_i K_i\varphi$	4
$\neg K_i\varphi \rightarrow K_i\neg K_i\varphi$	5

Table 2 – Quelques axiomes en logique modale

Faisons un rapide survol de ces axiomes. L’axiome **K**, tout d’abord, exprime pour l’approche épistémique le fait qu’un agent peut faire des déductions. En effet, si un agent \mathcal{A}_i sait $\varphi \rightarrow \psi$ et sait φ , il peut en déduire ψ . L’axiome **D** précise que la connaissance d’un agent est consistante. Ainsi, si \mathcal{A}_i est sûr qu’une propriété φ est vraie, alors il ne peut pas être sûr aussi que sa négation $\neg\varphi$ est vraie elle aussi. L’axiome **T**, plus fort que **D**, stipule qu’un agent ne sait que des choses vraies. Ainsi, on parle bien de connaissances et non de croyances. Les deux

derniers axiomes, 4 et 5, énoncent des propriétés sur ce que l'agent sait de lui-même. Ainsi, 4 exprime le fait que si un agent \mathcal{A}_i sait une proposition φ (soit $K_i \varphi$), alors il sait qu'il la sait (soit $K_i K_i \varphi$). Inversement, pour l'axiome 5, si un agent ne sait pas une proposition φ (soit cette fois $\neg K_i \varphi$), il sait qu'il ne la sait pas (soit $K_i \neg K_i \varphi$).

Pour avoir un survol complet de la définition d'éléments usuels en logique modale, il faut inclure la règle de nécessité notée **Nec**, qui indique que les agents connaissent tous les théorèmes de la logique considérée. En d'autres termes, si $\vdash \varphi$, alors $\vdash K_i \varphi$.

À l'aide de ces axiomes, on peut donc définir plusieurs logiques. Nous allons dans un premier temps nous pencher sur la logique appelée usuellement **S5**. Il s'agit de la logique obtenue en utilisant les axiomes **K**, **T**, 4 et 5 ainsi que la règle **Nec**. Un ensemble de travaux considèrent qu'elle est adéquate pour modéliser la connaissance (on peut citer comme représentant de ce courant le livre [23]). Il s'agit d'une approche très forte de la connaissance, puisqu'alors, un agent a des connaissances très précises sur le monde. Il sait que ses connaissances sont exactes, et sait ce qu'il sait et ce qu'il ne sait pas. Dans notre optique, cette logique est trop forte, puisque si l'on suppose que les connaissances viennent des observations, l'axiome 5 ne peut être vérifié en général. Ce point sera d'ailleurs développé dans la dernière partie.

Un autre ensemble de logiques à considérer est constitué des logiques modales intuitionnistes [58]. Un grand nombre de travaux leur a été consacré ces dernières années, d'une part du fait de grand nombre d'applications (systèmes distribués [60], vérification de circuits [24], langages de programmation [20], etc.). D'autre part, de nombreuses études ont été faites pour exhiber des modèles de ces logiques. On peut en particulier s'attarder sur la logique **IS4** qui est la logique intuitionniste à laquelle on ajoute les axiomes **K**, **T** et 4 et **Nec**. Le type de modèle le plus souvent utilisé dans le cadre des logiques modales sont les structures de Kripke. Une structure de Kripke est composée d'un ensemble d'éléments appelés mondes, et de relations binaires entre ces mondes, une par agent. La grande souplesse de ces structures vient du fait que le type de relation modélisant une modalité K_i découle directement des axiomes utilisés. Ainsi, dans le cas de **S4**, les relations sont des ordres partiels, pour **S5**, il s'agit de relations d'équivalences.

On connaît depuis les travaux de Gödel l'existence de liens forts entre la logique modale **S4** et la logique intuitionniste puisqu'il est possible de traduire la seconde dans la première [33]. Des travaux plus récents ont identifié des modèles variés pour la logique intuitionniste **IS4**, incluant des constructions catégoriques [4, 5], des modèles computationnels [34] et géométriques [35, 36], des systèmes de représentation [12], etc.

2.3 Connaissances et observations

Dans le cadre des relations existant entre la connaissance et l'observation, il existe un certain nombre de travaux où la première notion est basée sur la seconde. Nous allons nous pencher sur les travaux portant sur les flux d'information (*information flow*). Initiés dans un cadre formel par F. Dretske [22], ils ont conduit à de riches développements. Ainsi, dans [3], J. Barwise et J. Seligman s'intéressent aux liens qui existent entre différentes façons de représenter la réalité, et étudient comment mettre en relation les théories logiques obtenues. Ils se basent sur un formalisme où chaque point d'observation est décrit par une *classification*. Cette structure correspond exactement à un espace de Chu, et est donc constituée de trois éléments : un ensemble A d'objets à classifier, un ensemble Σ_A de types utilisés pour cette classification, et une relation de typage \models_A entre A et Σ_A .

Exemple

On peut considérer le cas où l'on lance un dé à différents moments $t \in T$. Le résultat d'un tel lancer peut être exprimé par un nombre compris entre 1 et 6. Soit donc $\Sigma_T = \{1, \dots, 6\}$ l'ensemble des types associé aux éléments de T . On peut alors définir une relation de typage entre ces deux ensembles en affirmant que $t \models_T n$ si et seulement si le résultat du lancer de dé à l'instant t est le nombre n .

Une telle classification permet alors de définir une étude logique du comportement du système, en étudiant des relations vérifiées par les types de la classification. Pour cela, on définit un séquent comme étant une paire $\langle \Gamma, \Delta \rangle$ d'ensembles de types, et on dit qu'une classification $\mathbf{A} = \langle A, \Sigma_A, \models_A \rangle$ vérifie ce séquent (ce que l'on note $\Gamma \vdash_{\mathbf{A}} \Delta$) si et seulement si :

$$\forall x \in A, (\forall \gamma \in \Gamma, x \models_A \gamma) \Rightarrow \exists \delta \in \Delta : x \models_A \delta$$

On peut donc associer à une classification \mathbf{A} la théorie $\text{Th}(\mathbf{A})$ définie comme l'ensemble des séquents vérifiés par \mathbf{A} .

De plus, il est possible de définir des relations entre différents classifications : étant donné deux classifications \mathbf{A} et \mathbf{B} , on définit un *infomorphisme* entre \mathbf{A} et \mathbf{B} comme une paire $\langle f^\wedge, f^\vee \rangle$ où l'on a $f^\wedge : \Sigma_A \rightarrow \Sigma_B$ et $f^\vee : B \rightarrow A$, et vérifiant :

$$\forall \alpha \in \Sigma_A, b \in B, f^\vee(b) \models_A \alpha \Leftrightarrow b \models_B f^\wedge(\alpha)$$

On peut remarquer que la définition d'un infomorphisme entre deux classifications est très proche de celle d'une correspondance de Galois semi-duale entre deux ensembles partiellement ordonnés. Il est d'ailleurs possible de développer ce parallèle et il a été montré dans [11, 10] que

les classifications pouvaient être rendues plus expressive en les transformant de façon à utiliser réellement des correspondances de Galois et non des isomorphismes.

Le fait de pouvoir définir sur les classifications une méthode pour étudier logiquement le phénomène de typage d'une part et une méthode pour relier entre elles différentes classifications d'autre part permet d'étendre l'étude logique à plusieurs classifications, et donne ainsi la possibilité de d'étudier la façon dont l'information (symbolisée par la relation de typage des différentes classifications) est manipulée, transportée, transformée entre différentes façons de représenter un même système.

Ces travaux montrent comment aborder de façon concrète le problème des relations existant entre différentes méthodes d'acquérir de la connaissance pour un même sujet, un même système. L'approche proposée est très flexible, et dans la suite, nous nous inspirerons grandement de sa structure.

Il convient pour finir d'évoquer les travaux mettant en relation la topologie et la calculabilité. Le slogan, exprimé par M. Smyth dans [59], est que les ouverts sont des propriétés semi-décidables. Cela a conduit au développement de la vision de la topologie (autrement appelée logique géométrique) comme une logique des observations finies [1]. Dans ce contexte, l'information est supposée provenir d'observations, et l'accent est mis sur l'aspect fini de la quantité d'information obtenue, ce qui limite la connaissance à des propriétés semi-décidables. Par exemple, si l'on a un nombre réel x et que l'on regarde ses décimales une par une (ce qui fait qu'à tout moment, on ne connaît qu'un nombre fini de décimales), savoir si x est différent de zéro ou non est un problème semi-décidable.

Cette approche diffère de celle que nous allons suivre, puisque nous mettrons l'accent sur le fait que les observations ne permettent de considérer que l'on a une description complète de système, et non sur le fait que les descriptions correspondent à une quantité finie d'informations. Cependant, ces deux notions sont liées, et il est intéressant d'avoir en tête ce parallèle.

2.4 Approche retenue

Après ce survol de différents travaux concernant des éléments que nous allons utiliser dans notre étude, nous allons donner quelques précisions sur les objets que nous voulons modéliser, en particulier les postulats que nous utiliserons.

Tout d'abord, qu'entendons-nous par système. Il s'agit pour nous d'une partie du monde dont les limites sont identifiées, et qui peut évoluer, se transformer, tout en restant identifiable. Ainsi, une lampe de poche peut être vue comme un système. Même si l'on déplace cette lampe

de poche, qu'on l'allume, etc., on reconnaît néanmoins que l'on a toujours une lampe de poche. Nous avons évoqué des modifications temporelles d'une lampe de poche, mais le fait de reconnaître un système va au delà de cela. Ainsi, on peut considérer plusieurs lampes de poches, de différentes formes, couleurs, puissances, on parlera de lampe de poche dans tous les cas, et la lampe de poche que l'on tient dans la main pourra être vue comme une *instance* du système.

Ainsi, un système possède une structure connue, qui est vérifiée par toutes les instances du système. C'est d'ailleurs cette structure que nous utiliserons comme caractérisation d'un système particulier. En utilisant cette structure, nous pourrions alors nous intéresser à la description d'instances particulières du système.

Nous allons faire un postulat sur l'état du système : dans le processus de description que nous allons faire, nous supposons que nous décrivons une instance donnée, et bien déterminée du système. Dans ce cas, toutes les observations que l'on va faire, toutes les informations que l'on va acquérir sur l'instance précise que l'on étudie (dit autrement, sur l'état du système) seront cohérentes entre elles. Ce postulat ne signifie pas une correction implicite des observations, mais illustre le fait que toutes les informations que l'on obtient proviennent d'observations du même système. En effet, cela revient à dire que les affirmations du type "*en effectuant cette observation-ci, nous avons trouvé ce résultat-là*" sont correctes, et correspondent effectivement à la réalité. Mais nous n'affirmons pas que le résultat de l'observation correspond lui à la réalité. Aussi, ce sont les observations qui sont réelles, pas leur résultat.

La description d'un système et de son état comporte donc deux parties : d'une part l'information que nous qualifierons d'externe ou de structurelle, qui comporte les données relatives à la structure du système, comme en particulier les états possibles du système, ainsi que les différentes façons de représenter, d'observer le système. On peut parler d'information structurelle complète, totale, puisque il s'agit des données de base que l'on considère, et tout ce que l'on pourra dire par la suite reposera sur un tel ensemble d'informations structurelles.

D'autre part, la description du système comporte l'information décrivant son état "actuel", réel. Ainsi, on a des données sur l'instance du système à laquelle on a précisément affaire. Cette information est incomplète, puisque l'on n'a qu'une description partielle de l'état du système. Cela correspond à un ensemble d'états possibles du système. On se base donc sur une approche "mondes possibles" à la Kripke où une description peut être vue comme la caractérisation d'un ensemble d'états que l'on considère comme tous possibles.

Cette approche aura de l'importance pas la suite. En effet, elle conduit à considérer comme naturelle la notion de conjonction d'informations, qui se traduit par l'intersection des ensembles de possibilités. Au contraire, la disjonction apparaît comme moins préférable, puisque affirmer que *le système n'est pas dans l'état d_1 , ou bien il n'est pas dans l'état d_2* n'est pas une description "constructive" de l'état du système, comme peut l'être une conjonction.

Ces précisions étant données, nous allons présenter quelques exemples simples illustrant notre approche, avant de définir les structures algébriques que nous utiliserons pour la formalisation et l'étude de systèmes observés.

Chapitre 3

Exemple : Des cartes et des enveloppes

Avant d'entrer dans la présentation du formalisme que nous voulons développer, nous allons illustrer dans ce chapitre quelques uns des éléments principaux de ces travaux dans un exemple. Pour cela, considérons une expérience où l'on a deux candidats : Alice et Bernard, deux cartes, l'une rouge et l'autre noire, ainsi que deux enveloppes. On commence par mélanger les cartes, on en dispose une dans chacune des deux enveloppes, et on les distribue aux deux candidats. Chacun est alors libre d'ouvrir son enveloppe comme il le désire et regarder quelle est la carte en sa possession. De plus, on impose comme condition que aucun des candidats ne peut voir le contenu de l'autre enveloppe, mais ils peuvent voir si l'autre a ouvert son enveloppe ou non.

3.1 Définition du système de représentations

Intéressons-nous tout d'abord à la façon dont on peut représenter la connaissance pour l'un des deux candidats, disons pour Alice. Tout d'abord, nous allons considérer deux descriptions, que nous nommerons N et R, et qui correspondent respectivement au cas où Alice trouve une carte *Noire* ou *Rouge* dans son enveloppe. Ces deux descriptions correspondent au cas où Alice a ouvert son enveloppe. Définissons justement une autre description, nommée A (comme *Alice*), qui correspond au fait qu'Alice a effectivement ouvert son enveloppe, sans donner d'indication sur le contenu de celle-ci. Il s'agit d'une description moins précise que les deux premières. On a donc : $N \leq A$ et $R \leq A$. Par contre, comme il n'y a qu'un seul résultat possible pour l'ouverture de l'enveloppe, les descriptions N et R sont incompatibles, c'est-à-dire qu'il n'existe pas de descriptions plus précises que les deux en même temps.

Puisqu'Alice voit si Bernard a ouvert son enveloppe, introduisons une description B pour le cas où Bernard a effectivement ouvert son enveloppe. Il est possible d'être dans un cas où Alice a ouvert son enveloppe (et connaît le résultat, donc la description correspondante est N ou R) et où Bernard a lui aussi ouvert son enveloppe. On décrit ces deux cas par deux nouvelles

descriptions NB et RB.

Enfin, ajoutons la description vide \top pour compléter cet ensemble de descriptions. On obtient une représentation \mathcal{R}_A (voir la figure 3.1).

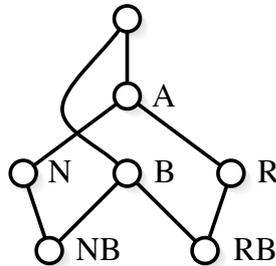


Figure 3.1 – Représentation associée à Alice

La description \mathcal{R}_B associée Bernard est exactement la même que \mathcal{R}_A . Cependant, pour garder des noms de descriptions suffisamment “parlants”, nous inversons les A et les B dans les noms des descriptions. Comment relier les descriptions des deux représentations entre elles ?

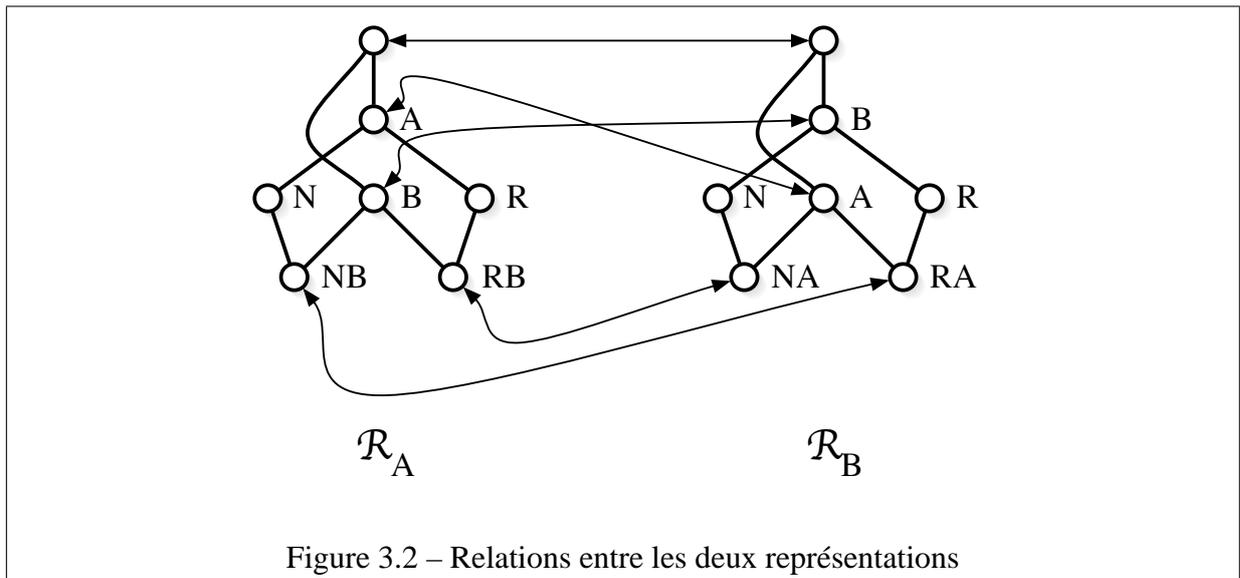
Tout d’abord, les descriptions vides \top_A et \top_B , correspondant toutes deux à une absence d’informations, peuvent être mise en relation naturellement.

La description A de \mathcal{R}_A , qui correspond au fait qu’Alice a ouvert son enveloppe peut être mise en relation avec la description A de \mathcal{R}_B qui correspond au fait que Bernard a vu qu’Alice a ouvert son enveloppe. Il en est de même pour les description B. Par contre, pour les descriptions N et R des deux descriptions, elles ne peuvent pas être mises en relation avec une description de l’autre représentation, puisque aucun des candidats de l’expérience ne peut savoir le contenu de l’autre enveloppe. Il y a néanmoins de l’information qui transite entre les deux représentations dans ce cas là, car si Alice a ouvert son enveloppe et trouvé la carte noire (description N de \mathcal{R}_A), alors en particulier elle a ouvert son enveloppe ($N \leq A$) et donc Bernard a vu Alice ouvrir son enveloppe. Aussi, la description N de \mathcal{R}_A entraîne la description A de \mathcal{R}_B , laquelle n’entraîne en retour que $A \in \mathcal{R}_A$.

Enfin, la description NB de \mathcal{R}_A exprime le fait qu’Alice a ouvert son enveloppe et trouvé la carte noire, et elle a vu Bernard ouvrir la sienne. Dans ce cas, Bernard a forcément trouvé la carte rouge et sait aussi que tous deux ont ouvert leur enveloppe. En d’autres termes, on associe NB à RA, et RB à NA.

Ces correspondances entre descriptions sont illustrées dans la figure 3.2.

Nous avons ainsi défini précisément notre système de représentations, c’est-à-dire un ensemble de représentations (une par point de vue, ici une par candidat), ainsi que les fonctions permettant de les relier. Ces fonctions sont déduites des relations entre descriptions que nous avons données. Ainsi, on a deux fonctions $f_{B|A} : \mathcal{R}_A \rightarrow \mathcal{R}_B$ et $f_{A|B} : \mathcal{R}_B \rightarrow \mathcal{R}_A$, qui vérifient



par exemple, d’après la discussion précédente :

$$\begin{aligned} f_{B|A}(\text{NB}) &= \text{RA} & f_{A|B}(\text{RA}) &= \text{NB} \\ f_{B|A}(\text{N}) &= \text{A} & f_{B|A}(\text{A}) &= \text{A} \end{aligned}$$

Ces fonctions sont croissantes, et vérifient de plus :

$$d_A \leq f_{A|B} \circ f_{B|A}(d_A) \quad d_B \leq f_{B|A} \circ f_{A|B}(d_B)$$

On trouve là une version particulière d’une propriété que nous retrouverons par la suite, et qui sera très importante, puisqu’elle exprime le fait que l’application des fonctions $f_{\cdot|_i}$ se traduit par une perte plus ou moins importante des informations contenues dans les descriptions.

À l’aide de cette structure, nous allons pouvoir commencer une étude précise du comportement logique que l’on peut associer à cette expérience.

3.2 Traduction logique

3.2.1 Définition du vocabulaire

Pour faire cette étude, nous devons d’abord définir les “briques de base”, c’est-à-dire les propositions élémentaires que nous utiliserons. Nous proposons d’en définir quatre. Tout d’abord, il y a globalement deux comportements possibles pour l’expérience : soit Alice reçoit la carte noire, soit elle reçoit la rouge. Nous allons symboliser ces deux comportements par deux propositions atomiques : *AliceNoire* et *AliceRouge* qui correspondent respectivement – comme leur intitulé l’indique – à une carte noire ou rouge dans l’enveloppe d’Alice.

Nous allons introduire deux autres propositions atomiques pour symboliser les actions possibles d’Alice et de Bernard, c’est-à-dire pour symboliser le fait qu’ils ont ouvert leur enveloppe. Soient donc deux propositions atomiques *AliceOuvert* et *BernardOuvert*.

Ainsi, notre vocabulaire de base est constitué de quatre propositions : *AliceNoire*, *AliceRouge*, *AliceOuvert* et *BernardOuvert*. Nous allons pouvoir maintenant combiner celles-ci en des propositions plus compliquées pour exprimer des propriétés plus complexes se rapportant au système. En utilisant les connecteurs logiques usuels (la conjonction \wedge , la disjonction \vee , l’implication \rightarrow et la négation \neg), on peut facilement écrire de telles propositions. Étant donné une proposition φ , nous écrirons $\vdash \varphi$ pour indiquer que le système vérifie cette proposition. Plus précisément, nous voulons que cela corresponde à des propositions qui sont toujours vérifiées par les expériences, c’est-à-dire ne dépendant pas de la carte placée dans l’enveloppe d’Alice ni de l’ordre dans lequel sont ouvertes les enveloppes. On veut donc avoir de cette façon des informations *a priori* sur le déroulement de l’expérience.

On peut exprimer simplement quelques propositions de cette sorte, comme par exemple :

$$\begin{aligned} \vdash \neg(\textit{AliceNoire} \wedge \textit{AliceRouge}) \\ \vdash \textit{AliceNoire} \vee \textit{AliceRouge} \end{aligned}$$

La première proposition exprime le fait que seule une carte se trouve dans l’enveloppe d’Alice, alors que la seconde indique qu’il y a au moins une carte. Elles illustrent bien le rôle que l’on veut attribuer aux propositions que l’on considère, c’est-à-dire que quelle que soit la façon dont se déroule l’expérience, ces deux propositions seront *a priori* vérifiées. Nous disons *a priori* puisque dans la suite, nous présenterons une façon formelle pour vérifier si une propriété φ est bien telle que $\vdash \varphi$ ou non. Aussi nous *pensons* que les propriétés précédentes sont vérifiées par le système mais plus tard, nous pourrions essayer de le prouver.

Dans le langage des propositions que nous allons utiliser par la suite, nous allons introduire deux autres opérateurs (notés K_A et K_B) qui permettront d’indiquer précisément la connaissance d’Alice ou de Bernard, soit de l’un des observateurs du système. Ainsi, $K_A \varphi$ sera la proposition associée au fait qu’Alice a suffisamment d’informations pour prouver que φ est vraie. Un exemple de propriété que l’on peut exprimer à propos du système et utilisant ces opérateurs peut être :

$$\textit{AliceNoire} \wedge \textit{AliceOuvert} \rightarrow K_A \textit{AliceNoire}$$

En langage courant, cette proposition signifie que si la carte noire est dans l’enveloppe d’Alice (*AliceNoire*) et si Alice a ouvert son enveloppe (*AliceOuvert*), alors elle sait quelle carte est dans son enveloppe. De même, on peut avoir :

$$\textit{AliceOuvert} \rightarrow K_A \textit{AliceOuvert} \wedge K_B \textit{AliceOuvert}$$

La signification est que si Alice a ouvert son enveloppe, alors à la fois Alice et Bernard le savent. Il s'agit là de la traduction sous forme logique d'une condition utilisée pour définir l'expérience. Nous espérons donc qu'elle sera effectivement vérifiée par le système.

3.2.2 Liens avec les descriptions

Comme annoncé plus tôt, nous allons relier de façon précise les propositions logiques que l'on pourra considérer avec des ensembles de descriptions de notre système de représentation. Cette façon de procéder permet de traduire mathématiquement l'un des postulats de départ que nous avons énoncé, qui est que la connaissance se base sur des observations. Ici, les informations liées directement aux observations faites par Alice et Bernard sont représentées par les descriptions des représentations \mathcal{R}_A et \mathcal{R}_B définies plus haut, et ce seront elles seules qui seront utilisées pour étudier le système. Comme il y a deux représentations, nous définirons donc deux fonctions de traduction d'une proposition vers un ensemble de descriptions. Si l'on note \mathcal{L} l'ensemble des propositions que l'on peut écrire (sans se soucier de leur sens ou de leur véracité), nous aurons donc une fonction $\llbracket \cdot \rrbracket_A : \mathcal{L} \rightarrow \wp(\mathcal{R}_A)$ qui associe une proposition φ à l'ensemble $\llbracket \varphi \rrbracket_A$ des descriptions de \mathcal{R}_A qui permettent à Alice de prouver que φ est vérifiée, et de même une fonction $\llbracket \cdot \rrbracket_B : \mathcal{L} \rightarrow \wp(\mathcal{R}_B)$ pour Bernard.

Propositions atomiques Pour effectuer cette traduction, il nous faut tout d'abord indiquer à quoi correspondent les propositions atomiques que nous avons définies précédemment. Commençons par *AliceNoire*. Cette proposition indique que la carte d'Alice est la noire. Quel peut être l'ensemble $\llbracket \text{AliceNoire} \rrbracket_A$? Comme on vient de l'indiquer, il s'agit des descriptions de \mathcal{R}_A correspond à une situation où Alice a assez d'informations pour prouver qu'elle a bien la carte noire. D'après la définition de \mathcal{R}_A que l'on a donnée, on a :

$$\llbracket \text{AliceNoire} \rrbracket_A = \{N, NB\}$$

De façon similaire, on peut écrire $\llbracket \text{AliceRouge} \rrbracket_A = \{R, RB\}$. Si l'on se penche maintenant sur l'interprétation de ces propositions atomiques en termes de descriptions de \mathcal{R}_B , on s'aperçoit que la situation s'inverse, puisque pour Bernard soit sûr qu'Alice a la carte noire, il faut qu'il ait ouvert son enveloppe et ait trouvé la carte rouge. On a donc :

$$\llbracket \text{AliceNoire} \rrbracket_B = \{R, RA\} \quad \llbracket \text{AliceRouge} \rrbracket_B = \{N, NA\}$$

Traisons maintenant le cas de la proposition *AliceOuvert*, et commençons par le point de vue de Bernard. Si Alice a ouvert son enveloppe, Bernard l'a vu, et sa connaissance correspond donc à une description au moins plus petite que A. Aussi, on peut écrire $\llbracket \text{AliceOuvert} \rrbracket_B = \{A, NA, RA\}$. Du point de vue d'Alice, nous considérerons que dès qu'elle a ouvert l'enveloppe,

elle en connaît le contenu, c'est-à-dire que si l'enveloppe est ouverte, alors Alice sait si la carte à l'intérieur est la noire ou bien la rouge. On pose donc :

$$\llbracket AliceOuvert \rrbracket_A = \{N, R, NB, RB\}$$

En particulier, cela signifie que la description A de \mathcal{R}_A n'est pas dans $\llbracket AliceOuvert \rrbracket_A$. De même, on définit :

$$\llbracket BernardOuvert \rrbracket_A = \{B, NB, RB\} \quad \llbracket BernardOuvert \rrbracket_B = \{N, R, NA, RA\}$$

Connecteurs logiques Une fois que la traduction a été définie pour les propositions de base, nous allons maintenant voir comme étendre la traduction à des propositions plus compliquées. Commençons par la conjonction : étant donné deux propositions φ et ψ , quelles sont les descriptions d'une représentation \mathcal{R}_i permettant de prouver $\varphi \wedge \psi$? Ce sont exactement les descriptions appartenant à la fois à la traduction $\llbracket \varphi \rrbracket_i$ de φ et à $\llbracket \psi \rrbracket_i$ de ψ . On déduit donc simplement la formule :

$$\llbracket \varphi \wedge \psi \rrbracket_i = \llbracket \varphi \rrbracket_i \cap \llbracket \psi \rrbracket_i$$

On traite la disjonction de façon similaire, pour obtenir la formule $\llbracket \varphi \vee \psi \rrbracket_i = \llbracket \varphi \rrbracket_i \cup \llbracket \psi \rrbracket_i$. En effet, il est naturel d'interpréter la disjonction $\varphi \vee \psi$ comme l'ensemble des descriptions permettant de prouver soit φ , soit ψ . L'implication a un traitement un peu plus délicat, car il faut prendre en compte explicitement une propriété des ensembles de la forme $\llbracket \varphi \rrbracket_i$ que l'on n'a pas évoquée jusqu'à présent. Ainsi, si une description $d \in \mathcal{R}_i$ permet de prouver une propriété φ (soit en d'autres termes une description d telle que $d \in \llbracket \varphi \rrbracket_i$), alors toute description d' plus précise que d doit aussi permettre de prouver φ . En d'autres termes :

$$d' \leq d \Rightarrow d \in \llbracket \varphi \rrbracket_i \Rightarrow d' \in \llbracket \varphi \rrbracket_i$$

Cette propriété dite *de clôture* est vérifiée par les traductions que l'on a donné des propositions atomiques, et les traductions de la conjonction et de la disjonction la conservent. Pour l'implication, on pourrait comme précédemment s'inspirer d'une interprétation ensembliste standard en posant :

$$\llbracket \varphi \rightarrow \psi \rrbracket_i = \{d \mid d \in \llbracket \varphi \rrbracket_i \Rightarrow d \in \llbracket \psi \rrbracket_i\}$$

Mais l'ensemble à droite de l'égalité ne vérifie pas toujours la propriété précédente, c'est-à-dire que l'on peut avoir deux description d et d' telles que $d' \leq d$, $d \in \llbracket \varphi \rightarrow \psi \rrbracket_i$ mais $d' \notin \llbracket \varphi \rightarrow \psi \rrbracket_i$. Cela signifie que l'ensemble $\llbracket \varphi \rightarrow \psi \rrbracket_i$ est trop grand, et ne prends pas en compte la propriété de clôture que l'on veut imposer. Nous sommes donc obligés de modifier la définition, pour obtenir :

$$\llbracket \varphi \rightarrow \psi \rrbracket_i = \{d \mid \forall d' \leq d, d' \in \llbracket \varphi \rrbracket_i \Rightarrow d' \in \llbracket \psi \rrbracket_i\}$$

Il reste à définir une traduction pour les opérateurs K_i . Cela se fait en revenant à la définition des fonctions $f_{\cdot|\cdot}$ qui permettent de relier les descriptions de représentations différentes. Ainsi, si l'on a deux descriptions $d_A \in \mathcal{R}_A$ et $d_B \in \mathcal{R}_B$ telles que $f_{B|A}(d_A) = d_B$, que cela signifie t'il ? Que d_B est la description de \mathcal{R}_B qui correspond de façon la plus proche à l'état décrit par d_A , mais sans donner plus d'informations que d_A . Ainsi, si $d_B \in \llbracket K_A \varphi \rrbracket_B$, cela signifie que d_B , transformée en la description $f_{A|B}(d_B)$ de \mathcal{R}_A est suffisante pour s'assurer que φ est vérifiée. On a donc :

$$\llbracket K_i \varphi \rrbracket_j = \{ d \mid f_{ij}(d) \in \llbracket \varphi \rrbracket_i \}$$

On suppose de plus que l'on a $\llbracket K_i \varphi \rrbracket_i = \llbracket \varphi \rrbracket_i$ pour traiter aussi le cas où les deux indices sont égaux. Nous avons donc défini des traductions pour la plupart des opérateurs logiques. Une liste complète sera donnée plus tard (voir table 6 p.65).

Signification du symbole \vdash Maintenant que l'on a défini une méthode de traduction pour les propositions en termes de descriptions, comment peut-on traduire le fait qu'une proposition φ vérifie $\vdash \varphi$? D'après la discussion faite plus haut, elle signifie que quelque soit le déroulement de l'expérience, on peut prouver que φ est vérifiée, soit en d'autres termes que pour un observateur i donné, quelle que soit la description $d \in \mathcal{R}_i$ associée à sa connaissance, on a $d \in \llbracket \varphi \rrbracket_i$. Ainsi, on définit pour le cas présent :

$$\vdash \varphi \Leftrightarrow \forall i \in \{A, B\}, \llbracket \varphi \rrbracket_i = \mathcal{R}_i$$

À partir de ces définitions, nous pouvons tout d'abord énoncer deux propriétés simples qui nous serviront par la suite. Tout d'abord, une propriété de la forme $\varphi \rightarrow \psi$ vérifie $\vdash \varphi \rightarrow \psi$ si et seulement si on a : $\forall i, \llbracket \varphi \rrbracket_i \subseteq \llbracket \psi \rrbracket_i$. De plus, on a $\vdash \neg \varphi$ si et seulement si $\forall i, \llbracket \varphi \rrbracket_i = \emptyset$.

3.3 Quelques propriétés du système

Nous allons pouvoir utiliser le formalisme précédent pour vérifier formellement si des propriétés sont vérifiées par le système ou non. Commençons par une proposition proposée précédemment : $\neg(\text{AliceNoire} \wedge \text{AliceRouge})$.

Pour la représentation \mathcal{R}_A , on a $\llbracket \text{AliceNoire} \rrbracket_A = \{N, NB\}$ et $\llbracket \text{AliceRouge} \rrbracket_A = \{R, RB\}$. Ainsi, l'interprétation de la proposition précédente dans \mathcal{R}_A est l'ensemble vide. On vérifie facilement qu'il en est de même pour \mathcal{R}_B . Donc on a bien :

$$\vdash \neg(\text{AliceNoire} \wedge \text{AliceRouge})$$

Par contre, on montre que : $\not\vdash \text{AliceNoire} \vee \text{AliceRouge}$ puisque que par exemple :

$$A \notin \llbracket \text{AliceNoire} \vee \text{AliceRouge} \rrbracket_A$$

Cela mets en lumière un aspect intéressant de la modélisation de la connaissance dans ce contexte, puisqu'illustrant le fait que le tiers exclu (qui stipule que pour toute proposition φ , on a $\vdash \varphi \vee \neg\varphi$) n'est pas vérifié. La logique classique ne s'applique donc pas à notre approche, et il faut utiliser la logique intuitionniste à la place. Par contre, on peut avoir une version plus faible de cette proposition :

$$\vdash \neg(\neg AliceNoire \wedge \neg AliceRouge)$$

Cette dernière peut s'écrire de façon équivalente $\vdash \neg\neg(AliceNoire \vee AliceRouge)$. On voit ici apparaître la double négation dont le rôle sera développé longuement plus tard dans ce mémoire.

Regardons maintenant la proposition $AliceNoire \wedge AliceOuvert \rightarrow K_A AliceNoire$. Vérifier que sa traduction vers \mathcal{R}_A est égal à l'ensemble en entier se fait directement, puisque $\llbracket K_A AliceOuvert \rrbracket_A = \llbracket AliceOuvert \rrbracket_A$. Qu'en est-il pour \mathcal{R}_B . Puisque l'on a défini $\llbracket AliceNoire \rrbracket_A$ comme étant $\{N, NB\}$, on a :

$$\llbracket K_A AliceNoire \rrbracket_B = \{d \mid f_{A|B}(d) \in \{N, NB\}\} = \{RA\}$$

Ayant par ailleurs $\llbracket AliceOuvert \rrbracket_B = \{A, RA, NA\}$ et $\llbracket AliceNoire \rrbracket_B = \{R, RA\}$, on a bien l'inclusion de $\llbracket AliceOuvert \wedge AliceNoire \rrbracket_B$ dans $\llbracket K_A AliceNoire \rrbracket_B$. On a donc vérifié formellement que l'on a :

$$\vdash AliceNoire \wedge AliceOuvert \rightarrow K_A AliceNoire$$

3.4 Étude statique et dynamique

Pour finir cet exemple d'utilisation des structures de représentations que nous développerons par la suite, nous allons donner quelques indications pour l'étude du comportement dynamique d'un tel système. Ce que nous qualifions de "dynamique" ici consiste en les assertions de la forme "si à un moment donné φ est vrai, alors peut-on en déduire que ψ l'est aussi ?".

Dans ce qui précède, les propositions φ telles que $\vdash \varphi$ sont, d'après la définition de \vdash que l'on a donnée, les propositions qui sont vérifiées par le système quel que soit son état, son évolution. Dans le cas de l'expérience des deux cartes par exemple, cela ne dépend pas de la carte mise dans l'enveloppe d'Alice.

Comment faire alors pour trouver les propriétés vérifiées par l'expérience dans le cas où c'est la carte rouge qui est dans l'enveloppe d'Alice ? Une première idée pourrait être de s'autoriser à utiliser une hypothèse supplémentaire, soit $\vdash AliceRouge$ par exemple. Malheureusement, cette méthode ne marche pas. En effet, avec la méthode de traduction que l'on a donné, une propriété est que si l'on a $\vdash \varphi$, cela entraîne $\vdash K_A \varphi$ (et $\vdash K_B \varphi$). On donne ainsi à notre hypothèse

le statut de propriété *toujours* vérifiée par le système, ce que l'on ne veut pas. Ici, ce n'est pas parce que l'on considère que l'enveloppe d'Alice contient la carte rouge que forcément Alice et Bernard le savent. C'est pourtant le cas si l'on suppose $\vdash \text{AliceRouge}$.

Une autre méthode, plus convaincante, est de traduire une interrogation de la forme “*si à un moment donné φ est vrai, alors peut-on en déduire que ψ l'est aussi ?*” par la prouvabilité de la proposition $\varphi \rightarrow \psi$. Ainsi, *a-t'on $\vdash \varphi \rightarrow \psi$?* devient la question à considérer. Considérons par exemple l'assertion *si Alice sait que son enveloppe contient la carte rouge, alors elle sait que si Bernard ouvre son enveloppe, il saura aussi que celle d'Alice contient la carte rouge*. Ainsi, peut-on relier $K_A \text{AliceRouge}$ à l'implication $K_A (\text{BernardOuvert} \rightarrow K_B \text{AliceRouge})$? La réponse est positive, puisque l'on a :

$$\vdash K_A \text{AliceRouge} \rightarrow K_A (\text{BernardOuvert} \rightarrow K_B \text{AliceRouge})$$

On peut remarquer que la première méthode qui consistait à supposer l'hypothèse supplémentaire $\vdash K_A \text{AliceRouge}$ (et donc restreindre l'ensemble des descriptions à considérer) pour déduire $\vdash K_A (\text{BernardOuvert} \rightarrow K_B \text{AliceRouge})$ est trop forte, puisque l'on peut même déduire dans ce cas $\vdash K_B \text{AliceRouge}$, ce qui signifierait que Bernard gagne automatiquement de l'information lorsqu'Alice ouvre sa carte.

On voit donc que même si notre formalisme ne s'occupe que de propriétés vérifiées tout le temps par le système étudié (ce que l'on peut qualifier de propriétés statiques ou structurelles), il permet néanmoins d'appréhender de façon puissante le comportement dynamique de celui-ci, puisque les propriétés du type *ceci entraîne cela* peut se traduire sous forme d'une implication correspondant à une propriété statique.

Nous allons maintenant aborder la présentation du formalisme proprement dit, avec en premier lieu la définition précise des structures de représentations et de systèmes de représentations que nous utiliserons par la suite.

Deuxième partie

Formalisation

Dans cette partie, nous allons introduire les structures élémentaires qui seront utilisées comme base pour modéliser des systèmes. Dans un premier temps, nous définirons la structure de *représentation*. Il s'agit d'une sorte d'ensemble partiellement ordonné, dont les éléments correspondent à des descriptions de l'état du système étudié. Une représentation est utilisée pour la formalisation d'un système depuis un unique point d'observation. Aussi, dans ce cas, les seuls éléments qui entrent en compte sont d'une part la structure que l'on connaît du système, laquelle fournit un ensemble d'états et de descriptions possibles, et d'autre part une notion d'ordre partiel sur ces descriptions permettant de les comparer suivant la quantité d'informations qu'elles fournissent. L'usage de ce type de structure très simple pour modéliser la connaissance a pour première implication que la logique de base est dans ce contexte la logique intuitionniste et non la logique classique. Cela traduit le fait que dans notre approche, seules les observations fournissent des informations, et l'absence d'observation d'une propriété ne peut être utilisée pour prouver que celle-ci est fausse.

Une fois cette structure définie, nous allons l'étendre pour considérer l'existence de plusieurs points d'observations possibles pour un système donné. Pour cela, nous allons d'une part utiliser plusieurs représentations – une par point de vue – et d'autre part définir un ensemble de fonctions permettant de relier ces représentations entre elles pour refléter le fait qu'il s'agit d'un unique système qui est vu de multiples façons. Ces fonctions précisent la façon dont “coule” l'information entre les différents points d'observation, pour reprendre une terminologie issue de [22, 3]. C'est ainsi que nous définirons la structure de *système de représentations*, qui nous permettra d'étudier formellement comment il est possible de relier les connaissances d'un même système selon plusieurs points de vue. En particulier, nous nous intéresserons à la façon dont une information obtenue depuis un premier point de vue peut être utilisée pour un second point de vue. Cela nous permettra de préciser le mode d'étude de l'observation d'un système en faisant abstraction de celui-ci pour se concentrer sur les relations entre les différentes observations.

Les principaux éléments de cette partie sont présentés plus succinctement dans [13].

Chapitre 4

Une représentation

4.1 Présentation

Nous allons dans un premier temps nous intéresser à la notion de représentation. Une représentation est la formalisation d'une façon de représenter les différents états du système que l'on observe. Une façon de représenter un de ces états est d'en donner une description. Aussi, dans un premier temps, on peut considérer qu'une représentation d'un système est un ensemble de descriptions. Dans cette définition, on ne spécifie pas le fait qu'une description doit être complète, ni même aussi précise que possible. Ainsi, partant d'une première description d , il est possible d'en obtenir d'autres en supprimant de d certaines informations. Mais les descriptions ainsi obtenues sont moins précises que d . Cela permet d'introduire une notion que l'on peut ajouter de façon naturelle à un ensemble de descriptions, celle de comparaison suivant la précision et la quantité d'informations fournie par les descriptions. Cette comparaison correspond à une relation d'ordre partiel entre les descriptions, et nous permet d'obtenir notre première définition :

Definition 1 (Représentation)

Une représentation est une paire $\langle R, \leq \rangle$ où R est un ensemble de descriptions et \leq est un ordre partiel sur les descriptions.

Cette définition est en fait exactement celle d'un *e.p.o.* (ensemble partiellement ordonné). Cependant, le nom de représentation permet de préciser le fait que les éléments du poset que l'on considère sont des descriptions d'un système. De plus, nous utiliserons la convention que pour deux descriptions comparables $d_1 \leq d_2$, la plus précise des deux est d_1 .

Grâce à cette définition très simple et très générale, on peut définir une grande variété de représentations.

4.1.1 Exemples

Resultat d'un dé

Un premier exemple peut être donné par le cas simple où l'on veut représenter (éventuellement de façon partielle) le résultat d'un lancer de dé. Les différents états possibles sont donc les nombres compris entre 1 et 6. On peut dans ce cas utiliser comme description partielle de l'état du dé tout sous-ensemble non vide de $\{1, \dots, 6\}$, à la condition que l'un de ces nombres est le résultat effectif du dé.

L'ensemble des descriptions à considérer est donc $\wp(\{1, \dots, 6\}) \setminus \{\emptyset\}$. Pour obtenir une représentation, il faut de plus équiper cet ensemble d'un ordre partiel. Mais supposons que l'on a deux descriptions $d_1 = \{1, 2, 4\}$ et $d_2 = \{2, 4\}$. Comme une description est vue comme un ensemble de résultats possibles du dé lancé, on peut affirmer dans ce cas que d_2 est une description plus précise que d_1 , puisque l'on gagne l'information que le résultat n'est pas 1. Ainsi, il paraît clair dans ce cas que l'ordre à utiliser sur cet ensemble est l'inclusion, soit si $d_1 \subseteq d_2$ alors d_1 est plus précise que d_2 .

La représentation utilisée dans ce cas-là est donc de la forme $\langle \wp(\{1, \dots, 6\}) \setminus \{\emptyset\}, \subseteq \rangle$.

Parties d'un ensemble

De façon plus générale, étant donné un ensemble E , on peut considérer la représentation formée par ses sous-ensembles non vides $\langle \wp(E) \setminus \{\emptyset\}, \subseteq \rangle$. Une description peut être vue comme un ensemble d'états qui sont tous des candidats plausibles pour l'état réel du système. Le complément de la description est donc un ensemble d'états dont on sait qu'ils ne peuvent pas être l'état du système.

On remarque à nouveau que dans ce cas, on ne considère pas l'ensemble vide \emptyset comme une description, puisque l'on considère qu'il existe toujours un état (soit un élément x de E) qui correspond effectivement à l'état du système, et que les descriptions partielles de cet état doivent au moins contenir cet élément x .

Il est possible de considérer une multitude de variations de cette définition, en prenant des sous-ensembles de l'ensemble des parties et en tenant compte de la structure de l'ensemble E de départ. Par exemple, suivant les cas, on peut considérer l'ensemble des ouverts, des parties connexes, des ensembles finis, des sous-espaces d'un espace vectoriel, etc. La liste est très longue.

Ensemble de propositions

Étant donné un langage logique \mathcal{L} et une relation de déduction \vdash , on peut définir un ordre partiel sur les sous-ensembles de \mathcal{L} clos par déduction en posant $\Gamma \leq \Delta \Leftrightarrow \forall \delta \in \Delta, \Gamma \vdash \delta$. On obtient alors une représentation en considérant les sous-ensembles consistants de \mathcal{L} (c'est-à-dire ne permettant pas de prouver la proposition fautive \perp) ordonnés par \vdash .

Cette sorte de représentation est en fait très générale, puisque permet de comprendre une description comme un ensemble de faits vérifiés par l'état du système. De plus, elle nous sera très utile par la suite pour les preuves de théorèmes de complétude.

On peut remarquer que l'ordre défini ici correspond à l'inverse de l'inclusion ensembliste. On retrouve alors exactement le dual de la représentation de la forme ensemble de parties, puisque l'on peut associer une proposition à l'ensemble des états du système vérifiant cette proposition, et si une proposition en implique une seconde, elle est inversement vérifiée par moins d'états du système. C'est la raison pour laquelle il y a un renversement de l'ordre entre les deux types de représentations.

4.1.2 Enrichissement de la définition

On définit une représentation comme n'étant rien de plus qu'un e.p.o., la seule notion importante étant la comparaison entre les descriptions suivant leur précision et l'information fournie. En général, il est intéressant d'enrichir la structure d'e.p.o. en définissant des opérations supplémentaires tenant compte de l'ordre. Les opérations les plus simples de ce genre sont l'union \vee et l'intersection \wedge (ou pour prendre un vocabulaire plus général et plus adapté aux e.p.o., les bornes inférieures et supérieures), ainsi que les constantes représentant l'élément maximum \top et minimum \perp pour l'ordre partiel. Nous allons ici explorer l'application de ces opérations aux représentations.

Description vide

Commençons par l'adjonction d'un élément maximum pour l'ordre. On l'appelle traditionnellement *top* et on le note \top . En termes de descriptions, comment peut-on interpréter un tel élément? Étant maximal pour l'ordre \leq , on vérifie en d'autres termes : $\forall d, d \leq \top$. Suivant le rôle attribué à \leq , on en déduit que pour toute description d , l'information contenue dans la description maximale \top est présente dans d . \top peut donc être vue comme une description ne contenant que des informations communes à toutes les descriptions, en particulier, elle ne peut pas contenir d'information relative à l'état du système. En fait, on peut même aller plus loin et considérer que \top ne fournit absolument aucune information. En effet, si l'on veut décrire un système en termes d'une description appartenant à une représentation, la représentation elle-même

doit être prise en compte, et c'est la présence de cette représentation et de la structure qui va avec qui peut être considérée comme symbolisant l'information *a priori* concernant le système, présente quelque soit la description retenue. Ainsi, on peut considérer qu'alors \top ne fournit aucune information, l'information commune étant donnée par la représentation. Cette description correspond donc à l'absence totale d'information sur l'état du système, et sera parfois appelée *description vide* par la suite.

Conjonction de descriptions

Dans l'optique retenue pour la définition des représentations, une description peut être vue comme une liste de faits, d'informations que l'on tient pour vraies concernant l'état du système. Cette approche permet de facilement considérer une nouvelle opération que l'on peut intituler conjonction. En effet, si l'on a deux descriptions d_1 et d_2 décrivant l'état du système, on peut construire une nouvelle description correspondant à l'ensemble des faits fournis par d_1 et d_2 . Notons cette nouvelle description $d_1 \wedge d_2$, et étudions les propriétés qu'elle vérifie. Tout d'abord, cette description contient plus d'informations que les descriptions de départ d_1 et d_2 . On a donc $d_1 \wedge d_2 \leq d_1$ et $d_1 \wedge d_2 \leq d_2$. De plus, supposons l'existence d'une description d_3 plus précise que d_1 et que d_2 . Alors tous les faits fournis par ces deux descriptions sont aussi donnés par d_3 . Mais comme la description $d_1 \wedge d_2$ correspond exactement à l'ensemble des informations données par d_1 et d_2 , on en déduit que $d_3 \leq d_1 \wedge d_2$.

Il faut remarquer que cette opération est définie de façon partielle sur l'ensemble des descriptions. En effet, deux descriptions données peuvent très bien ne pas correspondre à un même état du système, et il se peut qu'aucun état ne corresponde aux deux descriptions. On parlera alors de descriptions incompatibles, et de descriptions compatibles sinon.

Exemple

Reprenons l'exemple du dé présenté plus tôt, et représenté par $\langle \wp(\{1, \dots, 6\}) \setminus \{\emptyset\}, \subseteq \rangle$. Dans ce cas, l'opération de conjonction correspond exactement à l'intersection ensembliste. En effet, si $d_1 = \{1, 2, 4\}$ et $d_2 = \{2, 3, 4, 5\}$, alors la conjonction des deux descriptions nous informe que le résultat du lancer de dé est soit 2, soit 4, ce qui est donné par $d_1 \wedge d_2 = \{2, 4\}$. On peut de plus considérer une troisième description $d_3 = \{5, 6\}$. Les descriptions d_1 et d_3 sont incompatibles, puisque leur intersection est vide, alors qu'une description doit toujours être un sous-ensemble non vide de $\{1 \dots 6\}$, puisque contenant le résultat du dé.

Exemple

Dans le cadre du représentation dont les descriptions sont des sous-ensembles consistents de propositions d'un langage \mathcal{L} , l'opération de conjonction se traduit par la réunion en-

sembliste. En effet, si une description est une liste de faits vérifiés par le systèmes, la conjonction de deux descriptions revient à considérer une liste plus longue.

On retrouve bien ici le renversement d'ordre par rapport à une représentation du type "mondes possibles", puisque l'on interprète ici la conjonction par la réunion et non l'intersection d'ensembles. De plus, là encore l'opération est partielle, puisque la réunion de deux ensembles de propositions doit être consistante.

Ainsi, le domaine de définition de l'opération de conjonction est l'ensemble des paires compatibles. Si l'on note $d_1 \frown d_2$ pour indiquer que deux descriptions sont compatibles, on a :

$$d_1 \frown d_2 \Leftrightarrow \exists d_3 : d_3 \leq d_1 \text{ et } d_3 \leq d_2$$

$$d_1 \frown d_2 \Rightarrow \forall d_3, (d_3 \leq d_1 \text{ et } d_3 \leq d_2) \Leftrightarrow d_3 \leq d_1 \wedge d_2$$

Si de plus, on a un élément maximum \top pour cet ordre, on a $\forall d, d \frown \top$ et la conjonction vérifie $d \wedge \top = d$. Enfin, suivant la façon dont le domaine de la conjonction est défini (c'est-à-dire en enlevant les paires incompatibles qui conduiraient à des états absurdes), on a la propriété suivante :

$$\forall d_1 \leq d_2, \forall e_1 \leq e_2, d_1 \frown e_1 \Rightarrow d_2 \frown e_2$$

Par la suite, on pourra pour cette opération parler de conjonction ou bien d'intersection, cette dernière appellation venant du fait que si l'on associe à une description l'ensemble des états possibles correspondants (ce qui était le cas pour l'exemple du dé), alors les états correspondant à la conjonction de deux descriptions sont donnés par l'intersection des ensembles d'états associés à chacune des descriptions de départ.

Ces deux opérations, la conjonction et la constante maximum, peuvent être, comme on l'a vu, définies assez naturellement en termes de représentations. Cela vient du fait que l'on peut voir la construction d'une description comme l'activité où, partant des informations a priori sur le système, on ajoute différentes informations une à une pour arriver à la description désirée. Leur présence n'étant pas nécessaire dans l'étude que nous voulons faire des structures basées sur les représentations, nous ne les supposerons pas dans les définitions de cette partie. Cependant, ces opérations sont tellement naturelles qu'elles seront souvent présente dans les exemples que l'on pourra donner. Ce n'est que dans la prochaine partie, lors de la définition de la notion de hiérarchie de représentations, que l'on imposera explicitement la présence de ces opérations.

Nous allons considérer maintenant les deux autres opérations, la réunion \vee et la constante minimale \perp , et voir qu'elles sont plus délicates à considérer dans le cadre des représentations.

Commençons par la constante minimale, appelée généralement *bottom* et notée \perp , description duale de \top présentée plus haut. Par définition, elle vérifie $\forall d, \perp \leq d$, ce qui signifie que toutes les descriptions que l'on peut donner de l'état du système sont moins précises que cette description \perp . Une telle description affaiblirait trop la notion de représentation. Pour montrer cela, revenons sur l'un des postulats de base, qui est que le système étudié est toujours dans un état bien défini. Introduisons de plus la notion de couverture d'une représentation \mathcal{R} : une couverture est un ensemble de descriptions $C \subseteq \mathcal{R}$ telle que $\forall d \in \mathcal{R}, \exists c \in C : d \frown c$. Le postulat peut être traduit par le fait qu'étant donné une couverture C , quelque soit l'état du système, il existe un élément $c \in C$ décrivant partiellement cet état. Or, s'il existe un minimum global \perp , le singleton $\{\perp\}$ constitue une couverture de la représentation, et donc \perp , qui est la plus précise des descriptions disponibles, ne permet pas de différencier les différents états possibles du système. La représentation n'a donc absolument aucun intérêt, puisque aucune des descriptions ne fournit effectivement d'information sur le système. Elles sont toutes équivalentes à la description vide, ne donnant que des informations a priori.

Aussi, même si l'argumentation que l'on vient de donner n'est qu'approximative, on comprend aisément que la présence d'un minimum global conduit au fait que toutes les descriptions sont équivalentes à une description vide. Par la suite, nous ne rencontrerons donc jamais de représentation avec un minimum global.

La situation est moins nette pour le cas de la disjonction \vee (aussi appelée réunion), comme nous l'avons évoqué dans la présentation des postulats de départ. Tout d'abord, dans de nombreux cas, il sera possible de définir une opération de disjonction sur la représentation. En effet, une classe intéressante de représentation concerne celles comportant une description vide, une opération d'intersection et vérifiant la condition de chaîne croissante. Cette condition stipule que toute description est à une distance finie du sommet \top , ou plus précisément qu'il existe une borne à la cardinalité des chaînes strictement croissantes allant de d vers \top . Cette classe de représentations correspond au cas où une description est obtenue à partir de la description vide en ajoutant des informations, et ce un nombre fini de fois. On peut alors définir une disjonction en posant :

$$d_1 \vee d_2 = \bigwedge \{d \mid d_1 \leq d \text{ et } d_2 \leq d\}$$

Il s'agit alors bien d'une fonction totale, la condition de chaîne croissante nous assurant en effet que la conjonction est bien définie même si l'ensemble apparaissant dans la définition est infini puisque, l'élément final étant à une distance finie du sommet, seul un nombre fini d'éléments de l'ensemble intervient dans le calcul effectif de la conjonction. On notera que le cas où la représentation contient un nombre fini de descriptions est inclus dans le cas précédent.

Nous voyons que dans ce cas spécifique, il est possible de définir une opération de disjonction en fonction de la conjonction. Cependant, dans le cas général, la définition ci-dessus fait

intervenir la conjonction d'un ensemble infini, qui n'est pas toujours définie. Or, comme nous l'avons expliqué plus tôt, la conjonction semble se définir très naturellement dans le cadre des représentations et en se basant sur la définition de la disjonction en fonction de la conjonction que l'on vient de voir, nous pensons que la disjonction découle effectivement de cette dernière, mais que cette dépendance ne permet pas toujours de la définir. Aussi, nous pensons que c'est la conjonction qu'il faut considérer en priorité, et que la disjonction, si elle peut être définie, en dépend.

Une autre approche pour expliquer cette préférence, est que si l'on se place dans le cadre de l'observation d'un système, et que l'on part du principe que toute observation correspond à un élément de réalité de ce système, étant donné deux observations, on peut *a priori* définir une nouvelle observation correspondant à la conjonction des deux précédentes, et pas à la disjonction.

Il serait aussi possible de considérer d'autres opérations sur les e.p.o., comme par exemple la pseudo-complémentation (défini de la façon suivante : étant donné deux éléments d_1 et d_2 , le pseudo-complément $d_1 \rightarrow d_2$ est l'unique élément vérifiant $\forall d, d \wedge d_1 \leq d_2 \Leftrightarrow d \leq d_1 \rightarrow d_2$). Ces opérations sont trop spécifiques pour être définies sur les représentations, mais nous les retrouverons lors de l'étude logique des représentations.

Avant cela, nous allons d'abord présenter plusieurs méthodes générales pour combiner et transformer les représentations.

4.2 Construction de représentations

Pour construire une représentation associée à un système composé de plusieurs parties mises en relation, on peut procéder de la façon suivante : on construit tout d'abord une représentation correspondant à chacune des sous-parties, que l'on regroupe pour n'en former qu'une. Chaque partie est alors, du point de vue de la représentation obtenue, indépendante des autres. On applique une transformation qui réduit la taille de la représentation, afin de ne garder qu'une partie des descriptions pour exprimer les relations entre les différentes parties. Ainsi, si le système est composé de deux sous-parties A et B et qu'une représentation d_A de A correspond à un état qui implique un état de B décrit par d_B , alors l'opération de réduction supprimera toutes les descriptions donnant d_A mais pas d_B .

Cette méthode de construction peut donc s'effectuer à partir de deux opérations générales sur les représentations que nous allons définir maintenant.

Produit cartésien

Tout d'abord, intéressons-nous à la première construction, qui permet de regrouper ensemble des représentations correspondant intuitivement à des sous-parties d'un système. Ainsi, on se donne deux représentations \mathcal{R}_A et \mathcal{R}_B (où l'on note $\mathcal{R}_i = \langle R_i, \leq_i \rangle$), chacune correspondant à une sous-partie. Une description de l'état du système peut alors être la donnée d'une paire de descriptions $\langle d_A, d_B \rangle$, chacune correspondant à une sous-partie. On reconnaît la construction du produit cartésien pour les ensembles de descriptions, que l'on étend à la structure de représentation en considérant l'ordre produit. On obtient donc la représentation produit :

$$\mathcal{R}_A \otimes \mathcal{R}_B = \langle R_A \times R_B, \leq_{A \otimes B} \rangle$$

où l'ordre $\leq_{A \otimes B}$ est le produit point par point des deux ordres de départ (ainsi, la comparaison $\langle d_1, d_2 \rangle \leq_{A \otimes B} \langle e_1, e_2 \rangle$ est équivalente à la conjonction $d_1 \leq_A e_1$ et $d_2 \leq_B e_2$). Cette construction pouvant s'étendre à une famille quelconque de représentations, on peut définir dans le cas général $\prod \{ \mathcal{R}_i \mid i \in \mathcal{I} \}$.

Cette construction, très classique, permet de construire une représentation à partir d'une famille de représentations. Cependant, elle ne peut pas être utilisée toute seule. Dans le cas de produit de deux représentations, toute paire de descriptions est une description du produit, ce qui implique que des informations concernant l'une des parties n'entraînent aucune connaissance sur l'état de l'autre partie. Du point de vue de cette représentation produit, il n'existe aucune corrélation entre les différentes parties. Dans ce cas, l'étude de chacune des parties est indépendante des autres, et il n'est d'aucune utilité de les considérer ensemble. Pour remédier à cela et augmenter les informations exprimées par une représentation, nous allons maintenant présenter l'opération de raffinement.

Raffinage

Reprenons le cas où l'on a deux représentations \mathcal{R}_A et \mathcal{R}_B . Supposons que ces représentations correspondent à deux parties d'un système, et qu'il existe une relation entre l'état de la partie A et l'état de B . Pour simplifier, on pourra considérer que la relation en question permet, connaissant une description de l'état de A , de déduire des informations sur l'état de B . Cette relation peut s'exprimer de façon fonctionnelle, et ainsi si l'on a une description $d_A \in \mathcal{R}_A$, on en déduit de l'information sur B décrite par $f(d_A) \in \mathcal{R}_B$. Dans ces conditions, certaines descriptions de $\mathcal{R}_A \otimes \mathcal{R}_B$ peuvent être supprimées, précisément, les descriptions $\langle d_A, d_B \rangle$ telles que $d_B \not\leq f(d_A)$. En effet, si une part d'information fournie par $f(d_A)$ n'est pas présente dans d_B , cette description est vue comme trop "pauvre". En présence d'une opération de conjonction, on peut la transformer en une bonne représentation $\langle d_A, d_B \wedge f(d_A) \rangle$.

Dans le cas général (sans supposer une opération de conjonction), on peut donc définir l'opération de raffinement sur une représentation \mathcal{R} à partir d'une fonction $\phi : \mathcal{R} \mapsto \mathcal{R}$ en prenant comme nouvel ensemble de descriptions $\{d \mid d \leq \phi(d)\}$, et en restreignant l'ordre à cet ensemble. On notera pour simplifier $\phi(\mathcal{R})$ la représentation obtenue. Pour l'exemple précédent, on avait $\phi_f(\langle d_A, d_B \rangle) = \langle d_A, f(d_A) \rangle$. On notera que cette opération de raffinement est idempotente : on a $\phi \circ \phi(\mathcal{R}) = \phi(\mathcal{R})$.

Plusieurs renforcements de la définition peuvent être définis. En particulier, le raffinement que l'on vient d'introduire ne permet pas en général, étant donné une description $d \in \mathcal{R}$, de dire comment la transformer en une description de $\phi(\mathcal{R})$. Elle fournit uniquement un critère de sélection que l'on applique aux descriptions. Pourtant, l'opération de raffinement peut aussi être vue comme un ajout d'information structurelle à l'observation du système, et l'on pourrait trouver naturel de savoir comment transformer une description en intégrant l'information structurelle que l'on ajoute. On peut donc supposer que la fonction utilisée pour faire le raffinement permet en plus de transformer les descriptions. Cela revient à associer à toute description d une description $\phi(d) \in \phi(\mathcal{R})$. On peut facilement se convaincre qu'une telle fonction est un opérateur de clôture inférieure, c'est-à-dire qu'elle est croissante, intensive et idempotente : croissante car si l'on a deux descriptions comparables $d_1 \leq d_2$, les informations déduites de d_2 et de l'information structurelle ajoutée peuvent aussi s'obtenir à partir de d_1 ; intensive (c'est-à-dire $\forall d, \phi(d) \leq d$) car l'application de la fonction revient à ajouter de l'information ; idempotente (c'est-à-dire $\forall d, \phi \circ \phi(d) = \phi(d)$) enfin car appliquer la fonction revient à ajouter l'information structurelle, et le faire une deuxième fois n'apporte rien par rapport à la première fois. Dans ce cas, la représentation obtenue consiste en les descriptions fixées par ϕ .

Dans le cas du produit de deux représentations avec une opération de conjonction, une fonction $f : \mathcal{R}_A \rightarrow \mathcal{R}_B$ permet en fait de définir une telle opération de clôture inférieure comme on l'a vu plus haut, en posant $\phi_f : \langle d_A, d_B \rangle \mapsto \langle d_A, d_B \wedge f(d_A) \rangle$. Cette remarque permet d'observer le fait qu'une opération de clôture inférieure n'est pas forcément une fonction totale, puisque dans le cas ϕ_f , la conjonction n'est que partielle. Cela traduit le fait qu'en ajoutant de l'information structurelle, on peut découvrir que des descriptions jusqu'alors acceptables sont en fait absurdes. Cependant, la découverte de descriptions absurdes est la seule cause de partialité de la fonction, et de façon similaire à l'opération de conjonction, il est raisonnable d'exiger que le domaine de définition d'une telle fonction ϕ soit un idéal : $\forall d_1 \leq d_2, d_1 \in \text{dom}(\phi) \Rightarrow d_2 \in \text{dom}(\phi)$.

On peut s'interroger sur l'application successive de plusieurs raffinages, et de leur éventuelle commutativité. Notamment, comme on a défini deux sortes d'opérations de raffinement, on peut voir la façon dont elles diffèrent. Dans le premier cas, on n'utilise une fonction de raffinement ϕ

que comme un sélecteur de descriptions, pour transformer une représentation \mathcal{R} en l'ensemble $\phi(\mathcal{R}) = \{d \mid d \leq \phi(d)\}$. Dans ce cas, il y a toujours commutativité entre les différents raffinages, puisque on a en général :

$$\phi_1 \circ \phi_2(\mathcal{R}) = \{d \mid d \leq \phi_1(d) \text{ et } d \leq \phi_2(d)\} = \phi_2 \circ \phi_1(\mathcal{R})$$

Dans le cas de la définition étendue, la situation est cependant moins concluante. En effet, rien n'impose qu'après avoir appliqué ϕ_1 puis ϕ_2 , on doive obtenir une description qui est toujours point fixe de ϕ_1 .

Exemple

Soit \mathbb{N} l'ensemble des entiers naturels, ordonnée de façon usuelle. C'est bien une représentation. On définit maintenant deux fonctions sur les entiers :

$$f : x \mapsto 2 \left\lceil \frac{x}{2} \right\rceil - 1 \quad g : x \mapsto 2 \left\lfloor \frac{x}{2} \right\rfloor$$

où $\lceil n \rceil$ (resp. $\lfloor n \rfloor$) représente la partie entière par excès (resp. par défaut) de n . Il est facile de s'assurer que ces fonctions sont bien des opérateurs de clôture inférieure, f ayant pour points fixes les nombres pairs, et g et les nombres impairs. L'application des fonctions ne commute pas, puisqu'on a en particulier : $\forall x, f \circ g(x) = g \circ f(x) \pm 1$.

On l'a vu, le raffinement d'une représentation peut se traduire par l'ajout d'informations structurelles concernant le système étudié. Cette opération a donc deux actions sur une représentation : d'une part, elle sélectionne les descriptions qui ne sont pas absurdes, d'autre part, elle précise quelles sont les descriptions qui peuvent être considérées comme suffisamment riches, incluant l'information structurelle que l'on a ajoutée.

La définition de ces deux opérations permet donc bien d'exprimer la façon dont en général on peut définir l'étude d'un système à partir de ses sous-parties. En effet, on commence d'abord par étudier chacune de ces sous-parties (ce qui correspond à la définition des représentations \mathcal{R}_i), puis on considère le système dans son ensemble, en prenant en compte toutes les sous-parties (on effectue le produit des \mathcal{R}_i) et en exprimant de façon précise les corrélations qu'il peut y avoir entre ces parties (on applique alors un ensemble de raffinages correspondant à ces corrélations).

De cette manière, et en considérant la variété des représentations que l'on peut définir, on voit qu'il est possible d'exprimer un très grand nombre de systèmes sous forme de représentations, et ce de façon très riche grâce à l'opération de raffinement. Nous allons maintenant voir comment il est possible d'étudier le comportement d'un système que l'on a formalisé de la sorte.

4.3 Logique associée aux représentations

De manière générale, pour étudier de façon logique le comportement d'un système formalisé par une représentation, on se donne un ensemble dénombrable Ψ de propositions liées à ce système, propositions dites atomiques. À partir de cet ensemble, on va définir le langage \mathcal{L}_Ψ en considérant toutes les propositions obtenues à partir de Ψ en utilisant les opérateurs logiques usuels (soit la conjonction \wedge , la disjonction \vee , l'implication \rightarrow et la négation \neg , ainsi que la constante fausse \perp). Pour cette définition, les opérateurs s'appliquent à des propositions et ne doivent pas être confondus avec les opérateurs discutés lors de la définition des représentations.

À partir de ce langage, nous allons associer aux propositions des ensembles de descriptions de la représentation retenue, et cela nous permettra de caractériser la façon dont se comportent les propositions en fonction de cette association avec une représentation. En particulier, nous étudierons la partie de ce comportement qui ne dépend pas de la représentation choisie mais seulement du fait que l'on utilise une représentation.

Définition de la fonction d'interprétation

Ainsi, soit \mathcal{R} une représentation. On va définir une fonction dite d'interprétation notée $\llbracket \cdot \rrbracket$ qui va associer à chaque proposition $\varphi \in \mathcal{L}_\Psi$ un ensemble de descriptions de \mathcal{R} . Le sens que l'on peut donner à cette fonction est le suivant : pour une proposition φ donnée, $\llbracket \varphi \rrbracket$ est exactement l'ensemble des descriptions $d \in \mathcal{R}$ pour lesquelles on peut s'assurer que φ est vraie (nous utilisons ici le terme *s'assurer* volontairement, pour ne pas préciser s'il s'agit de savoir, croire, prouver ou autre).

Suivant cette définition, on peut donner une première caractéristique de la fonction d'interprétation. Si l'on a deux descriptions $d, d' \in \mathcal{R}$ telles que d' est plus précise que d (c'est-à-dire $d' \leq d$), alors si d suffit pour être sûr que φ est vraie, alors d' , étant plus précise, suffit aussi. Ainsi, $\llbracket \varphi \rrbracket$ est un idéal de \mathcal{R} , c'est-à-dire un ensemble clos vers le bas. Si l'on note $\wp^\downarrow(\mathcal{P})$ l'ensemble des idéaux d'un ensemble partiellement ordonné (ou *epo*) \mathcal{P} , alors la fonction $\llbracket \cdot \rrbracket$ est à valeurs dans $\wp^\downarrow(\mathcal{R})$. Sachant cela, nous allons définir cette fonction de façon inductive, suivant la structure du terme donné en argument.

Tout d'abord, il faut préciser l'interprétation des propositions atomiques. Pour cela, la seule façon est de préciser "à la main" les descriptions correspondant à chaque proposition atomique. On a donc besoin d'une fonction $\nu : \Psi \rightarrow \wp^\downarrow(\mathcal{R})$. Le cas de la disjonction se traite simplement. Ainsi, supposons que $d \in \llbracket \varphi \vee \psi \rrbracket$. Cela signifie que la description d permet de s'assurer que φ ou que ψ est vraie. Ainsi, on a $\llbracket \varphi \vee \psi \rrbracket = \llbracket \varphi \rrbracket \cup \llbracket \psi \rrbracket$. De façon similaire, on trouve que $\llbracket \varphi \wedge \psi \rrbracket = \llbracket \varphi \rrbracket \cap \llbracket \psi \rrbracket$.

Le cas de l'implication est un peu plus complexe. En effet, on pourrait à nouveau définir

l'interprétation de $\varphi \rightarrow \psi$ en s'inspirant de la sémantique ensembliste usuelle : $\llbracket \varphi \rightarrow \psi \rrbracket = \{d \mid d \in \llbracket \varphi \rrbracket \Rightarrow d \in \llbracket \psi \rrbracket\}$. Cependant, en général, il ne s'agit pas d'un idéal. Nous devons donc restreindre cette définition en prenant le plus grand idéal inclus dans cet ensemble. On a donc :

$$\llbracket \varphi \rightarrow \psi \rrbracket = \{d \mid \forall d' \leq d, d' \in \llbracket \varphi \rrbracket \Rightarrow d' \in \llbracket \psi \rrbracket\}$$

De même, on obtient pour la négation $\llbracket \neg \varphi \rrbracket = \{d \mid \forall d' \leq d, d' \notin \llbracket \varphi \rrbracket\}$. On complète la définition en interprétant la constante fausse \perp par l'ensemble vide. Les règles de définition de $\llbracket \cdot \rrbracket$ sont données dans la table 3, où l'on note $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu}$ pour préciser que cette fonction est définie à valeurs dans $\wp^\downarrow(\mathcal{R})$, en utilisant une fonction d'interprétation atomique $\nu : \Psi \rightarrow \wp^\downarrow(\mathcal{R})$.

$$\begin{aligned} \forall \psi \in \Psi, \llbracket \psi \rrbracket_{\mathcal{R}, \nu} &= \nu(\psi) \\ \llbracket \varphi \vee \psi \rrbracket_{\mathcal{R}, \nu} &= \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} \cup \llbracket \psi \rrbracket_{\mathcal{R}, \nu} \\ \llbracket \varphi \wedge \psi \rrbracket_{\mathcal{R}, \nu} &= \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} \cap \llbracket \psi \rrbracket_{\mathcal{R}, \nu} \\ \llbracket \varphi \rightarrow \psi \rrbracket_{\mathcal{R}, \nu} &= \{d \mid \forall d' \leq d, d' \in \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} \Rightarrow d' \in \llbracket \psi \rrbracket_{\mathcal{R}, \nu}\} \\ \llbracket \neg \varphi \rrbracket_{\mathcal{R}, \nu} &= \{d \mid \forall d' \leq d, d' \notin \llbracket \varphi \rrbracket_{\mathcal{R}, \nu}\} \\ \llbracket \perp \rrbracket_{\mathcal{R}, \nu} &= \emptyset \end{aligned}$$

Table 3 – Définition de $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu} : \mathcal{L}_\Psi \rightarrow \wp^\downarrow(\mathcal{R})$

Notions sémantiques de vérité

En utilisant ces fonctions d'interprétation, on peut définir plusieurs notions de “vérité”. Tout d'abord, étant donné une représentation \mathcal{R} , une fonction d'interprétation atomique ν et une description $d \in \mathcal{R}$, on dit que d valide φ pour $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu}$ si $d \in \llbracket \varphi \rrbracket_{\mathcal{R}, \nu}$ (dit autrement, une description d valide une proposition si intuitivement, elle permet de s'assurer que cette proposition est vraie). Au niveau des fonctions d'interprétations, φ est valide pour $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu}$ ($\mathcal{R}, \nu \models_R \varphi$) si toutes les descriptions de \mathcal{R} valident cette proposition : $\forall d \in \mathcal{R}, d \in \llbracket \varphi \rrbracket_{\mathcal{R}, \nu}$ (c'est-à-dire $\llbracket \varphi \rrbracket_{\mathcal{R}, \nu} = \mathcal{R}$). Enfin, de façon générale, une proposition φ est valide pour les représentations si quelque soit la fonction d'interprétation considérée, la proposition est validée par celle-ci : $\models_R \varphi$ si et seulement si $\forall \mathcal{R}, \nu, \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} = \mathcal{R}$.

Cette notion de vérité définie de façon purement sémantique à partir des représentations est, comme on va le voir, équivalente à la prouvabilité pour la logique intuitionniste.

La logique intuitionniste

Cette logique, développée initialement par Brouwer puis par Heyting[41] dans les années 1930, repose sur la notion de constructibilité en mathématiques. Ainsi, un objet dont on peut prouver l'existence pour cette logique peut être construit explicitement, ce qui fait que l'on peut toujours exhiber les objets dont on parle. Cela interdit en particulier les raisonnements par l'absurde.

Dans le cas de l'interprétation en fonction des représentations, le constructivisme est incarné en particulier par la notion d'observation. L'observation d'un fait est considérée comme une preuve de la véracité de celui-ci, et toutes les propositions vraies sont alors des combinaisons de propositions primitives qui correspondent à des observations. La règle du tiers-exclu, qui est à l'origine du raisonnement par l'absurde et s'écrit $\varphi \vee \neg\varphi$, doit alors être abandonnée.

Exemple

Si l'on considère une pièce que l'on lance dans un jeu de pile ou face, tant que l'on n'a pas observé la pièce après le lancer, on ne peut pas dire quel est le résultat. Aussi, on ne peut pas considérer la proposition $\text{Pile} \vee \text{Face}$ comme étant vraie, ne pouvant pas affirmer que Pile est vraie, ni que Face l'est.

Cependant, une version équivalente du tiers-exclu signifie que la double négation est équivalente à l'identité ($\neg\neg\varphi \rightarrow \varphi$). Donc si l'on abandonne la règle du tiers-exclu, on abandonne aussi cette seconde règle. La double-négation acquiert, comme on le verra plus tard, un rôle très particulier.

Exemple

Dans le cas du jeu de pile ou face, elle permet de considérer la propositions $\neg\neg(\text{Pile} \vee \text{Face})$, qui exprime le fait que même si l'on ne regarde pas le résultat du lancer, on peut affirmer que le résultat ne peut pas être autre chose que pile ou face.

La logique intuitionniste repose principalement sur une notion très forte de l'implication, qui est ici un connecteur de base, et non une abbréviation (en logique classique, $\varphi \rightarrow \psi$ est un raccourci de $\neg\varphi \vee \psi$). Cette spécificité implique que pour la partie propositionnelle de cette logique, ce que l'on considère ici, on ne peut pas utiliser de tables de vérités pour vérifier la validité d'une proposition, contrairement à la logique classique. La notion de preuve est ici plus importante que la notion de vérité, et on en donne une formulation sous forme du calcul des séquents à la Gentzen dans la table 4.

$$\begin{array}{c}
\frac{}{\Gamma, \varphi \vdash \varphi} \text{Id.} \quad \frac{}{\Gamma, \perp \vdash \varphi} \perp L \\
\\
\frac{\Gamma \vdash \varphi \quad \Delta, \varphi \vdash \psi}{\Gamma, \Delta \vdash \psi} \text{Cut} \quad \frac{\Gamma, \varphi, \varphi \vdash \psi}{\Gamma, \varphi \vdash \psi} \text{Contract.} \\
\\
\frac{\Gamma, \varphi \vdash \vartheta \quad \Gamma, \psi \vdash \vartheta}{\Gamma, \varphi \vee \psi \vdash \vartheta} \vee L \quad \frac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi} \vee R_1 \quad \frac{\Gamma \vdash \psi}{\Gamma \vdash \varphi \vee \psi} \vee R_2 \\
\\
\frac{\Gamma, \varphi \vdash \vartheta}{\Gamma, \varphi \wedge \psi \vdash \vartheta} \wedge L_1 \quad \frac{\Gamma, \psi \vdash \vartheta}{\Gamma, \varphi \wedge \psi \vdash \vartheta} \wedge L_2 \quad \frac{\Gamma \vdash \varphi \quad \Gamma \vdash \psi}{\Gamma \vdash \varphi \wedge \psi} \wedge R \\
\\
\frac{\Gamma \vdash \varphi \quad \Gamma, \psi \vdash \vartheta}{\Gamma, \varphi \rightarrow \psi \vdash \vartheta} \rightarrow L \quad \frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \rightarrow \psi} \rightarrow R
\end{array}$$

Table 4 – Logique intuitionniste formulée en calcul des séquents

Équivalence entre \vdash_{IL} et $\models_{\mathcal{R}}$

Le modèle usuel pour la logique intuitionniste est la catégorie des algèbres de Heyting complètes, c'est-à-dire la catégorie des treillis complets pseudo-complémentés. Le fait que l'ensemble des idéaux d'un poset est justement une algèbre de Heyting complète permet de se convaincre que l'interprétation à l'aide de représentations comme on vient de le définir réalise bien un modèle pour la logique intuitionniste. Nous allons donner cependant une preuve précise de ce résultat, car elle servira de prototype pour les preuves où l'on considérera aussi les connecteurs de modalité.

Théorème 1

Toute proposition $\varphi \in \mathcal{L}_{\Psi}$ est prouvable avec la logique intuitionniste si et seulement si elle est valide pour les représentations :

$$\forall \varphi, \vdash_{\text{IL}} \varphi \Leftrightarrow \models_{\mathcal{R}} \varphi$$

Preuve Succinctement, on montre que l'interprétation est correcte, puisque chaque axiome de la logique est vérifiée par les interprétations $\llbracket \cdot \rrbracket$. Réciproquement, on utilise la méthode classique consistant à définir une interprétation canonique $\llbracket \cdot \rrbracket_c$ pour laquelle les propositions vraies sont exactement celles prouvables par IL. La preuve complète est donnée en annexe B.1. \square

Dans la suite, nous allons enrichir l'expressivité de la formalisation de systèmes en donnant la possibilité d'avoir plusieurs points de vue, à quoi correspondront plusieurs représentations, et en donnant des méthodes pour relier ces différentes représentations entre elles. De même, on enrichira la logique en introduisant les connecteurs de modalité au langage de propositions, pour symboliser la présence des multiples représentations et dont le comportement dépendra de la façon dont les représentations sont reliées. Cependant, à chaque fois, la base sera l'usage de représentations, seul l'agencement de celles-ci changera, ce qui implique que les logiques que l'on trouvera seront toutes basées sur la logique intuitionniste. Les règles logiques que l'on ajoutera pour caractériser leurs relations ne concerneront que les opérateurs modaux.

Chapitre 5

Deux représentations

Dans un premier temps, pour introduire l'étude des relations entre différentes représentations, nous allons étudier le cas simple où l'on a seulement deux représentations. De cette manière, nous allons pouvoir introduire le formalisme utilisé pour relier les représentations ensemble. Ce formalisme va être défini comme un développement de l'idée utilisée pour le raffinement des représentations. Ainsi, en partant du produit cartésien $\mathcal{R}_A \otimes \mathcal{R}_B$, on avait raffiné ce produit en introduisant une fonction qui associe une description de \mathcal{R}_B à chaque description de \mathcal{R}_A .

Pour exprimer une relation dans chaque sens, on peut donc relier deux représentations en se donnant deux fonctions $f_{B|A} : \mathcal{R}_A \rightarrow \mathcal{R}_B$ et $f_{A|B} : \mathcal{R}_B \rightarrow \mathcal{R}_A$. L'interprétation de ces fonctions est, comme précédemment, qu'étant donnée une description d_A appartenant par exemple à \mathcal{R}_A , cela implique des informations sur l'autre description (ici, \mathcal{R}_B) décrites par $f_{B|A}(d_A)$. Avec cette interprétation, on peut exprimer quelques conditions qui doivent intuitivement être vérifiées par ces fonctions. Tout d'abord, ces fonctions doivent être croissantes, puisque si l'on a deux descriptions comparables $d_1 \leq d_2$, les informations fournies sur l'autre représentation par d_2 sont aussi fournies par d_1 . Si l'on prend une description $d_A \in \mathcal{R}_A$, la représentation obtenue par transformation en appliquant $f_{B|A}$ fournit une partie des informations présentes dans d , mais une partie seulement. Puisque l'on perd de l'information à nouveau en retransformant la description obtenue vers \mathcal{R}_A , on en déduit que la description finale $f_{A|B} \circ f_{B|A}(d_A)$ est moins précise que d_A , soit exprimé autrement $\forall d_A \in \mathcal{R}_A, d_A \leq f_{A|B} \circ f_{B|A}(d_A)$. Le raisonnement symétrique à partir de \mathcal{R}_B permet aussi de conclure que $\forall d_B \in \mathcal{R}_B, d_B \leq f_{B|A} \circ f_{A|B}(d_B)$.

Nous allons donc partir de cette situation pour étudier les relations entre représentations : on se donne deux représentations \mathcal{R}_A et \mathcal{R}_B et deux fonctions $f_{B|A}$ et $f_{A|B}$ vérifiant les conditions précédentes. Pour simplifier les notations, nous supposons aussi la présence de deux fonctions $f_{A|A}$ et $f_{B|B}$ définies comme l'identité sur \mathcal{R}_A et \mathcal{R}_B respectivement. Les conditions sur les

fonctions autres que la croissance peuvent donc s'écrire :

$$\forall i, j, k \in \{A, B\}, f_{k|i} \leq f_{k|j} \circ f_{j|i}$$

Avec cela, nous allons étudier différentes situations de relations entre deux représentations. En particulier, on peut étudier deux cas extrêmes de relation entre deux représentations. Il s'agit tout d'abord d'une relation que l'on peut qualifier d'"horizontale", où les deux représentations sont totalement déconnectées l'une de l'autre. Aucune information ne circule alors entre les deux, et les descriptions de l'une des représentations ne donne aucune information exprimables dans l'autre représentation. À l'opposé, deux représentations peuvent être en relation "verticale", auquel cas l'une des représentations apparaît comme une approximation de l'autre. Avant cela, nous allons adapter deux constructions ensemblistes classiques – la somme disjointe et le produit – aux représentations en tenant compte des fonctions les reliant.

5.1 Constructions à partir de deux représentations

5.1.1 Somme disjointe

Nous allons tout d'abord définir la notion de somme disjointe pour les représentations. Pour définir cette construction, il faut être capable de comparer toute paire de descriptions, quelle que soit leurs représentations d'origine. Si les descriptions proviennent de la même représentation, on peut les comparer en utilisant l'ordre correspondant à cette représentation. Aussi, pour que cet ordre soit défini dans tous les cas, il faut maintenant s'intéresser au cas où l'on compare les descriptions provenant des deux représentations. Soit donc deux descriptions $d_A \in \mathcal{R}_A$ et $d_B \in \mathcal{R}_B$. En utilisant la fonction de transformation $f_{B|A}$, on déduit de d_A la description $f_{B|A}(d_A)$ appartenant à \mathcal{R}_B . Elle représente la description la plus précise dans cette représentation basée sur les informations de d_A . Aussi, d_B est moins précise que d_A si et seulement si elle est moins précise que $f_{B|A}(d_A)$ dans \mathcal{R}_B . En considérant les autres cas de provenance pour les distributions, et en utilisant les fonctions $f_{A|A}$ et $f_{B|B}$, on obtient la définition de la relation $\leq_{A \oplus B}$ définie sur $\mathcal{R}_A \oplus \mathcal{R}_B$:

$$\forall i, j \in \{A, B\}, \forall d \in \mathcal{R}_i, \forall e \in \mathcal{R}_j, d \leq_{A \oplus B} e \Leftrightarrow f_{j|i}(d) \leq_j e$$

Grâce aux inégalités concernant la composition de ces fonctions, on prouve sans mal que la relation $\leq_{A \oplus B}$ ainsi définie est bien un pré-ordre. Pour ce pré-ordre, deux éléments différents d_1 et d_2 sont équivalents si et seulement si ils n'appartiennent pas à la même représentation (puisque les relations associées aux représentations sont des ordres partiels et non des pré-ordres), et vérifient, si $d_1 \in \mathcal{R}_A$ et $d_2 \in \mathcal{R}_B$, $d_A = f_{A|B}(d_B)$ et $d_B = f_{B|A}(d_A)$. Si l'on note $\simeq_{A \oplus B}$

cette relation d'équivalence, on peut alors définir la construction de somme disjointe sur les représentations :

$$\mathcal{R}_A \oplus \mathcal{R}_B = \left\langle (R_A \oplus R_B)_{/\simeq_{A \oplus B}}, \leq_{A \oplus B} \right\rangle$$

Cette construction permet donc d'utiliser conjointement les descriptions provenant de différentes représentations. On identifie les descriptions qui se transforment l'une en l'autre. De plus, l'ordre que l'on a défini permet de comparer le contenu informationnel de différentes descriptions.

Cependant, cette construction n'augmente pas le nombre ni la précision des descriptions, puisqu'à chaque fois, on se donne une unique description du système. Dans la construction suivante, on augmentera la puissance descriptive en s'autorisant à utiliser simultanément une description de chaque représentation.

5.1.2 Produit réduit

Dans les constructions appliquées aux représentations seules, on avait d'une part défini le produit cartésien de deux représentations, et le raffinage où l'on réduit le nombre des descriptions d'une représentation en les sélectionnant suivant un critère d'"intérêt". Nous allons ici combiner ces deux opérations en une seule pour obtenir le produit réduit de deux représentations. Pour cela, nous allons utiliser les fonctions $f_{i|j}$ pour effectuer dans la pratique la réduction du produit cartésien de \mathcal{R}_A et \mathcal{R}_B . Le critère de sélection des descriptions va être le suivant : si l'on considère la paire formée de deux description $\langle d_A, d_B \rangle$, on veut s'assurer qu'elle vérifie deux propriétés. D'une part, que les deux descriptions sont compatibles, c'est-à-dire qu'elles peuvent s'appliquer au même état du système (au moins dans l'état des connaissances du moment), et qu'elles sont maximales, c'est-à-dire que chaque description contient l'information donnée par l'autre description et exprimable dans la représentation. On remarque que cette seconde propriété implique la première. On peut donc définir le produit réduit à partir de l'ensemble de descriptions suivant :

$$\{ \langle d_A, d_B \rangle \mid d_A \leq f_{A|B}(d_B) \text{ et } d_B \leq f_{B|A}(d_A) \}$$

Cette construction n'est qu'un cas particulier de l'application du produit et du raffinage sur les représentations. En effet, le produit réduit ainsi défini est exactement la représentation $\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B)$ où l'on a défini :

$$f : \langle d_A, d_B \rangle \mapsto \langle f_{A|B}(d_B), f_{B|A}(d_A) \rangle$$

Dans cette construction, il est possible de retrouver chacune des représentations initiales. De façon plus précise, pour chaque description d'une des représentations de départ (prenons par

exemple $d_A \in \mathcal{R}_A$), il existe une description du produit réduit $\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B)$ fournissant exactement la même quantité d'information. En effet, considérons la description $\langle d_A, f_{B|A}(d_A) \rangle$. Elle appartient au produit réduit, puisque l'on a :

$$f_{B|A}(d_A) \leq f_{B|A}(d_A) \quad d_A \leq f_{A|B} \circ f_{B|A}(d_A)$$

De plus, l'inégalité de gauche qui est en fait une égalité prouve qu'il s'agit de la plus petite description de $\phi_f(\mathcal{R}_A \times \mathcal{R}_B)$ contenant les informations fournies par d_A . Or, puisque elle n'est définie qu'à partir de d_A , on comprends qu'elle ne contient pas plus d'information, si ce n'est l'information structurelle ajoutée par le biais des $f_{i|j}$.

Ces considérations nous montrent qu'il existe pour chaque description de chaque représentation une description du produit réduit décrivant le même état et étant minimale. Cela n'était pas vrai pour un produit cartésien dans le cas général, il fallait pour cela la présence de descriptions minimales (les descriptions vides \top) pour chacune des représentations de départ. Cette relation entre un produit réduit et l'une des représentations de départ est en fait plus forte que cela, et nous verrons que la seconde est une approximation de la première, une fois que l'on aura défini cette notion.

5.2 Cas extrêmes de relations entre représentations

Nous allons maintenant étudier deux cas de relations extrêmes entre représentations et voir la façon dont cela se traduit pour la somme disjointe et le produit réduit.

5.2.1 Cas horizontal

Commençons tout d'abord par ce que l'on a défini comme une relation d'horizontalité entre les deux représentations \mathcal{R}_A et \mathcal{R}_B . Rappelons que cela signifie qu'une description d'une des représentations ne fournit absolument aucune information exprimable dans l'autre représentation. Cette définition implique une contrainte assez forte sur chaque description, qui est la présence d'une description minimale. En effet, la condition précédente peut se traduire, pour le sens de A vers B par exemple, par l'inégalité suivante :

$$\forall d_A \in \mathcal{R}_A, \forall d_B \in \mathcal{R}_B, d_B \leq_B f_{B|A}(d_A)$$

Si d_A ne contient aucune information exprimable dans \mathcal{R}_B , cela signifie que $f_{B|A}(d_A)$ ne fournit aucune information, et l'on a donc $\forall d_B, d_B \leq f_{B|A}(d_A)$. Or, $f_{B|A}(d_A)$ est bien une description de \mathcal{R}_B , et l'inégalité que l'on vient de considérer nous montre qu'il s'agit de la description vide.

Ainsi, on vient de montrer que dans le cas relation horizontale entre représentations, chacune des deux représentations de départ possède une description vide, et chaque fonction de transformation $f_{A|B}$ et $f_{B|A}$ envoie les descriptions de la représentation de départ vers la description vide de l'autre transformation. Partant de cela, nous allons pouvoir caractériser la somme disjointe et le produit réduit de ces représentations.

Pour la somme disjointe, l'ordre $\leq_{A \oplus B}$ vérifie pour les descriptions provenant de représentations différentes :

$$\forall d_A \in \mathcal{R}_A, \forall d_B \in \mathcal{R}_B, \begin{cases} d_A \leq_{A \oplus B} d_B \Leftrightarrow d_B = \top_B \\ d_B \leq_{A \oplus B} d_A \Leftrightarrow d_A = \top_A \end{cases}$$

On en déduit que les seuls éléments équivalents pour $\simeq_{A \oplus B}$ sont les descriptions vides \top_A et \top_B (en dehors de la réflexivité), et donc la somme disjointe $\mathcal{R}_A \oplus \mathcal{R}_B$ est simplement la juxtaposition des deux représentations de départ où l'on a identifié les descriptions vides.

Pour le produit réduit, la fonction de sélection ϕ_f envoie toutes les paires de descriptions sur la paire $\langle \top_A, \top_B \rangle$, ce qui fait qu'il n'y a aucune réduction qui s'applique effectivement. Le produit réduit est alors exactement le produit cartésien des deux représentations (on a $\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B) = \mathcal{R}_A \otimes \mathcal{R}_B$).

On voit bien que dans le cas d'une relation horizontale entre deux représentations, aucune information ne circule entre les deux représentations, il n'y a aucune relation entre les descriptions de représentations différentes et les deux systèmes correspondant peuvent s'étudier séparément.

5.2.2 Cas vertical

Nous allons maintenant nous intéresser à l'autre cas extrême de relation entre deux représentations. Cette fois, il s'agit de verticalité, où l'on considère qu'une des représentations (nous supposons qu'il s'agit de \mathcal{R}_A) est exactement "sous" l'autre (ici, il s'agira de \mathcal{R}_B). Intuitivement, dans la relation d'ordre partiel utilisée pour comparer les descriptions d'une même représentation, une description est "sous" une autre si elle contient plus d'information que celle-ci. C'est cette idée que nous allons adopter pour les représentations. On veut signifier que \mathcal{R}_A est plus précise que \mathcal{R}_B . Cela peut se traduire par le fait que \mathcal{R}_A a une plus grande expressivité, qu'elle peut donner des descriptions avec une meilleure précision, etc. Mais en considérant de plus qu'une représentation peut contenir pour chaque description les descriptions partielles où l'on oublie des informations, cela revient à dire que chaque description de \mathcal{R}_B peut être exactement exprimée dans \mathcal{R}_A . Dans ce cas, on dira réciproquement que \mathcal{R}_B est une *approximation* de \mathcal{R}_A .

Mais la donnée des fonctions $f_{i|j}$ (où $i, j \in \{A, B\}$) nous fournit pour chaque description $d_B \in \mathcal{R}_B$ le candidat la représentant dans \mathcal{R}_A , il s'agit de $f_{A|B}(d_B)$. Si cette dernière description représente fidèlement d_B , cela signifie que l'on retrouve celle-ci en retraduisant la description obtenue dans \mathcal{R}_B . En d'autres termes, la relation de verticalité entre \mathcal{R}_A et \mathcal{R}_B peut s'exprimer de la façon suivante :

$$\forall d_B \in \mathcal{R}_B, d_B = f_{B|A} \circ f_{A|B}(d_B)$$

C'est cette condition que nous allons considérer pour étudier la somme disjointe et le produit réduit dans ce cas.

Pour la somme disjointe, on en déduit que chaque description d_B de \mathcal{R}_B est équivalente à $f_{A|B}(d_B)$ pour la relation $\simeq_{A \oplus B}$. Si l'on quotiente $\mathcal{R}_A \oplus \mathcal{R}_B$ par $\simeq_{A \oplus B}$ on retombe sur une représentation isomorphe à \mathcal{R}_A . Ainsi, si la représentation \mathcal{R}_A est plus précise que \mathcal{R}_B au sens que l'on vient de définir, on a $\mathcal{R}_A \oplus \mathcal{R}_B = \mathcal{R}_A$.

On va avoir une situation analogue pour le produit réduit. En effet, soit une paire de descriptions $\langle d_A, d_B \rangle$ appartenant au produit réduit $\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B)$. Par définition de celui-ci, les descriptions d_A et d_B vérifient $d_A \leq f_{A|B}(d_B)$ et $d_B \leq f_{B|A}(d_A)$. Mais si l'on applique $f_{B|A}$ à chaque membre de la première inégalité, on en déduit grâce à la croissance de la fonction, $f_{B|A}(d_A) \leq f_{B|A} \circ f_{A|B}(d_B) = d_B$. En particulier, cela implique que d_B est entièrement déterminé par d_A , puisque l'on a $d_B = f_{B|A}(d_A)$. Le produit réduit $\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B)$ est donc lui aussi isomorphe à \mathcal{R}_A , puisque l'on a, en oubliant l'ordre partiel sur les descriptions :

$$\phi_f(\mathcal{R}_A \otimes \mathcal{R}_B) = \{ \langle d_A, f_{B|A}(d_A) \rangle \mid d_A \in \mathcal{R}_A \}$$

Les constructions précédentes nous montrent donc que lorsque une représentation est l'approximation d'une autre, la première se comporte comme si elle était incluse dans la seconde. Dès lors, l'étude du système représenté peut uniquement se faire par le biais de la représentation la plus précise.

Cependant, la notion d'approximation, telle que nous l'avons définie, est très intéressante en tant que telle et, dans la prochaine section, nous allons l'étudier d'une manière plus générale.

5.3 Approximation et abstraction

Comme on l'a vu au chapitre 2.1, les notions très proches d'approximation et d'abstraction sont des sujets d'étude très développés dans de nombreux domaines des mathématiques et de l'informatique théorique. L'idée de base dans tous les cas est alors généralement la même : on se donne deux ensembles ainsi qu'une relation entre ces deux ensembles, permettant de

mettre en correspondance les éléments des différents ensembles. Alors, l'étude de l'approximation consiste en la détermination des informations que l'on a sur un ou plusieurs éléments de l'ensemble de départ en fonction des éléments du second ensemble qui leur sont reliés.

Dans le court survol que nous avons fait, nous avons présenté différentes façons de formaliser l'idée d'approximation. Tout d'abord, intuitivement, on peut vouloir dire que B est un sous-ensemble bien choisi de A . Réciproquement, chaque élément de B est accessible depuis A . Dans le cadre de départ, avec un espace de Chu, cela peut donc se traduire par le fait que tout élément de B est en relation avec au moins un élément de A : $\forall b, \exists a : a\mathcal{R}b$. En terme de fonctions \triangleleft et \triangleright utilisées en analyse de concept, cela revient à imposer $\triangleright\triangleleft B = B$, et pour une correspondance de Galois (semi-duale ou non), $\alpha \circ \gamma(x) = x$. On parle dans le dernier cas de surjections de Galois.

On voit donc que dans ces différents formalismes, la notion d'approximation s'exprime assez naturellement avec des conditions simples, et nous allons voir que ce que précédemment nous avons définis comme une approximation dans le cadre de la relation entre deux représentations correspond bien à cette notion dans l'acception que l'on vient de donner.

5.3.1 Approximation et représentations

On définit une approximation d'une représentation par une autre comme le cas où l'on a deux représentations \mathcal{R}_A et \mathcal{R}_B et des fonctions croissantes $\{f_{i,j} \mid i, j \in \{A, B\}\}$ vérifiant en particulier :

$$\begin{aligned} \forall d_A \in \mathcal{R}_A, d_A &\leq f_{A|B} \circ f_{B|A}(d_A) \\ \forall d_B \in \mathcal{R}_B, d_B &= f_{B|A} \circ f_{A|B}(d_B) \end{aligned}$$

Soient dans ce cas deux descriptions $d_A \in \mathcal{R}_A$ et $d_B \in \mathcal{R}_B$. Si $d_A \leq f_{A|B}(d_B)$, on en déduit $f_{B|A}(d_A) \leq f_{B|A} \circ f_{A|B}(d_B) = d_B$. Réciproquement, si $f_{B|A}(d_A) \leq d_B$, l'application de $f_{A|B}$ nous donne $d_A \leq f_{A|B} \circ f_{B|A}(d_A) \leq f_{A|B}(d_B)$. On a donc montré que les fonctions $f_{A|B}$ et $f_{B|A}$ forment une correspondance de Galois semi-duale :

$$\forall d_A \in \mathcal{R}_A, \forall d_B \in \mathcal{R}_B, f_{B|A}(d_A) \leq d_B \Leftrightarrow d_A \leq f_{A|B}(d_B)$$

Il s'agit en fait d'une surjection de Galois semi-duale, puisque $f_{B|A}$ est une surjection (ayant $f_{B|A} \circ f_{A|B} = \text{id}_B$).

On retrouve donc bien pour les représentations la construction d'approximation qui se présente dans le cas général. Cela n'est pas étonnant, puisque les fonctions de transformations entre les deux représentations correspondent à l'idée d'une relation entre celles-ci. Dans la suite, nous

allons reprendre l'étude logique des représentations, en nous intéressant au cas de l'approximation. De par la généralité de cette construction, l'étude que nous allons faire pourra donc s'adapter à un grand nombre de situations et de formalisations de la notion d'approximation.

5.3.2 Logique de l'approximation

Au chapitre précédent, nous avons étudié le comportement logique associé aux représentations, et nous avons montré que l'on retrouvait exactement la logique intuitionniste. Maintenant, dans le cas d'une approximation, nous allons considérer deux représentations, ainsi que les relations imposées entre elles. Mais avant cela, nous allons présenter une autre façon d'exprimer l'approximation d'une représentation par une autre.

Dans l'idée même d'approximation, on a vu que l'ensemble abstrait (celui qui "approxime") peut être vu comme inclus dans l'ensemble concret (celui qui est "approximé"). Aussi, on peut penser qu'il est possible de ne considérer qu'un ensemble, en précisant comment transformer les éléments concrets en éléments correspondant à l'ensemble abstrait. Dans le cas de l'approximation d'une représentation par une autre, une description concrète d_A est transformée en description abstraite $f_{B|A}(d_A)$. Or, celle-ci est exactement équivalente à la description concrète $f_{A|B} \circ f_{B|A}(d_A)$. En effet, on a :

$$\begin{aligned} f_{B|A}(f_{A|B} \circ f_{B|A}(d_A)) &= (f_{B|A} \circ f_{A|B})(f_{B|A}(d_A)) = f_{B|A}(d_A) \\ f_{A|B}(f_{B|A}(d_A)) &= f_{A|B} \circ f_{B|A}(d_A) \end{aligned}$$

Cela qui signifie que les deux descriptions correspondent à exactement la même description du système. On va donc associer chaque description abstraite d_B à la description concrète $f_{A|B}(d_B)$, et nous allons définir la fonction $\rho = f_{A|B} \circ f_{B|A}$ pour représenter l'application de l'approximation. En effet, si l'on a $d_A \in \mathcal{R}_A$, $\rho(d_A)$ représente la description obtenue en transformant d_A dans \mathcal{R}_B (il s'agit alors de $f_{B|A}(d_A)$) mais exprimée dans \mathcal{R}_A en utilisant $f_{A|B}$. La fonction ρ ainsi définie est un opérateur de clôture supérieure, c'est-à-dire qu'elle vérifie :

$$\begin{aligned} \forall d, d &\leq \rho(d) \\ \forall d, \rho(d) &= \rho \circ \rho(d) \\ \forall d_1 \leq d_2, \rho(d_1) &\leq \rho(d_2) \end{aligned}$$

En particulier, l'idempotence de ρ (c'est-à-dire que $\rho \circ \rho = \rho$ vient du fait à nouveau que $f_{B|A} \circ f_{A|B}$ est l'identité sur \mathcal{R}_B .

La donnée d'une représentation et d'un opérateur de clôture supérieure est équivalente à la donnée de l'approximation d'une représentation \mathcal{R}_A par une autre \mathcal{R}_B par le biais des fonctions de transformation $f_{A|B}$ et $f_{B|A}$. En effet, comme on vient de le définir, on définit ρ en fonction

des $f_{i|j}$. Réciproquement, on définit $\mathcal{R}_B = \{d \in \mathcal{R}_A \mid d = \rho(d)\}$, $f_{B|A}$ est l'identité sur \mathcal{R}_A restreinte à \mathcal{R}_B , et $f_{A|B}$ est la fonction ρ .

Grâce à cette équivalence, on va pouvoir remplacer l'étude des approximations d'une représentation par une autre par l'étude des représentations munies d'un opérateur de clôture supérieure.

Pour étudier la logique associée à l'approximation de représentations, il faut de plus étendre le langage. Pour cela, nous allons ajouter un opérateur \Box , dont le rôle sera de personnifier le processus d'approximation. Ainsi $\Box\varphi$ signifie que l'on est sûr que φ est vraie, même après application de l'approximation. Nommons le $\mathcal{L}_{\Box, \Psi}$. Comme c'était le cas pour les représentations, nous allons définir une fonction d'interprétation de $\mathcal{L}_{\Box, \Psi}$ vers $\wp^\perp(\mathcal{R})$. Pour la définir, nous allons avoir besoin d'une fonction d'interprétation atomique ν comme précédemment, ainsi que de l'opérateur de clôture supérieure ρ pour exprimer l'opérateur \Box . On peut ainsi définir l'interprétation de $\Box\varphi$. Comme on l'a dit, $\llbracket \Box\varphi \rrbracket$ correspond à l'ensemble des descriptions qui, même après approximation, valident φ . Cela donne la définition suivante :

$$\llbracket \Box\varphi \rrbracket_{\mathcal{R}, \rho, \nu} = \{d \mid \rho(d) \in \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu}\}$$

On retrouve la définition complète de $\llbracket \cdot \rrbracket_{\mathcal{R}, \rho, \nu}$ dans la table 5 en ajoutant les définitions pour les autres opérateurs logiques.

$$\begin{aligned} \forall \psi \in \Psi, \llbracket \psi \rrbracket_{\mathcal{R}, \rho, \nu} &= \nu(\Psi) \\ \llbracket \varphi \vee \psi \rrbracket_{\mathcal{R}, \rho, \nu} &= \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu} \cup \llbracket \psi \rrbracket_{\mathcal{R}, \rho, \nu} \\ \llbracket \varphi \wedge \psi \rrbracket_{\mathcal{R}, \rho, \nu} &= \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu} \cap \llbracket \psi \rrbracket_{\mathcal{R}, \rho, \nu} \\ \llbracket \varphi \rightarrow \psi \rrbracket_{\mathcal{R}, \rho, \nu} &= \{d \mid \forall d' \leq d, d' \in \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu} \Rightarrow d' \in \llbracket \psi \rrbracket_{\mathcal{R}, \rho, \nu}\} \\ \llbracket \neg\varphi \rrbracket_{\mathcal{R}, \rho, \nu} &= \{d \mid \forall d' \leq d, d' \notin \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu}\} \\ \llbracket \perp \rrbracket_{\mathcal{R}, \rho, \nu} &= \emptyset \\ \llbracket \Box\varphi \rrbracket_{\mathcal{R}, \rho, \nu} &= \{d \mid \rho(d) \in \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu}\} \end{aligned}$$

Table 5 – Définition de $\llbracket \cdot \rrbracket_{\mathcal{R}, \rho, \nu} : \mathcal{L}_{\Box, \Psi} \rightarrow \wp^\perp(\mathcal{R})$

Caractérisation de la logique

Comme précédemment, nous allons définir une notion de vérité, avec le symbole $\models_{\mathcal{AR}}$ défini ainsi : $\models_{\mathcal{AR}} \varphi \Leftrightarrow \forall \mathcal{R}, \rho, \nu, \llbracket \varphi \rrbracket_{\mathcal{R}, \rho, \nu} = \mathcal{R}$. Si l'on ne tient pas compte de l'opérateur \Box , on a montré que la logique obtenue est exactement la logique intuitionniste. L'ajout de cet opérateur

va nous obliger à ajouter quelques axiomes pour caractériser son comportement. Nous allons montrer qu'il se comporte exactement comme un opérateur de modalité vérifiant **S4 + KV**. On rappelle les axiomes et règles correspondants dans la liste suivante :

$\Box(\varphi \rightarrow \psi) \rightarrow \Box\varphi \rightarrow \Box\psi$	K
$\Box\varphi \rightarrow \varphi$	T
$\Box\varphi \rightarrow \Box\Box\varphi$	4
$\Box(\varphi \vee \psi) \rightarrow \Box\varphi \vee \Box\psi$	KV
$\frac{\vdash \varphi}{\vdash \Box\varphi}$	Nec

On peut les relier précisément aux propriétés de la définition d'un opérateur de clôture supérieure. Ainsi, la règle de nécessité vient du fait que ρ est une fonction totale, et que l'image réciproque de \mathcal{R} est l'ensemble lui-même. L'axiome de Kripke **K** exprime la croissance de ρ . L'axiome **T** correspond à l'extensivité de ρ (i.e. $d \leq \rho(d)$), puisque $d \in \llbracket \Box\varphi \rrbracket \Rightarrow \rho(d) \in \llbracket \varphi \rrbracket \Rightarrow d \in \llbracket \varphi \rrbracket$. L'axiome **4** correspond lui à l'idempotence de ρ . On a même, en l'utilisant conjointement à l'axiome **T**, $\Box\varphi \leftrightarrow \Box\Box\varphi$ qui est la conséquence au niveau de l'interprétation de l'égalité $\llbracket \Box\Box\varphi \rrbracket = \llbracket \Box\varphi \rrbracket$. Enfin, l'axiome **KV** vient de l'interprétation sous forme d'ensembles.

De façon plus précise, on a le théorème suivant :

Théorème 2

*Toute proposition $\varphi \in \mathcal{L}_{\Box, \Psi}$ est prouvable pour la logique **IS4 + KV** si et seulement si elle est valide pour les approximations de représentations :*

$$\forall \varphi \in \mathcal{L}_{\Box, \Psi}, \vdash_{\mathbf{IS4+KV}} \varphi \Leftrightarrow \models_{\mathcal{AR}} \varphi$$

Preuve La preuve complète se trouve en annexe B.2. □

Ainsi, dans le cas de deux représentations, une situation particulière où l'une des représentations est l'approximation de l'autre donne un comportement logique très proche de **IS4**. Nous allons maintenant nous tourner vers des cas où le nombre de représentations n'est pas limité. Il n'y a alors pas forcément de représentation plus précise que toutes les autres, et nous verrons alors qu'il faudra affaiblir cette logique pour rendre compte du comportement général d'un tel système.

Chapitre 6

Plusieurs représentations

Dans le chapitre précédent, nous avons étudié le cas où deux représentations étaient présentes et reliées entre elles par le biais de fonctions transformant les descriptions. En particulier, nous avons caractérisé le comportement logique d'un tel système dans le cas où l'une des représentations est l'approximation de l'autre.

Nous allons maintenant nous intéresser à une étude plus générale de cette catégorie de structures, composées de représentations reliées entre elles par des fonctions de transformation des descriptions. Pour cela et de manière à rester le plus général, nous allons définir un *système de représentations* d'une façon aussi simple que possible. Dans la version de base, la seule condition imposée sur les fonctions de transformation est la croissance. Pour reprendre l'argument donné au chapitre précédent, on justifie cela en présentant les fonctions de transformation comme sélectionnant une partie des informations présentes dans la description de base, et en les traduisant dans une autre représentation. Si maintenant nous avons deux descriptions comparables $d_1 \leq d_2$, les informations sélectionnées à partir de d_2 seront aussi présentes après sélection à partir de d_1 , donc après traduction vers la nouvelle représentation, les nouvelles descriptions obtenues sont elles aussi comparables. Nous imposons de plus que les fonctions de traduction allant d'une représentation vers elle-même soient l'identité, puisqu'avec l'interprétation que l'on en a donné, toutes les informations présentes dans une description peuvent subsister, puisqu'il n'y a pas de perte liée à la traduction. Ces considérations sur les fonctions de transformation ne s'appliquaient qu'au cas où il n'y avait que deux représentations, et nous allons maintenant nous en inspirer pour définir une structure similaire dans le cas où il y a un ensemble quelconque de représentations :

Definition 2

Un système de représentations est un triplet $\langle \mathcal{I}, \{ \mathcal{R}_i \mid i \in \mathcal{I} \}, \{ f_{i|j} \mid i, j \in \mathcal{I} \} \rangle$, où \mathcal{I} est un ensemble d'indices, pour chaque indice $i \in \mathcal{I}$, \mathcal{R}_i est une représentation et les fonctions $\{ f_{i|j} : \mathcal{R}_j \rightarrow \mathcal{R}_i \}$ définies entre chaque paire de représentations sont croissantes, avec en particulier $\forall i \in \mathcal{I}, f_{i|i} = \text{id}_i$.

À partir de cette définition, nous allons étudier divers raffinements possibles, et exprimer de façon logique les modifications qui en découlent dans le comportement de ces structures. Mais pour cela, nous allons tout d'abord adapter le langage logique à l'existence de multiples représentations, c'est-à-dire de plusieurs façons d'observer et d'acquérir des informations.

Modification des notations

Dans le cadre de l'approximation d'une représentation par une autre, on avait étendu le langage en ajoutant un opérateur \square symbolisant le processus d'approximation. Ainsi, on avait une représentation de base, ainsi qu'une autre, moins précise, et l'on utilisait cet opérateur \square pour parler de l'information accessible depuis cette représentation moins précise. Le point important à remarquer ici est que l'opérateur \square est associé à une représentation. C'est cette idée que nous allons développer dans le cadre général. Ainsi, nous allons définir le langage $\mathcal{L}_{\mathcal{I},\Psi}$ pour étudier de façon logique les systèmes de représentation avec un ensemble \mathcal{I} indexant les représentations. Ce langage est obtenu à partir de \mathcal{L}_{Ψ} par l'ajout d'un ensemble d'opérateurs unaires $\{K_i\}_{i \in \mathcal{I}}$, chacun de ces opérateurs K_i jouant vis-à-vis de la représentation correspondante \mathcal{R}_i le même rôle que \square vis-à-vis de la représentation approximante. Ils permettent donc d'exprimer la présence d'information accessible depuis un point de vue donné : $K_i \varphi$ indique que la propriété φ peut être vérifiée depuis le point de vue associé à la représentation \mathcal{R}_i .

On retrouve avec cette construction une interprétation des opérateurs de modalité similaire à celle de la logique épistémique. Dans le cadre de cette logique, on a un ensemble d'agents et les opérateurs de modalité K_i correspondent à la connaissance de l'agent correspondant. Ainsi, $K_i \varphi$ signifie alors que l'agent \mathcal{A}_i "sait" que φ est vraie.

6.1 Cas sans composition

Nous allons tout d'abord nous intéresser à la version brute de la définition d'un système de représentations que l'on vient de donner. Pour cette version, il n'y a aucune relation *a priori* entre les différentes fonctions de transformation. Nous allons cependant voir que l'interprétation des formules de $\mathcal{L}_{\mathcal{I},\Psi}$ vers un système de représentations va permettre d'identifier certains comportements des opérateurs K_i . Mais du fait que l'on ne peut pas relier des descriptions appartenant à des représentations différentes, il en résulte que ces comportements ne vont pas concerner les K_i pour des indices différents.

Cela implique que si l'on veut interpréter les formules de notre langage en termes de descriptions, il faut donner une interprétation pour chaque représentation. En reprenant la construction déjà utilisée dans les chapitres précédents, nous allons donc définir un ensemble de fonctions d'interprétation $\llbracket \cdot \rrbracket_i : \mathcal{L}_{\mathcal{I},\Psi} \rightarrow \wp^{\downarrow}(\mathcal{R}_i)$. Pour chacune de ces fonctions, il faut de plus se donner

une fonction d'interprétation atomique $\nu_i : \Psi \rightarrow \wp^\downarrow(\mathcal{R}_i)$. Cela nous permet de définir l'interprétation des formules de \mathcal{L}_Ψ vers les idéaux d'un système de représentations. Cependant, comme on a étendu le langage en ajoutant les opérateurs de modalité pour obtenir $\mathcal{L}_{\mathcal{I},\Psi}$, il nous faut maintenant définir l'interprétation d'une formule du type $K_i \varphi$. Pour cela, intéressons-nous à l'interprétation d'une formule de ce type. Aussi, soit $d \in \llbracket K_i \varphi \rrbracket_j$. Il s'agit d'une description de \mathcal{R}_j dont on sait que transformée en description de \mathcal{R}_i , elle valide φ . On peut donc en déduire que $\llbracket K_i \varphi \rrbracket_j$ est exactement l'ensemble des descriptions de \mathcal{R}_j qui, transformées vers \mathcal{R}_i par $f_{i|j}$, se trouvent dans $\llbracket \varphi \rrbracket_i$. On en déduit la définition générale de l'interprétation de $\mathcal{L}_{\mathcal{I},\Psi}$ vers un système de représentations $\mathcal{S} = \langle \mathcal{I}, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$ suivant un ensemble de fonctions d'interprétation atomique $\nu = \{\nu_i\}$, qui est donnée dans la table 6.

$$\begin{aligned}
\forall \psi \in \Psi, \llbracket \psi \rrbracket_{\mathcal{S},\nu,j} &= \nu_j(\Psi) \\
\llbracket \varphi \vee \psi \rrbracket_{\mathcal{S},\nu,j} &= \llbracket \varphi \rrbracket_{\mathcal{S},\nu,j} \cup \llbracket \psi \rrbracket_{\mathcal{S},\nu,j} \\
\llbracket \varphi \wedge \psi \rrbracket_{\mathcal{S},\nu,j} &= \llbracket \varphi \rrbracket_{\mathcal{S},\nu,j} \cap \llbracket \psi \rrbracket_{\mathcal{S},\nu,j} \\
\llbracket \varphi \rightarrow \psi \rrbracket_{\mathcal{S},\nu,j} &= \{d \mid \forall d' \leq_j d, d' \in \llbracket \varphi \rrbracket_{\mathcal{S},\nu,j} \Rightarrow d' \in \llbracket \psi \rrbracket_{\mathcal{S},\nu,j}\} \\
\llbracket \neg \varphi \rrbracket_{\mathcal{S},\nu,j} &= \{d \mid \forall d' \leq_j d, d' \notin \llbracket \varphi \rrbracket_{\mathcal{S},\nu,j}\} \\
\llbracket \perp \rrbracket_{\mathcal{S},\nu,j} &= \emptyset \\
\llbracket K_i \varphi \rrbracket_{\mathcal{S},\nu,j} &= \{d \mid f_{i|j}(d) \in \llbracket \varphi \rrbracket_{\mathcal{S},\nu,i}\}
\end{aligned}$$

Table 6 – Définition de $\{\llbracket \cdot \rrbracket_{\mathcal{S},\nu,i} : \mathcal{L}_{\mathcal{I},\Psi} \rightarrow \wp^\downarrow(\mathcal{R}_i)\}_{i \in \mathcal{I}}$

6.1.1 Caractérisation de l'interprétation

Cette interprétation va nous permettre de trouver quelques indications sur le comportement des opérateurs K_i . Tout d'abord, comme l'indépendance des différentes représentations le laisse présager, nous pouvons penser qu'il n'y a pas d'axiomes utilisant explicitement plusieurs K_i différents servant à modéliser leur comportement. Nous verrons par la suite que cette intuition est juste. Intéressons-nous donc aux axiomes standards de la logique modale, puisqu'il ne mettent pas en jeu plusieurs modalités différentes.

Tout d'abord, l'axiome **K** est bien vérifié, puisque partant de $(\varphi \rightarrow \psi) \wedge \varphi \rightarrow \psi$, on a pour une description $d \in \mathcal{R}_j$, si $d \in \llbracket K_i (\varphi \rightarrow \psi) \wedge K_i \varphi \rrbracket_j$, cela implique sur $f_{i|j}(d) \in \llbracket (\varphi \rightarrow \psi) \wedge \varphi \rrbracket_i$, ce qui implique que $f_{i|j}(d) \in \llbracket \psi \rrbracket_i$, soit de façon équivalente $d \in \llbracket K_i \psi \rrbracket_j$. Cet axiome est très important, puisqu'il exprime le fait que l'on peut faire des déductions localement, c'est-à-dire que l'ensemble des informations que l'on a à partir d'un point de vue donné est clos

par déduction. La règle de nécessité **Nec** est elle aussi vérifiée, puisque $\llbracket \varphi \rrbracket_i = \mathcal{R}_i$ implique que $\llbracket K_i \varphi \rrbracket_j = \mathcal{R}_j$. Enfin, l'axiome **KV** : $K_i(\varphi \vee \psi) \rightarrow K_i \varphi \vee K_i \psi$ est aussi vérifié, puisque l'on a $f_{i|j}(x) \in A \cup B \Leftrightarrow (f_{i|j}(x) \in A \text{ ou } f_{i|j}(x) \in B)$.

Ces trois axiomes et règles viennent du fait que l'on a défini l'interprétation d'une formule comme un ensemble de descriptions et ne sont que les conséquences de propriétés ensemblistes.

Nous allons maintenant nous pencher sur les deux familles principales d'axiomes de modalité, qui sont **D** et **T** d'une part, et **4** et **5** d'autre part, qui ne dépendent eux que de la façon dont sont reliées ensembles les différentes descriptions. Pour cela, et afin d'alléger l'interprétation des différents axiomes, nous allons introduire un ensemble d'agents $\{\mathcal{A}_i\}$, chaque agent correspondant à un point d'observation. Ainsi, dans la suite, la phrase "l'agent \mathcal{A}_i sait que φ est vraie" signifiera que l'on considère une description de \mathcal{R}_i validant φ (i.e. une description d telle que $d \in \llbracket \varphi \rrbracket_i$).

Les axiomes **D** et **T**

Étudions maintenant le couple d'axiomes **D** et **T**. On rappelle que le premier **D** : $K_i \varphi \rightarrow \neg K_i \neg \varphi$ exprime le fait que la connaissance obtenue depuis un point de vue donné est consistante. En effet, avec **K**, il est équivalent à $\neg K_i \perp$. Sous cette forme, il est facile de montrer qu'il est vérifié, puisque $\llbracket K_i \perp \rrbracket_j = \{d \mid f_{i|j}(d) \in \emptyset\} = \emptyset$.

Le second (**T** : $K_i \varphi \rightarrow \varphi$) est plus puissant, puisqu'il permet d'affirmer des choses sur le système lui-même, et non plus sur les observations que l'on peut en faire. Or, cela revient à faire des généralisations inconsidérées.

La raison majeure pour laquelle l'axiome **T** n'est pas vérifié est qu'il n'existe pas de contraintes sur les différentes interprétations d'une même proposition atomique. Ainsi, soit $\psi \in \Psi$ et définissons $\nu_1(\psi) = \mathcal{R}_1$ et $\nu_2(\psi) = \emptyset$. Un calcul simple montre que :

$$\begin{aligned} \llbracket K_1 \psi \rightarrow \psi \rrbracket_2 &= \{d \mid \forall d' \leq d, f_{1|2}(d') \in \llbracket \psi \rrbracket_1 \Rightarrow d' \in \llbracket \psi \rrbracket_2\} \\ &= \{d \mid \forall d' \leq d, f_{1|2}(d') \in \mathcal{R}_1 \Rightarrow d' \in \emptyset\} \\ &= \emptyset \end{aligned}$$

Mais si l'axiome **T** n'est pas vérifié, on peut constater qu'une version plus faible l'est. En effet, considérons l'axiome suivant, que nous appellerons **T₁** :

$$\mathbf{T}_1 : K_i K_i \varphi \rightarrow K_i \varphi$$

Il exprime le fait que si un agent sait qu'il sait qu'une proposition φ est vraie, alors il sait qu'elle est vraie. Cet axiome exprime précisément le fait que $\forall i, f_{i|i} = \text{id}_i$. Or, si cette condition a

été imposée dans la définition des systèmes de représentations, elle apparaît comme naturelle, et l'axiome \mathbf{T}_1 apparaît lui aussi comme naturel. En particulier, si l'on reprend l'idée de la présence d'agents aux différents points d'observations, comme on n'impose pas de restrictions à l'"intelligence" de celui-ci (que ce soit sa puissance déductive ou sa capacité de représentation), on ne peut pas faire a priori de distinction entre ce que sait un agent et ce qu'il sait qu'il sait. L'axiome \mathbf{T}_1 en est une facette, l'axiome 4 que l'on va voir maintenant en est une autre.

Les axiomes 4 et 5

On se retrouve avec ces axiomes dans une situation similaire aux axiomes précédents. Le premier est en effet vérifié alors que le second est délicat et non vérifié dans le cas général.

L'axiome 4 : $K_i \varphi \rightarrow K_i K_i \varphi$, tout d'abord, est souvent appelé *axiome d'introspection positive*. Il signifie qu'un agent sait les informations qu'il sait, il peut prouver qu'il peut prouver un fait. Combiné à l'axiome \mathbf{T}_1 , cela donne $K_i \varphi \leftrightarrow K_i K_i \varphi$, qui est exactement l'expression du fait que $f_{i|i}$ est l'identité de \mathcal{R}_i , ce qui revient aux égalités $\forall \varphi, \llbracket \varphi \rrbracket_i = \llbracket K_i \varphi \rrbracket_i$.

L'axiome 5 : $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$, est appelé *axiome d'introspection négative*. Il signifie qu'un agent peut prouver qu'il ne peut pas prouver une proposition donnée. Il s'agit aussi, comme \mathbf{T} , d'un axiome trop fort. Dans le chapitre précédent, on a vu qu'il n'était même pas vérifié dans le cas de l'approximation d'une représentation par une autre, cas pourtant très restrictif puisque l'on pouvait alors considérer que la représentation la plus précise correspondait à une description complète de la réalité, ou en tout cas suffisamment complète (ce qui fait que l'axiome \mathbf{T} était lui vérifié). Dans le cas général, la situation n'est pas meilleure.

Exemple

En reprenant l'exemple utilisé lors de l'étude de l'axiome \mathbf{T} , on peut s'intéresser à l'interprétation $\llbracket \neg K_2 \varphi \rightarrow K_2 \neg K_2 \varphi \rrbracket_1$. Pour cette interprétation, on a $\llbracket \neg K_2 \varphi \rrbracket_2 = \llbracket \neg \varphi \rrbracket_2 = \emptyset$ ce qui entraîne $\llbracket K_2 \neg K_2 \varphi \rrbracket_1 = \emptyset$. Cependant, dans la représentation \mathcal{R}_1 , on a $\llbracket K_2 \varphi \rrbracket_1 = \llbracket \varphi \rrbracket_1$, ce qui entraîne que $\llbracket \neg K_2 \varphi \rrbracket_1 \neq \emptyset$. Ainsi, la proposition :

$$\neg K_2 \varphi \rightarrow K_2 \neg K_2 \varphi$$

n'est pas valide.

Cet axiome n'est donc pas vérifié par les systèmes de représentations tels que nous venons de les définir. En fait, nous pensons que quelque soit l'extension de la définition que l'on peut considérer, cet axiome ne sera jamais vérifié. Cela vient de l'idée de baser la connaissance sur l'observation. Nous développerons ce point plus précisément dans la troisième partie.

Axiomes de localisation

Comme on l'a vu, l'étude du comportement des opérateurs K_i et de leur interprétation entraînent la vérification des axiomes 4 et \mathbf{T}_1 , du fait que l'on a $\llbracket \varphi \rrbracket_i = \llbracket K_i \varphi \rrbracket_i$. Les répercussions de cette égalité sur la logique sont en fait plus fortes, puisque l'on a alors $\llbracket \varphi \leftrightarrow K_i \varphi \rrbracket_i = \mathcal{R}_i$, et donc :

$$\forall \varphi, \forall i, j, \llbracket K_i (\varphi \leftrightarrow K_i \varphi) \rrbracket_j = \mathcal{R}_j$$

Ainsi, un axiome plus puissant que $K_i \varphi \leftrightarrow K_i K_i \varphi$ est vérifié, que l'on peut écrire de la façon suivante :

$$\mathbf{L} : K_i (K_i \varphi \leftrightarrow \varphi)$$

que l'on peut décomposer en les deux axiomes suivants :

$$\mathbf{L}_T : K_i (K_i \varphi \rightarrow \varphi) \quad \mathbf{L}_4 : K_i (\varphi \rightarrow K_i \varphi)$$

Nous avons intitulé cet axiome \mathbf{L} comme *localisation*, puisqu'avec l'interprétation telle qu'on l'a définie, un observateur a des informations sur le point de vue qu'il utilise et ces informations sont présentes de façon interne à la logique.

Règle d'universalité

Pour clore ce survol d'axiomes vérifiés par l'interprétation $\llbracket \cdot \rrbracket$, nous allons présenter un dernier élément, qui est une règle et non un axiome, et qui est aussi vérifié par cette interprétation. D'après sa définition, on a en effet : $\llbracket \varphi \rrbracket_i = \mathcal{R}_i$ qui implique $\forall j, \llbracket K_i \varphi \rrbracket_j = \mathcal{R}_j$. On peut même obtenir une équivalence en quantifiant universellement :

$$(\forall i, \llbracket \varphi \rrbracket_i = \mathcal{R}_i) \Leftrightarrow (\forall i, j, \llbracket K_i \varphi \rrbracket_j = \mathcal{R}_j)$$

On peut s'en inspirer pour définir ce que nous nommons la règle d'universalité, qui spécifie qu'une proposition est vraie si elle est vue comme vraie de partout. Plus formellement, on a :

$$\mathbf{Univ} : (\forall i, \vdash K_i \varphi) \Rightarrow \vdash \varphi$$

On vérifie sans peine que cette règle est vérifiée par $\llbracket \cdot \rrbracket$, d'après sa définition.

Cette règle permet de caractériser certaines relations existant entre les propriétés vérifiées par le système (de la forme $\vdash \psi$), et les propriétés considérées comme étant vérifiées par un observateur (de la forme $\vdash K_i \psi$). Ainsi, si \mathcal{I} est fini, cette propriété est équivalente à :

$$\vdash \left(\bigwedge_{i \in \mathcal{I}} K_i \varphi \right) \rightarrow \varphi$$

En effet, on a par $\mathbf{L}_T \forall i, \vdash K_i(K_i \varphi \rightarrow \varphi)$ ce qui entraîne $\forall i, \vdash K_i \left(\bigwedge_j K_j \varphi \rightarrow \varphi \right)$ dont on déduit la proposition précédente en appliquant **Univ**. Réciproquement, si l'on a $\forall i, \vdash K_i \varphi$, alors on en déduit $\vdash \bigwedge_i K_i \varphi$, et donc $\vdash \varphi$ avec $\vdash \left(\bigwedge_i K_i \varphi \right) \rightarrow \varphi$.

De façon duale, on a $\forall i, \vdash K_i(\varphi \rightarrow K_i \varphi)$ dont on déduit $\forall i, \vdash K_i(\varphi \rightarrow \bigvee_j K_j \varphi)$ puis $\vdash \varphi \rightarrow \bigvee_j K_j \varphi$ en appliquant **Univ** à nouveau. On voit donc que pour si une propriété est vérifiée par le système, alors au moins un des points d'observation permet de s'assurer qu'elle est vraie. Cette dernière propriété illustre bien la différence qui existe entre une propriété vérifiée par le système de façon globale (donc telle que $\vdash \varphi$) et une propriété vérifiée de façon dynamique (donc intervenant dans une expression de la forme $\vdash \varphi \rightarrow \dots$). Dans le premier cas, la règle de nécessité nous dit que si on a $\vdash \varphi$, donc si φ est vérifiée de façon globale, ou encore structurelle par le système, alors ce fait est accessible par tous les points d'observation. À l'inverse, si φ est vérifiée à un moment par le système, alors on peut seulement affirmer que ce fait est accessible par au moins l'un des points d'observation, ce qui est traduit par $\vdash \varphi \rightarrow \bigvee_i K_i \varphi$. De plus, ce dernier cas n'est vrai que si \mathcal{I} est fini, alors que la règle de nécessité s'applique même si l'ensemble des indices est infini.

Nous verrons d'ailleurs dans le chapitre suivant que si \mathcal{I} est infini, on peut supprimer la règle d'universalité sans appauvrir la logique.

6.1.2 Caractérisation de la logique

Nous venons de voir qu'avec une interprétation vers un système de représentation, la logique intuitionniste (puisque l'on est toujours basé sur une interprétation sous forme d'idéaux d'un poset), ainsi que les axiomes **K**, **D**, **L**, **KV** et les règles **Nec** et **Univ** sont vérifiés. Appelons la logique correspondante \mathbf{OL}^- comme *logique observationnelle* (observational logic), l'exposant $-$ indiquant que nous ajouterons un axiome dans la section suivante pour obtenir ce que nous considérons être une bonne définition de ce que peut être une logique de l'observation. Nous allons maintenant caractériser complètement la logique associée aux systèmes de représentations. Pour cela, nous définissons \models_S par $\models_S \varphi \Leftrightarrow \forall S, \nu, i, \llbracket \varphi \rrbracket_{S, \nu, i} = \mathcal{R}_i$ avec les notations usuelles.

Nous pouvons alors énoncer le théorème caractérisant précisément ce comportement.

Théorème 3

Toute proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$ est prouvable pour la logique \mathbf{OL}^- si et seulement si elle est valide pour les systèmes de représentations :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I}, \Psi}, \vdash_{\mathbf{OL}^-} \varphi \Leftrightarrow \models_S \varphi$$

Preuve La preuve complète se trouve en annexe B.3. □

Dans cette première version du système de représentation, il n'y avait pas de contraintes exprimant des relations entre les fonctions de transformation des descriptions. Nous allons maintenant étudier comment l'ajout de quelques contraintes va modifier la logique correspondante.

6.2 Cas avec composition

Pour étudier comment renforcer la définition d'un système de représentations en ajoutant des relations entre les fonctions de transformation, revenons sur leur rôle. Comme nous l'avons défini, on peut considérer que la transformée $f_{i|j}(d)$ d'une description $d \in \mathcal{R}_j$ contient moins d'informations que la description de départ. En effet, il est possible qu'une part des informations de d ne soit pas exprimable dans la nouvelle représentation \mathcal{R}_i . Mais puisque ces deux descriptions n'appartiennent pas à la même représentation, il est difficile de les comparer directement pour exprimer le fait que l'une est plus ou moins précise que l'autre.

Nous allons donc définir une condition traduisant le fait que d est plus précise que $f_{i|j}(d)$ en utilisant une troisième représentation \mathcal{R}_k et en transformant ces deux descriptions dans cette dernière. Il est alors possible de les comparer, et l'on veut alors écrire :

$$f_{k|j}(d) \leq f_{k|i}(f_{i|j}(d))$$

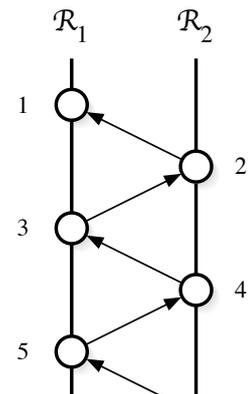
On obtient donc la règle de composition que l'on veut imposer aux systèmes de représentations pour exprimer le fait que le processus de traduction d'une représentation vers une autre entraîne une perte d'informations. En écrivant cette inégalité de façon générale, on a donc :

$$\forall i, j, k \in \mathcal{I}, \forall d \in \mathcal{R}_i, f_{k|i}(d) \leq_k f_{k|j} \circ f_{j|i}(d) \quad (6.1)$$

Un cas particulier de celle-ci survient lorsque $k = i$, cela donne $d \leq_i f_{i|j} \circ f_{j|i}(d)$, ce qui est naturel puisque comme on l'a vu, l'application d'une fonction de transformation perd de l'information et rend une description moins précise. Cependant, on pourrait s'attendre à voir vérifiée l'égalité :

$$f_{i|j} \circ f_{j|i} \circ f_{i|j} \circ f_{j|i}(d) = f_{i|j} \circ f_{j|i}(d)$$

En effet, la description obtenue pourrait correspondre intuitivement à la description dont les informations sont celles fournies par d qui peuvent être exprimées à la fois dans \mathcal{R}_i et \mathcal{R}_j . Ce n'est pas le cas, puisque cette interprétation est erronée. En effet, l'ensemble des informations de d exprimables dans \mathcal{R}_j ne correspond pas forcément à une description de \mathcal{R}_i , aussi lorsque l'on retransforme vers \mathcal{R}_i , on ne garde une fois de plus qu'une partie des informations présentes dans $f_{j|i}(d)$.



Exemple

On se donne deux représentations, \mathcal{R}_1 (resp. \mathcal{R}_2) correspondant à l'ensemble des nombres impairs (resp. pairs), ordonnés par l'ordre usuel sur les entiers. Intuitivement, une description d d'une des deux représentations correspond à la minoration d'un nombre x donné. Ainsi, la description $1 \in \mathcal{R}_1$ peut s'interpréter comme l'assertion $x \geq 1$. Maintenant, si l'on sait que $x \geq d_1$ pour $d_1 \in \mathcal{R}_1$ (c'est-à-dire que d_1 correspond à un nombre impair), la meilleure borne inférieure paire que l'on peut en donner est $d_2 = d_1 - 1$ (en identifiant une description et le nombre entier associé). Ainsi, les deux fonctions $f_{1|2}$ et $f_{2|1}$ sont les restrictions de la fonction précédentes aux nombres pairs et impairs. En particulier, $f_{1|2} \circ f_{2|1}(d) = d - 2 < d$.

Cette inégalité permet de relier entre elles les différentes représentations de façon simple, et exprime des relations fortes entre les fonctions de transformation (comme le montre l'inégalité $d \leq f_{i|j} \circ f_{j|i}(d)$). Elle impose en fait une condition très forte sur les différents points de vue, qui est qu'ils correspondent tous maintenant à l'observation de la *même* instance du système, et surtout que de l'information est conservée lorsque l'on fait passer une description d'une représentation à une autre par le biais d'un $f_{i|j} : d \leq f_{i|j} \circ f_{j|i}(d)$ exprime que la connaissance qu'a l'agent \mathcal{A}_i des informations de \mathcal{A}_j peut être utilisée pour ses observations. Cela va impliquer pour la suite que la logique correspondant aux systèmes de représentations dont les $f_{i|j}$ vérifient les inégalités précédentes sera la logique OL^- à laquelle on ajoute un axiome concernant plusieurs opérateurs de modalité K_i .

6.2.1 Renforcement de la logique

Nous allons donc étudier concrètement quel axiome va être ajouté. Pour cela, nous allons tout d'abord donner une nouvelle définition :

Definition 3

Un système compositionnel de représentations est un triplet $\langle \mathcal{I}, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$, où \mathcal{I} est un ensemble d'indices, pour chaque indice $i \in \mathcal{I}$, \mathcal{R}_i est une représentation et les fonctions $\{f_{i|j} : \mathcal{R}_j \rightarrow \mathcal{R}_i\}$ définies entre chaque paire de représentations sont croissantes, avec en particulier $\forall i \in \mathcal{I}, f_{i|i} = \text{id}_i$. On impose de plus qu'elles vérifient :

$$\forall i, j, k \in \mathcal{I}, \forall d \in \mathcal{R}_i, f_{k|i}(d) \leq_k f_{k|j} \circ f_{j|i}(d)$$

On définit le symbole \models_{sc} pour exprimer la validation par des systèmes compositionnels de représentations. Il correspond à la même fonction d'interprétation que celle définie pour les systèmes de représentations, mais en se limitant aux seuls systèmes compositionnels.

La structure de l'inégalité que l'on impose aux fonctions de transformations constitue une indication pour trouver la forme de l'axiome que l'on va ajouter à la logique OL^- . En effet, elle signifie que pour toute interprétation $\llbracket \varphi \rrbracket_k \subseteq \wp^\downarrow(\mathcal{R}_k)$, pour toute description $d \in \mathcal{R}_i$, on a :

$$f_{k|j} \circ f_{j|i}(d) \in \llbracket \varphi \rrbracket_k \Rightarrow f_{k|i}(d) \in \llbracket \varphi \rrbracket_k$$

Cette implication correspond à la proposition $K_j K_k \varphi \rightarrow K_k \varphi$. Nommons \mathbf{T}_2 cet axiome. Il s'agit d'un axiome plus fort que \mathbf{T}_1 , mais plus faible que \mathbf{T} . Il s'agit en fait d'une restriction de ce dernier, restriction portant sur la forme des assertions que l'on peut faire sur le système et qui oblige à toujours garder au moins un opérateur K_i dans les affirmations que l'on peut faire sur l'état du système.

Le théorème suivant montre que l'on obtient effectivement la logique associée aux systèmes compositionnels en ajoutant \mathbf{T}_2 à OL^- , pour obtenir la logique OL . On rappelle qu'elle est la logique obtenue en ajoutant à la logique intuitionniste les règles de nécessité et d'universalité, ainsi que les axiomes \mathbf{K} , \mathbf{D} , \mathbf{L} , \mathbf{T}_2 et \mathbf{KV} . Ces axiomes et règles sont rappelés dans le table 7.

$K_i(\varphi \rightarrow \psi) \rightarrow K_i \varphi \rightarrow K_i \psi$	K
$K_i \varphi \rightarrow \neg K_i \neg \varphi$	D
$K_i K_j \varphi \rightarrow K_j \varphi$	T₂
$K_i(\varphi \leftrightarrow K_i \varphi)$	L
$K_i(\varphi \vee \psi) \rightarrow K_i \varphi \vee K_i \psi$	KV
$\frac{\vdash \varphi}{\vdash K_i \varphi}$	Nec
$\frac{\forall i, \vdash K_i \varphi}{\vdash \varphi}$	Univ

Table 7 – Axiomes et règles de la logique OL

Théorème 4

Toute proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$ est prouvable pour la logique OL si et seulement si elle est valide pour les systèmes compositionnels :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I}, \Psi}, \vdash_{\text{OL}} \varphi \Leftrightarrow \models_{sc} \varphi$$

Preuve La preuve complète se trouve en annexe B.4. □

La structure de système compositionnel semble être une bonne candidate comme formalisme de représentation des observations d'un système. Elle est basée sur des représentations qui sont des structures algébriques très générales, et les conditions que l'on impose sur les relations entre différentes représentations sont aussi peu restrictives que possible. Dans la suite, nous omettrons le terme "compositionnel", et tous les systèmes de représentation que nous rencontrerons seront compositionnels, à moins que le contraire ne soit précisé.

Chapitre 7

Quelques variations logiques autour de OL

Nous allons dans ce chapitre nous attarder sur la logique de l'observation OL que l'on vient d'identifier. Tout d'abord, nous allons présenter une version en calcul des séquents à la Gentzen. Ensuite, nous nous intéresserons à quelques propriétés vérifiées par OL mettant en lumière diverses manipulations que l'on peut faire sur l'ensemble des indices \mathcal{I} , avec en particulier un résultat exprimant le fait que pour prouver une proposition donnée, on peut se restreindre à un ensemble fini d'indices, ce qui entraînera certains résultats pour le cas où \mathcal{I} est infini.

7.1 Formulation en calcul des séquents

La formulation de la logique sous forme de calcul des séquents [29] permet d'appréhender plus facilement les aspects déduction et prouvabilité. Ainsi, nous utiliserons cette formulation pour exprimer certains résultats sur l'utilisation des coupures et de la règle d'universalité dans le processus de preuve.

Dans un premier temps, nous allons cependant nous restreindre au cas où l'ensemble des indices \mathcal{I} est fini. En effet, la règle d'universalité peut poser des problèmes à cause de la présence d'une possible infinité de prémisses (puisque sa définition est la possibilité d'inférer $\vdash \varphi$ des prémisses $\forall i, \vdash K_i \varphi$). Nous verrons plus loin qu'il est possible d'utiliser la formulation en système des séquents même dans le cas d'un ensemble \mathcal{I} infini, mais commençons donc par nous restreindre au cas fini.

7.1.1 Localisation des propositions

Avant de définir les règles composant le calcul des séquents, nous allons introduire une notation qui permettra une gestion des indices de localisation intervenant dans les termes que

nous auront à manipuler. Ainsi, les propositions qui interviendront dans les séquents seront écrites sous la forme $[\varphi]_{\Lambda}$ où Λ est une suite finie d'indices. Intuitivement, un terme de la forme $[\varphi]_{i_1, \dots, i_n}$ correspond au terme $K_{i_n} \dots K_{i_1} \varphi$.

Pour cela, définissons tout d'abord l'ensemble \mathcal{I}^* des mots de \mathcal{I} , soit plus précisément l'ensemble des suites finies à valeurs dans \mathcal{I} . Nous identifierons en particulier le mot vide qui sera noté ε , et l'opération de concaténation sera elle notée \circ . Nous allons cependant apporter une petite modification, qui est une conséquence de l'équivalence $K_i K_i \varphi \leftrightarrow K_i \varphi$. Aussi, nous nous restreindrons aux mots n'ayant pas deux fois la même lettre consécutivement, ce qui revient à quotienter l'ensemble des mots par la relation d'équivalence définie par $\Lambda_1 \circ i \circ i \circ \Lambda_2 \simeq \Lambda_1 \circ i \circ \Lambda_2$. Pour l'opération de concaténation, on considère donc que $i \circ (i \circ \Lambda)$ est égal à $i \circ \Lambda$. Nous continuerons néanmoins à noter cet ensemble simplement \mathcal{I}^* .

Nous allons de plus définir deux ordres partiels sur \mathcal{I}^* . Tout d'abord, l'ordre \leq correspond au cas où un mot est obtenue à partir d'un autre en enlevant un nombre arbitraire de lettres. Ainsi, si l'on note $e = e(1) \circ e(2) \circ \dots \circ e(n)$ (avec éventuellement $n = 0$), on pose :

$$e \leq f \Leftrightarrow \exists \{\omega_i \in \mathcal{I}^*\}_{0 \leq i \leq n} : f = \omega_0 \circ e(1) \circ \omega_1 \dots \circ e(n) \circ \omega_n$$

De façon équivalente, $e \leq f$ si et seulement si il existe une fonction pos strictement croissante sur les indices, telle que $\forall i \in \{1, \dots, n\}, e(i) = f(\text{pos}(i))$.

Nous aurons de plus besoin d'un second ordre partiel, noté lui \leq_* qui est une restriction au cas où les deux mots commencent par la même lettre s'il ne sont pas vide. Ainsi, on pose :

$$e \leq_* f \Leftrightarrow \begin{cases} e = f = \varepsilon & \text{ou} \\ \exists i \in \mathcal{I}, e', f' \in \mathcal{I}^* : e = i \circ e' \text{ et } f = i \circ f' \text{ et } e' \leq f' \end{cases}$$

On peut remarquer que cet ordre retranscrit la façon dont les opérateurs modaux peut se composer à la lumière de l'axiome \mathbf{T}_2 . En effet, si $\Lambda' \leq_* \Lambda$, on a $\vdash_{\text{OL}} K_{\Lambda} \varphi \rightarrow K_{\Lambda'} \varphi$ où l'on a défini la notation K_{Λ} par $K_{i \circ \Lambda} \varphi = K_i K_{\Lambda} \varphi$ et $K_{\varepsilon} \varphi = \varphi$.

7.1.2 Règles à la “Gentzen” pour OL

Les règles sont données dans la table 8 p.77. Elles peuvent se diviser en trois parties. Les règles *Axiom* et \perp d'une part qui permettent l'introduction des éléments de base. Viennent ensuite les règles structurelles (*Weak*, *Contract*, *Cut*, *Loc* et *Univ*) qui ne modifient pas les propositions apparaissant dans les séquents, et enfin les règles logiques qui précisent la façon dont on peut introduire les différents connecteurs logiques (\wedge , \vee , \rightarrow et K).

Nous allons tout d'abord montrer que ce système de séquent permet bien d'obtenir la logique OL, en commençant par le résultat de complétude :

$$\begin{array}{c}
\frac{\Lambda' \leq_* \Lambda}{[\varphi]_\Lambda \vdash [\varphi]_{\Lambda'}} \text{Axiom} \quad \frac{\Lambda' \leq \Lambda}{[\perp]_\Lambda \vdash [\varphi]_{\Lambda'}} \perp \\
\\
\frac{\Gamma \vdash [\psi]_\Lambda}{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda} \text{Weak} \quad \frac{\Gamma, [\varphi]_\Lambda, [\varphi]_\Lambda \vdash \psi}{\Gamma, [\varphi]_\Lambda \vdash \psi} \text{Contract} \quad \frac{\forall i, \Gamma_{oi} \vdash [\varphi]_{\Lambda \circ i}}{\Gamma \vdash [\varphi]_\Lambda} \text{Univ} \\
\\
\frac{\Gamma \vdash [\varphi]_\Lambda \quad \Delta, [\varphi]_{\Lambda'} \vdash \psi \quad \Lambda' \leq_* \Lambda}{\Gamma, \Delta \vdash \psi} \text{Cut} \quad \frac{\Gamma, [\varphi]_{\Lambda'} \vdash \psi \quad \Lambda' \leq_* \Lambda}{\Gamma, [\varphi]_\Lambda \vdash \psi} \text{Loc} \\
\\
\frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta \quad \Gamma, [\psi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \vee \psi]_\Lambda \vdash \vartheta} \vee L \quad \frac{\Gamma \vdash [\varphi]_\Lambda}{\Gamma \vdash [\varphi \vee \psi]_\Lambda} \vee R_1 \quad \frac{\Gamma \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \vee \psi]_\Lambda} \vee R_2 \\
\\
\frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \wedge \psi]_\Lambda \vdash \vartheta} \wedge L_1 \quad \frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \wedge \psi]_\Lambda \vdash \vartheta} \wedge L_2 \quad \frac{\Gamma \vdash [\varphi]_\Lambda \quad \Gamma \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \wedge \psi]_\Lambda} \wedge R \\
\\
\frac{\Gamma \vdash [\varphi]_\Lambda \quad \Gamma, [\psi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \rightarrow \psi]_\Lambda \vdash \vartheta} \rightarrow L \quad \frac{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \rightarrow \psi]_\Lambda} \rightarrow R \\
\\
\frac{\Gamma, [\varphi]_{i \circ \Lambda} \vdash [\psi]_{\Lambda'}}{\Gamma, [K_i \varphi]_\Lambda \vdash [\psi]_{\Lambda'}} KL \quad \frac{\Gamma \vdash [\varphi]_{i \circ \Lambda}}{\Gamma \vdash [K_i \varphi]_\Lambda} KR
\end{array}$$

Table 8 – Formulation en calcul des séquents de OL

Proposition 1

Pour toute proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$, on a :

$$\vdash_{\text{OL}} \varphi \Rightarrow \emptyset \vdash [\varphi]_\varepsilon$$

Preuve Cela vient du fait que les différents axiomes de OL sont prouvables dans ce calcul des séquents avec la localisation ε , et que la règle d'universalité peut aussi être exprimée dans

ce système. Pour les axiomes, nous en donnons les dérivations :

$$\mathbf{K} : \frac{\frac{\frac{\text{Axiom}}{[\varphi]_i \vdash [\varphi]_i} \text{Axiom} \quad \frac{\frac{\text{Axiom}}{[\psi]_i \vdash [\psi]_i} \text{Axiom}}{[\varphi]_i, [\psi]_i \vdash [\psi]_i} \text{Weak}}{[\varphi \rightarrow \psi]_i, [\varphi]_i \vdash [\psi]_i} \rightarrow L}{\frac{[K_i(\varphi \rightarrow \psi)]_\varepsilon, [K_i \varphi]_\varepsilon \vdash [K_i \psi]_\varepsilon}{\vdash [K_i(\varphi \rightarrow \psi) \rightarrow K_i \varphi \rightarrow K_i \psi]_\varepsilon} \rightarrow R} KR, KL$$

$$\mathbf{D} : \frac{\frac{\frac{\text{Axiom}}{[\perp]_i \vdash [\perp]_\varepsilon} \text{Axiom}}{[K_i \perp]_\varepsilon \vdash [\perp]_\varepsilon} KL}{\vdash [K_i \perp \rightarrow \perp]_\varepsilon} \rightarrow R \quad \mathbf{T}_2 : \frac{\frac{\frac{\text{Axiom}}{[\varphi]_{i \circ j} \vdash [\varphi]_i} \text{Axiom}}{[K_j K_i \varphi]_\varepsilon \vdash [K_i \varphi]_\varepsilon} KL}{\vdash [K_j K_i \varphi \rightarrow K_i \varphi]_\varepsilon} \rightarrow R$$

$$\mathbf{L}_T : \frac{\frac{\frac{\text{Axiom}}{[\varphi]_{i \circ i} \vdash [\varphi]_i} \text{Axiom}}{[K_i \varphi]_i \vdash [\varphi]_i} KL}{\vdash [K_i \varphi \rightarrow \varphi]_i} \rightarrow R}{\vdash [K_i(K_i \varphi \rightarrow \varphi)]_\varepsilon} KR \quad \mathbf{L}_4 : \frac{\frac{\frac{\text{Axiom}}{[\varphi]_i \vdash [\varphi]_{i \circ i}} \text{Axiom}}{[\varphi]_i \vdash [K_i \varphi]_i} KL}{\vdash [\varphi \rightarrow K_i \varphi]_i} \rightarrow R}{\vdash [K_i(\varphi \rightarrow K_i \varphi)]_\varepsilon} KR$$

Pour ces deux dérivations, on a fait apparaître explicitement des localisations de la forme $i \circ i$ qui sont en fait équivalentes à i , afin uniquement de rendre l'enchaînement des règles plus clair.

$$\mathbf{KV} : \frac{\frac{\frac{\text{Axiom}}{[\varphi]_i \vdash [\varphi]_i} \text{Axiom}}{[\varphi]_i \vdash [K_i \varphi]_\varepsilon} KR}{[\varphi]_i \vdash [K_i \varphi \vee K_i \psi]_\varepsilon} \vee R_1 \quad \frac{\frac{\frac{\text{Axiom}}{[\psi]_i \vdash [\psi]_i} \text{Axiom}}{[\psi]_i \vdash [K_i \psi]_\varepsilon} KR}{[\psi]_i \vdash [K_i \varphi \vee K_i \psi]_\varepsilon} \vee R_2}{[\varphi \vee \psi]_i \vdash [K_i \varphi \vee K_i \psi]_\varepsilon} \vee L}{\frac{[\varphi \vee \psi]_i \vdash [K_i \varphi \vee K_i \psi]_\varepsilon}{[K_i(\varphi \vee \psi)]_\varepsilon \vdash [K_i \varphi \vee K_i \psi]_\varepsilon} KL}{\vdash [K_i(\varphi \vee \psi) \rightarrow K_i \varphi \vee K_i \psi]_\varepsilon} \rightarrow R$$

Pour le cas de la règle *Univ*, nous allons anticiper sur la suite en utilisant la propriété 10 p.83, qui découle de la possibilité d'éliminer les coupures et permet de déduire une preuve de $\vdash [\varphi]_i$ à partir d'une preuve de $\vdash [K_i \varphi]_\varepsilon$. En effet, si l'on utilise la règle d'universalité pour déduire $\vdash_{OL} \varphi$ de $\forall i, \vdash_{OL} K_i \varphi$, on en déduit par récurrence que pour tout i , le séquent $\emptyset \vdash [K_i \varphi]_\varepsilon$.

Alors, d'après la proposition 10, on en déduit une preuve Π_i de $\vdash [\varphi]_i$, et on construit :

$$\frac{\Pi_i}{\forall i, \vdash [\varphi]_i} \text{Univ} \\ \vdash [\varphi]_\varepsilon$$

□

Nous venons de montrer que le système des séquents que nous venons de présenter est complet pour la logique OL. Nous allons maintenant nous pencher sur le résultat de correction. Pour cela, nous allons introduire deux notations pratiques, avant de présenter le résultat proprement dit.

Definition 4

Étant donné un mot $\Lambda \in \mathcal{I}^*$ et une proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$, on définit $K_\Lambda \varphi$ par récurrence par :

$$K_{i \circ \Lambda} \varphi = K_\Lambda K_i \varphi \quad K_\varepsilon \varphi = \varphi$$

De même, étant donné un système de représentations \mathcal{R} , on définit $f_{\Lambda|j}$ de façon récursive par :

$$f_{\Lambda \circ i|j} = f_{\Lambda|i} \circ f_{i|j} \quad f_{\varepsilon|j} = \text{id}_j$$

Cette dernière définition permet d'avoir la propriété suivante, qui relie l'ordre \leq_* avec les fonctions de transformation d'un système compositionnel de représentations :

Proposition 2

Pour tout système de représentations \mathcal{R} , pour $i \in \mathcal{I}$ et $\Lambda', \Lambda \in \mathcal{I}^*$ tels que $\Lambda' \leq_* \Lambda$, on a :

$$f_{\Lambda'|i} \leq f_{\Lambda|i}$$

Preuve C'est une conséquence directe de la définition de l'ordre \leq_* et de la règle de composition 6.1 p.70. □

Ces notations vont nous permettre d'exprimer le résultat de correction :

Proposition 3

Étant donné un séquent $\Gamma \vdash [\varphi]_\Lambda$ prouvable à l'aide des règles définies ci-dessus, un système de représentations \mathcal{R} et une interprétation atomique ν , on a :

$$\forall i, \bigwedge_j \{[\gamma_j]_{\Lambda_j}\}_i \subseteq \{[\varphi]_\Lambda\}_i$$

où $\Gamma = [\gamma_1]_{\Lambda_1}, \dots, [\gamma_n]_{\Lambda_n}$ et où l'on utilise la notation $\forall \varphi, \Lambda, i, \{[\varphi]_\Lambda\}_i = \{K_\Lambda \varphi\}_i$.

Preuve On prouve ce résultat par induction sur la hauteur de la preuve. Les cas de bases sont constitués par les règles *Axiom* et \perp . Pour la première règle, si l'on a $[\varphi]_{\Lambda} \vdash [\varphi]_{\Lambda'}$ avec $\Lambda' \leq_* \Lambda$, le résultat découle de la proposition 2 précédente. Pour la règle \perp , cela vient du fait que $\{[\perp]_{\Lambda}\}_i = \emptyset$.

Nous allons nous intéresser aux autres règles. Tout d'abord, on vérifie facilement que les règles *Weak*, *Contract*, *Cut* ainsi que les règles d'introduction des connecteurs \vee , \wedge et \rightarrow ne posent pas de problème.

On prouve la cas des règles *KL* et *KR* en remarquant :

$$\{[K_i \varphi]_{\Lambda}\} = \{[K_{\Lambda} K_i \varphi]\} = \{[\varphi]_{i \circ \Lambda}\}$$

La règle *Loc* est elle aussi une conséquence de la proposition 2. Il reste à prouver le cas de la règle *Univ*. Reprenons les notations utilisées pour la définition de la règle, et posons $i \in \mathcal{I}$. Par récurrence, on sait que $\forall j, \{\Gamma_{oi}\}_j \subseteq \{[\varphi]_{\Lambda \circ i}\}_j$. Pour le cas particulier où $j = i$, on a $\{\Gamma_{oi}\}_i = \{\Gamma\}_i$ et $\{[\varphi]_{\Lambda \circ i}\}_i = \{[\varphi]_{\Lambda}\}_i$, ce qui permet de conclure. \square

Nous avons donc prouvé que le système de séquents que nous avons introduit est correct et complet vis-à-vis de OL, ce que nous notons formellement par :

Théorème 5

Pour tout $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$, on a :

$$\vdash_{\text{OL}} \varphi \Leftrightarrow \emptyset \vdash [\varphi]_{\varepsilon}$$

Preuve On a en fait $\vdash_{\text{OL}} \varphi \Rightarrow \vdash [\varphi]_{\varepsilon} \Rightarrow \models_{\text{SC}} \varphi \Rightarrow \vdash_{\text{OL}} \varphi$. La première implication est donnée par la proposition 1, la seconde par 3 et la dernière par le théorème 4 p.72. \square

7.2 Propriétés et manipulations de preuves

Nous allons tout d'abord énoncer des propriétés concernant la localisation des différentes propositions apparaissant dans les preuves.

Proposition 4

Pour tout séquent $[\gamma_1]_{\Lambda_1}, \dots, [\gamma_n]_{\Lambda_n} \vdash [\varphi]_{\Lambda}$ prouvable avec le système précédent, on a :

$$\forall i, \Lambda \leq \Lambda_i$$

Preuve On prouve ce résultat par induction sur la hauteur de la preuve. On remarque que pour les cas de base, une condition est ajoutée spécifiquement pour que cette propriété soit vérifiée. Les autres règles pouvant poser une difficulté sont les règles *Cut*, *Loc* et *KL*. Pour ces deux dernières, une condition est donnée explicitement pour que la condition soit vérifiée. Pour la règle de coupure *Cut*, en reprenant les notations de la définition, on a par récurrence Λ' qui est supérieur à la localisation de ψ , et toutes les propositions intervenant dans Γ ont une localisation supérieure à Λ , et donc à Λ' , ce qui permet de conclure. \square

Proposition 5

Étant donné une preuve Π de $\Gamma \vdash [\varphi]_{\Lambda}$, pour tout séquent $\Delta \vdash [\psi]_{\Lambda'}$ apparaissant dans Π , on a $\Lambda \leq \Lambda'$.

Preuve Ce résultat se prouve lui aussi par induction sur la hauteur de la preuve. Pour cela, il faut remarquer en particulier que les seules règles agissant sur la localisation d'un terme à droite du symbole \vdash – soit *Cut*, *Univ*, $\rightarrow L$ et *KR* – ont deux actions possibles sur la localisation en regardant du bas vers le haut : soit on ajoute un indice à la localisation (règles *Univ* et *KR*), soit on a un séquent supplémentaire dont la localisation du terme de droite est la localisation d'un terme de gauche du séquent de départ (règles *Cut* et $\rightarrow L$). La proposition précédente s'applique alors. \square

Corollaire 5.1

Étant donné une preuve Π de $\Gamma \vdash [\varphi]_{\Lambda}$, toute proposition apparaissant dans Π a une localisation supérieure à Λ .

Preuve C'est une conséquence directe des deux propositions précédentes. \square

Après ces quelques résultats à propos de la façon dont les localisations sont organisées dans une preuve, nous allons maintenant nous pencher sur quelques manipulations simples que l'on peut faire pour une preuve.

Proposition 6

Étant donné une preuve Π de $\Gamma \vdash [\perp]_{\Lambda}$, il est possible de la transformer en une preuve Π' de $\Gamma \vdash [\varphi]_{\Lambda'}$ pour φ quelconque et $\Lambda' \leq \Lambda$. De plus, Π' et Π ont le même enchaînement de règles, à l'exception de règles Axiom qui peuvent être transformées en règles \perp .

Preuve Pour obtenir la preuve Π' , nous allons marquer le terme $[\perp]_{\Lambda}$ à changer en $[\varphi]_{\Lambda'}$, en écrivant $[[\perp]]_{\Lambda'}^{\Lambda'}$ et nous allons donner des règles de transformations pour faire remonter ce marquage le long de la preuve jusqu'à le faire disparaître.

La règle d'axiome se transforme de la façon suivante :

$$\frac{}{[\perp]_{\Lambda} \vdash [\perp]_{\Lambda'}^{\Lambda''}} \text{Axiom} \quad \mapsto \quad \frac{}{[\perp]_{\Lambda} \vdash [\varphi]_{\Lambda''}} \perp$$

La règle \perp se traduit de façon similaire. Pour les autres règles, la transformation se fait simplement en faisant remonter le marquage vers les termes concernés. Nous donnons quelques exemples de transformation :

$$\begin{array}{ccc} \frac{\Gamma \vdash [\psi]_{\Lambda} \quad \Delta, [\psi]_{\Lambda} \vdash [\perp]_{\Lambda'}}{\Gamma, \Delta \vdash [\perp]_{\Lambda'}^{\Lambda''}} \text{Cut} & \mapsto & \frac{\Gamma \vdash [\psi]_{\Lambda} \quad \Delta, [\psi]_{\Lambda} \vdash [\perp]_{\Lambda'}^{\Lambda''}}{\Gamma, \Delta \vdash [\varphi]_{\Lambda''}} \text{Cut} \\ \frac{\Gamma \vdash [\psi_1]_{\Lambda} \quad \Gamma, [\psi_2]_{\Lambda} \vdash [\perp]_{\Lambda'}}{\Gamma, [\psi_1 \rightarrow \psi_2]_{\Lambda} \vdash [\perp]_{\Lambda'}^{\Lambda''}} \rightarrow L & \mapsto & \frac{\Gamma \vdash [\psi_1]_{\Lambda} \quad \Gamma, [\psi_2]_{\Lambda} \vdash [\perp]_{\Lambda'}^{\Lambda''}}{\Gamma, [\psi_1 \rightarrow \psi_2]_{\Lambda} \vdash [\varphi]_{\Lambda''}} \rightarrow L \\ \frac{\forall i, \Gamma_{oi} \vdash [\perp]_{\Lambda'oi}}{\Gamma \vdash [\perp]_{\Lambda'}^{\Lambda''}} \text{Univ} & \mapsto & \frac{\forall i, \Gamma_{oi} \vdash [\perp]_{\Lambda'oi}^{\Lambda''oi}}{\Gamma \vdash [\varphi]_{\Lambda''}} \text{Univ} \end{array}$$

Finalement, la preuve Π' s'obtient en appliquant les transformations précédentes à :

$$\begin{array}{c} \Pi \\ \Gamma \vdash [\perp]_{\Lambda}^{\Lambda'} \end{array}$$

De plus, on montre par induction sur la hauteur de la preuve que la preuve obtenue est bien une preuve correcte. \square

Proposition 7

Étant donné une preuve Π de $\Gamma \vdash [\varphi]_{\Lambda}$, il est possible de la transformer en une preuve Π' de $\Gamma_{oi} \vdash [\varphi]_{\Lambda oi}$ où l'enchaînement des règles reste inchangé.

Preuve La transformation s'effectue simplement en ajoutant l'indice i à droite de toutes les localisation de toutes les propositions apparaissant dans la preuve. La preuve obtenue est bien correcte, puisqu'en particulier les conditions sur les localisation des différentes règles restent vérifiées. \square

Proposition 8

Étant donné une preuve Π de $\Gamma \vdash [\varphi]_{\Lambda}$, il est possible de la transformer en une preuve Π' de $\Gamma \vdash [\varphi]_{\Lambda'}$ où $\Lambda' \leq_{\star} \Lambda$. De plus, la structure de Π' ne diffère par rapport à Π que par l'ajout de règles Loc.

Preuve Cette proposition est une conséquence des propositions 4 et 5. Cependant, la preuve, bien que sans difficultés techniques, nécessite d'introduction de notations et de quelques propriétés les concernant. C'est la raison pour laquelle nous donnons la preuve dans l'annexe C.1 p.179 afin de ne pas trop alourdir cette partie. \square

Nous présentons maintenant la proposition principale d'élimination de coupures.

Théorème 6 (Élimination des coupures)

Étant donné une preuve Π de $\Gamma \vdash [\varphi]_{\Lambda}$, il est possible de transformer celle-ci en une preuve Π' du même séquent, mais n'utilisant pas la règle de coupure Cut.

Preuve La preuve de ce théorème, longue bien que sans réelles difficultés, est donnée dans l'annexe C.2 p.186. \square

7.3 Quelques propriétés sur les indices

La présence de la règle d'universalité, lorsque l'on a affaire à un ensemble d'indices \mathcal{I} infini, peut laisser craindre qu'il est impossible d'utiliser le formalisme de calcul des séquents, sous peine d'avoir une preuve infinie. Nous allons étudier quelques propriétés vérifiées par la logique OL, et nous montrerons en particulier que pour prouver la validité d'une proposition, on peut toujours se ramener à un ensemble fini d'indices, ce qui permettra de supprimer cette difficulté. Tout d'abord, montrons une déduction vraie en général peut être faite localement dans une représentation donnée :

Proposition 9

Pour tout ensemble fini de formules Γ , toute formule φ et pour tout indice i , on a :

$$\Gamma \vdash_{\text{OL}} \varphi \Rightarrow K_i \Gamma \vdash_{\text{OL}} K_i \varphi$$

Preuve C'est une conséquence directe de la proposition 7. \square

De même, on peut énoncer une propriété précisant que

Proposition 10

Pour toute proposition φ et tout indice i , si $\emptyset \vdash [K_i \varphi]_{\varepsilon}$ est prouvable, alors il en est de même pour $\emptyset \vdash [\varphi]_i$.

Preuve En effet, toute preuve sans coupure de $\vdash [K_i \varphi]_\varepsilon$ a comme règle à la racine soit une instance de *Univ*, soit de *KR*. Si c'est cette dernière qui est à la racine, on a directement le résultat voulu. Sinon, parmi les prémisses on a une preuve du séquent $\vdash [K_i \varphi]_i$. Si l'on regarde la règle utilisée pour obtenir cette conclusion, on remarque à nous que soit *Univ*, soit *KR* ont été utilisées. En répétant le processus, on arrive forcément à un moment où l'on a utilisé la règle *KR* pour déduire $\vdash [K_i \varphi]_{i \circ \dots \circ i}$ de $\vdash [\varphi]_{i \circ \dots \circ i}$, ne pouvant avoir un empilement infini de règles *Univ*. Maintenant, puisque l'on a identifié i et $i \circ \dots \circ i$, on en déduit que la preuve de $\vdash [\varphi]_{i \circ \dots \circ i}$ est en fait une preuve de $\vdash [\varphi]_i$. \square

Nous allons maintenant nous intéresser à quelques manipulations d'indices lors de preuves. Tout d'abord, on montre que les indices associés aux représentations n'ont aucune signification intrinsèque, et que l'on peut les renommer sans modifier la prouvabilité d'une proposition. On peut même fusionner plusieurs indices sans que cela n'ait d'incidence :

Proposition 11

Étant donné un ensemble d'indices \mathcal{J} , un sous-ensemble $\mathcal{I} \subseteq \mathcal{J}$, et une fonction $\sigma : \mathcal{I} \rightarrow \mathcal{J}$, on a :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I}, \Psi}, \vdash_{\text{OL}, \mathcal{I}} \varphi \Rightarrow \vdash_{\text{OL}, \sigma(\mathcal{I})} \sigma(\varphi)$$

où $\sigma(\varphi)$ désigne le terme obtenu à partir de φ en échangeant chaque indice i en $\sigma(i)$ (et donc chaque K_i en $K_{\sigma(i)}$).

Preuve Cela vient du fait qu'étant donné une preuve Π de $\vdash [\varphi]_\varepsilon$, la preuve obtenue en appliquant la substitution σ à tous les indices (que ce soit pour les K_i et les localisations Λ) est une preuve de $\vdash [\sigma(\varphi)]_\varepsilon$ pour l'ensemble d'indices $\sigma(\mathcal{I})$. \square

Introduisons la notation $\text{ind}(\varphi)$ qui est l'ensemble des indices apparaissant dans un terme φ donné. Cet ensemble est toujours fini, puisque φ l'est aussi. La proposition précédente permet de montrer :

Corollaire 11.1

Étant donné un ensemble d'indices \mathcal{I} et une proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$, pour tout ensemble \mathcal{J} tel que $\text{ind}(\varphi) \subseteq \mathcal{J} \subseteq \mathcal{I}$, on a :

$$\vdash_{\text{OL}, \mathcal{I}} \varphi \Rightarrow \vdash_{\text{OL}, \mathcal{J}} \varphi$$

Preuve On utilise la proposition 11 avec une fonction $\sigma : \mathcal{I} \rightarrow \mathcal{I}$ fixant \mathcal{J} et telle que $\sigma(\mathcal{I}) = \mathcal{J}$. On remarque en particulier que l'on a $\sigma(\varphi) = \varphi$. \square

Ce corollaire montre qu'il est correct de restreindre l'ensemble \mathcal{I} d'indices à un sous-ensemble fini. La proposition suivante va permettre d'étudier le cas inverse, et justifier le fait

que la prouvabilité pour un sous-ensemble particulier est complète vis-à-vis de l'ensemble \mathcal{I} en entier.

Proposition 12

Étant donné deux ensembles d'indices \mathcal{I} et \mathcal{J} tels que $\mathcal{J} \not\subseteq \mathcal{I}$, en posant $j \in \mathcal{J} \setminus \mathcal{I}$, on a :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I}, \Psi}, \vdash_{\text{OL}, \mathcal{I} \cup \{j\}} \varphi \Rightarrow \vdash_{\text{OL}, \mathcal{I} \cup \mathcal{J}} \varphi$$

Preuve Pour prouver ce résultat, nous allons utiliser le fait que la logique OL est modélisée par les systèmes de représentations. Par hypothèse, on a pour tout système \mathcal{S} et toute interprétation atomique ν , on a $\forall i \in \text{ind}(\varphi) \cup \{j\}, \llbracket \varphi \rrbracket_i = \mathcal{R}_i$.

Soit maintenant un système de représentations \mathcal{S} et une interprétation atomique ν donnés. Soit de plus un indice $i \in \mathcal{I} \cup \mathcal{J}$. Nous allons montrer que $\llbracket \varphi \rrbracket_i = \mathcal{R}_i$. On sait que c'est le cas si $i \in \mathcal{I} \cup \{j\}$. Sinon, définissons un nouveau système de représentations \mathcal{S}' en partant de \mathcal{S} et en intervertissant les indices i et j , c'est-à-dire en ayant :

$$\begin{aligned} \mathcal{R}'_i &= \mathcal{R}_j & \mathcal{R}'_j &= \mathcal{R}_i \\ \mathcal{R}'_k &= \mathcal{R}_k & k &\notin \{i, j\} \\ f'_{i|k} &= f_{j|k} & f'_{k|i} &= f_{k|j} & k &\notin \{i, j\} \\ f'_{j|k} &= f_{i|k} & k &\notin \{i, j\} \\ f'_{l|k} &= f_{l|k} & k, l &\notin \{i, j\} \end{aligned}$$

De même, on définit ν' en intervertissant i et j . D'après l'hypothèse de départ, on a $\llbracket \varphi \rrbracket_{\mathcal{S}', \nu', j} = \mathcal{R}'_j$. Or, l'expression ensembliste de $\llbracket \varphi \rrbracket_{\mathcal{S}', \nu', j}$ ne fait intervenir comme fonctions que les $f'_{k|j}$ et les $f'_{l|k}$ avec dans les deux cas $k, l \notin \{i, j\}$. En partant de cette expression et en intervertissant à nouveau i et j , on obtient l'interprétation de $\llbracket \varphi \rrbracket_{\mathcal{S}, \nu, i}$ qui vaut \mathcal{R}_i , puisque l'on avait $\llbracket \varphi \rrbracket_{\mathcal{S}', \nu', j} = \mathcal{R}'_j$. En généralisant ce résultat pour tout $i \notin \mathcal{I} \cup \{j\}$, on a montré :

$$\forall i \in \mathcal{I} \cup \mathcal{J}, \llbracket \varphi \rrbracket_i = \mathcal{R}_i$$

En généralisant maintenant pour tous les systèmes de représentations et toutes les interprétations atomiques, on obtient le résultat recherché. \square

En utilisant conjointement le corollaire 11.1 et la proposition 12, on peut montrer que la prouvabilité pour un ensemble infini d'indices \mathcal{I} est équivalente à la prouvabilité pour un sous-ensemble fini $\mathcal{J} \subseteq \mathcal{I}$ bien choisi.

Théorème 7

Étant donné un ensemble infini d'indices \mathcal{I} , soit $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$. Si l'on note $\text{ind}(\varphi)$ l'ensemble des indices apparaissant dans le terme φ , et pour $i \in \mathcal{I} \setminus \text{ind}(\varphi)$, on a :

$$\vdash_{\text{OL}, \text{ind}(\varphi) \cup \{i\}} \varphi \Leftrightarrow \vdash_{\text{OL}, \mathcal{I}} \varphi$$

Preuve La proposition 12 permet de prouver le sens \Rightarrow , tandis que le sens \Leftarrow se prouve avec le corollaire 11.1 en posant $\mathcal{J} = \text{ind}(\varphi) \cup \{i\}$. \square

7.4 Élimination de la règle d'universalité dans le cas infini

Pour finir, nous allons montrer que dans le cas où l'ensemble des indices est infini, on peut effectivement utiliser le système de séquents défini précédemment sachant qu'alors on peut se passer de la règle d'universalité.

Proposition 13

Étant donné un séquent $\Gamma \vdash [\varphi]_{\Lambda}$ dans lequel un indice i n'apparaît pas, et une preuve Π du séquent $\Gamma_{oi} \vdash [\varphi]_{\Lambda_{oi}}$ n'utilisant pas la règle Univ, il est possible de transformer Π en une preuve du séquent $\Gamma \vdash [\varphi]_{\Lambda}$.

Preuve Ce résultat se prouve par induction sur la hauteur de la preuve, en raisonnant par cas sur la règle utilisée à la racine.

Si la règle en question est *Axiom*, le séquent est de la forme $[\varphi]_{\Lambda_{oi}} \vdash [\varphi]_{\Lambda'_{oi}}$ avec $\Lambda' \circ i \leq_{\star} \Lambda \circ i$. Deux cas se présentent alors. Si $\Lambda' = \varepsilon$, comme alors $i \leq_{\star} \Lambda \circ i$, cela implique que $\Lambda = \varepsilon$ (puisque $\Lambda \circ i$ commence par i , et que Λ n'en contient pas). Si au contraire $\Lambda' \neq \varepsilon$, alors il en découle que $\Lambda' \leq_{\star} \Lambda$. Dans les deux cas, on peut donc transformer la preuve en :

$$\frac{\text{Axiom}}{[\varphi]_{\Lambda} \vdash [\varphi]_{\Lambda'}}$$

qui est correcte. Les règles \perp , *Weak* et *Contract* se traitent sans problème.

Penchons-nous maintenant sur la règle *Loc*. Reprenant les notations de la définition, nous partons d'une règle de la forme :

$$\frac{\Gamma_{oi}, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''_{oi}}}{\Gamma_{oi}, [\varphi]_{\Lambda_{oi}} \vdash [\psi]_{\Lambda''_{oi}}} \text{Loc}$$

avec $\Lambda' \leq_{\star} \Lambda \circ i$ et $\Lambda'' \circ i \leq \Lambda'$ d'après la proposition 4 p.80. Mais alors, de $\Lambda'' \circ i \leq \Lambda'$ il découle que Λ' contient au moins une occurrence de l'indice i , et comme $\Lambda' \leq \Lambda \circ i$ avec i

n'apparaissant pas dans Λ , on en déduit que Λ' est équivalent à un mot de la forme $\Lambda'_1 \circ i$ où i n'apparaît pas dans Λ'_1 et tel que $\Lambda'_1 \leq_* \Lambda$ (le raisonnement est la même que pour le cas de la règle *Axiom*). On conclut en appliquant la récurrence à $\Gamma_{oi}, [\varphi]_{\Lambda'_1 \circ i} \vdash [\psi]_{\Lambda'' \circ i}$ et en appliquant à nouveau la règle *Loc* pour passer de Λ'_1 à Λ .

Enfin, les règles d'introduction de connecteurs se traitent elles-aussi sans problèmes. \square

Cette proposition va maintenant être utilisée pour éliminer les instances de la règle *Univ* dans le cas où un indice i n'apparaît pas dans la conclusion d'une preuve.

Proposition 14

*Étant donné une preuve sans coupure Π d'un séquent $\Gamma \vdash [\varphi]_{\Lambda}$ dans lequel n'apparaît pas un indice i donné, il est possible de transformer Π en une autre preuve Π' du même séquent n'utilisant pas la règle d'universalité *Univ*.*

Preuve Pour prouver ce résultat, on procède par récurrence sur la longueur minimale de la preuve sans coupure d'un séquent donné, et l'on procède par cas suivant la règle utilisée à la racine.

Bien évidemment, le seul cas à développer est celui où la règle utilisée à la racine est alors *Univ*. Dans ce cas, on a par récurrence une preuve Π_i de $\Gamma_{oi} \vdash [\varphi]_{\Lambda \circ i}$ n'utilisant pas la règle *Univ*. On applique alors la proposition précédente pour obtenir une preuve Π'_i de $\Gamma \vdash [\varphi]_{\Lambda}$ ne contenant pas non plus d'instance de cette règle. \square

Nous allons maintenant nous pencher sur la dernière proposition de cette partie, qui permet d'utiliser le système de preuves en calcul des séquents même dans le cas d'ensemble d'indices infini, puisqu'alors on peut se passer de la règle d'universalité qui sinon poserait un problème.

Proposition 15

*Pour tout ensemble d'indices \mathcal{I} est infini, pour toute proposition $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$ telle que $\vdash_{\text{OL}} \varphi$, il existe une preuve de $\emptyset \vdash [\varphi]_{\varepsilon}$ n'utilisant pas la règle d'universalité *Univ*.*

Preuve Puisque \mathcal{I} est infini et $\text{ind}(\varphi)$ est fini, il existe un indice i dans $\mathcal{I} \setminus \text{ind}(\varphi)$. Alors, on en déduit grâce au théorème 7 qui $\vdash_{\text{OL}, \text{ind}(\varphi) \cup \{i\}} \varphi$. Alors, on en déduit qu'il existe une preuve Π de $\emptyset \vdash [\varphi]_{\varepsilon}$ sans utilisation de la règle *Univ* et pour l'ensemble d'indices $\text{ind}(\varphi) \cup \{i\}$. Or, comme cet ensemble est inclus dans \mathcal{I} , et que justement la règle *Univ* n'a pas été utilisée, on en déduit que Π est aussi une preuve du séquent $\emptyset \vdash [\varphi]_{\varepsilon}$ pour l'ensemble d'indices \mathcal{I} . \square

Ce chapitre ne fournit qu'un rapide survol de propriétés préliminaires vérifiées par la logique OL. Il existe de nombreux autres aspects qui pourraient être développés. Tout d'abord, on

pourrait s'interroger sur l'extension de OL au premier ordre. Il conviendrait alors de se poser un certain nombre de questions sur la marche à suivre pour faire cela, en gardant à l'esprit la façon donc cela pourrait s'interpréter en terme d'observations partielles et de multiples observateurs. En particulier, deux voies se dessinent, suivant que les valeurs que prendraient les variables dépendent ou non de la description choisie, et plus encore de la représentation choisie. De même, comme il est possible d'ajouter des quantifications pour les propositions, on pourrait envisager de quantifier aussi les indices de représentations. Cela pourrait permettre d'inclure explicitement au sein de la logique des informations sur la structure du système de représentation.

Un autre développement possible est l'inclusion d'autres opérateurs épistémiques. En particulier, dans [23], sont définis des opérateurs permettant de traiter la connaissance distribuée sur un ensemble d'agents (une information distribuée est telle qu'elle peut être reconstituée en mettant en commun les connaissances des différents agents), la connaissance partagée (précisant que tous membres d'un ensemble d'agents donnés connaissent une information particulière), ou encore la connaissance de sens commun où tous les agents connaissent une information φ , savent que les autres connaissent φ , savent que les autres savent que les autres connaissent φ , etc.

Cependant, nous gardons ces pistes pour des développements ultérieurs. Pour le moment, la logique OL est suffisante pour l'étude que l'on veut en faire. En effet, puisqu'elle émane du point de vue selon lequel la connaissance est basée sur des observations, seuls les opérateurs K_i tels que nous les avons définis sont utiles. De même, les quantificateurs ne sont pas nécessaires, puisque l'on se contente d'étudier les différentes informations que l'on peut avoir à partir de résultats d'observations donnés, et l'on n'effectue pas de généralisation à partir de cela.

Nous allons donc dans le chapitre suivant entamer une interprétation des résultats provenant de l'étude de la logique OL, pour exhiber certaines propriétés vérifiées par la connaissance dans l'approche que nous avons formalisée.

Chapitre 8

Quelques conséquences de la logique OL

8.1 La double négation intuitionniste

L'une des principales différences entre la logique classique et la logique intuitionniste est que pour cette dernière, la double négation n'est pas équivalente à l'identité. Elle en est juste une conséquence. Ainsi, on peut considérer cette double négation comme un nouvel opérateur logique unaire. Quelle est sa signification ?

Dans un premier temps, en utilisant sa définition et la logique intuitionniste, il est possible d'identifier un premier aspect de son comportement. Pour cela, considérons les assertions suivantes, qui sont facilement prouvables :

$$\begin{aligned} \vdash_{\mathbf{IL}} \varphi &\rightarrow \neg\neg\varphi \\ \vdash_{\mathbf{IL}} \neg\neg\varphi &\leftrightarrow \neg\neg\neg\neg\varphi \\ \vdash_{\mathbf{IL}} (\varphi &\rightarrow \psi) \rightarrow (\neg\neg\varphi \rightarrow \neg\neg\psi) \end{aligned}$$

Si l'on définit un pré-ordre \preceq par $\varphi \preceq \psi \Leftrightarrow \vdash \varphi \rightarrow \psi$, les propositions précédentes nous montrent que l'opérateur $\neg\neg$ agit comme un opérateur de clôture supérieure, puisqu'elles correspondent respectivement à l'extensivité, l'idempotence et la croissance.

Nous avons déjà rencontré des opérateurs de cette sorte dans les parties précédentes. En particulier, ils incarnaient dans la partie 5.3.2 l'idée d'approximation entre deux représentations. On peut donc penser qu'ici aussi, la double négation peut être associée à une forme d'abstraction sur les formules de notre langage logique.

Une application intéressante de cet opérateur apparaît dans le cas d'une formule de la forme $\varphi \vee \neg\varphi$. En effet, on a :

$$\forall \varphi, \vdash_{\mathbf{IL}} \neg\neg(\varphi \vee \neg\varphi)$$

Ainsi, une proposition $\neg\neg(\varphi \vee \neg\varphi)$ est vérifiée partout, quelque soit la description que l'on a de

l'état du système, donc on peut prouver cette proposition même si l'on n'a aucune information sur cet état (ce qui arrive avec la description \top).

Il existe une corrélation entre les descriptions vérifiant une proposition $\neg\neg\varphi$ et l'information que l'on possède pour pouvoir prouver φ . Avant de préciser cette relation, nous allons développer un exemple qui servira d'illustration.

8.1.1 Le protocole de transfert de bit

Cet exemple, tiré de [23], est la modélisation du protocole de transfert de bit entre deux personnes. Il s'agit tout d'abord d'un agent – appelé émetteur et noté *Send* – qui émet un bit b à un autre agent, le récepteur (noté *Recv*). Le protocole permet de s'assurer que le récepteur a de bonnes chances de recevoir la valeur b , même dans le cas où les communications sont difficiles. Il se déroule de la façon suivante : l'émetteur commence à émettre la valeur b de façon répétée. Il est possible qu'un bit émis n'arrive pas jusqu'au récepteur, ce qui explique que la valeur soit envoyée plusieurs fois. Dès que le récepteur reçoit le bit, il se met à envoyer à l'émetteur des accusés de réception, symbolisés par un message *Ack*. Dès que l'émetteur reçoit un accusé de réception, il arrête d'émettre la valeur b .

Ce protocole peut être formalisé par la hiérarchie de représentation donnée dans la figure 8.1. Les deux indices *Send* et *Recv* correspondent respectivement à la connaissance de l'émetteur et du récepteur. Un troisième indice – *Protocol* – symbolise l'ensemble du système. On peut remarquer que la représentation associée au protocole est le produit réduit des représentations associées à *Send* et à *Recv*.

La représentation correspondante, en plus de la description vide \top , est composée de deux séries de descriptions, correspondant aux deux valeurs possibles du bit à transmettre. Chacune de ces séries est constituée de trois descriptions, pour les différentes étapes du protocole.

La première description, celle au niveau *Send*, correspond à la première étape, où l'émetteur envoie de façon répétée le bit à transmettre. Puis, lorsque le récepteur obtient cette valeur, le système peut être décrit par une description du niveau *Recv*. À partir de ce moment, ce dernier se met à émettre régulièrement des accusés de réception. Le niveau *Ack* représente la fin du protocole, lorsque l'émetteur a reçu l'un de ces accusés.

Le déroulement de ce protocole peut être vu comme un échange successif d'informations entre les deux agents qui représentent les sous-systèmes. En particulier, la seule évolution de l'état de ces sous-systèmes est l'acquisition d'information. Il n'y a ni perte, ni transformation d'information. Aussi, la connaissance sur l'état général du système est croissante au cours du déroulement de ce protocole.

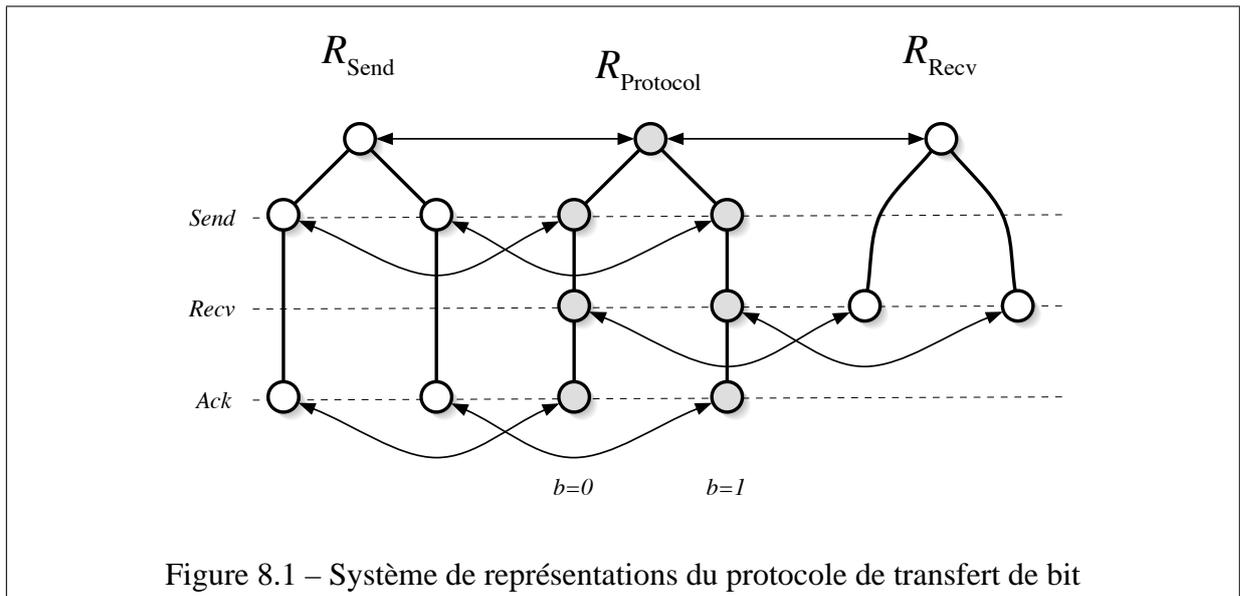


Figure 8.1 – Système de représentations du protocole de transfert de bit

Cet exemple va nous permettre d’illustrer le comportement de l’opération de double négation. On se souvient que l’un des postulats que l’on avait énoncés pour justifier l’utilisation des représentations est que l’on suppose que le système que l’on représente est dans un état bien défini, et qu’une description n’est que partielle vis-à-vis de cet état. Dans le cas de ce protocole, l’état du système est constitué par l’ensemble du déroulement du protocole, et ne correspond pas au système à un moment donné. Au contraire, l’état ne dépend pas du temps, et les différentes descriptions que l’on peut en faire, si elles dépendent du temps, découlent néanmoins de cet état “intemporel”.

Si par exemple nous nous intéressons à la proposition Ack_0 qui correspond aux situations où l’on est sûr que l’émetteur a reçu un accusé de réception pour un bit de valeur 0. Il en découle que son interprétation pour la représentation $\mathcal{R}_{\text{Protocol}}$ est le singleton $\{\langle Ack, b = 0 \rangle\}$. Dans ce cas, l’interprétation de la proposition $\neg\neg Ack_0$ est :

$$\{\langle Send, b = 0 \rangle, \langle Recv, b = 0 \rangle, \langle Ack, b = 0 \rangle\}$$

Il s’agit des états où il est inéluctable que dans le futur, on arrivera dans un état où la proposition Ack_0 sera vérifiée (puisque il faut que le récepteur connaisse la valeur 0 et que l’émetteur le sache), mais la description actuelle ne vérifie pas cette propriété (par exemple, pour l’état correspondant à la description $\langle Send, b = 0 \rangle$, la valeur n’a pas encore été correctement communiquée au récepteur). Ainsi, si la propriété $\neg\neg Ack_0$ est vérifiée, mais que Ack_0 n’est pas vérifiée, ou plus précisément n’est pas prouvée, c’est uniquement par manque d’information. Une fois que le protocole aura évolué suffisamment longtemps, on est sûr en effet qu’elle pourra être prouvée.

8.1.2 Interprétation de la double négation

Cet exemple permet de préciser l'interprétation que l'on peut donner à l'opération de double négation. Une première interprétation pourrait être que si pour un état donné, la propriété $\neg\neg\varphi$ est prouvable, alors φ le sera dans le futur. Dans l'exemple précédent, si à un moment donné on vérifie $\neg\neg\varphi$, alors à la fin de l'évolution du protocole, φ sera vérifiée.

Une autre interprétation possible est de dire que si une proposition $\neg\neg\varphi$ est vérifiée, alors la proposition φ est "forcément" vérifiée par le système, c'est-à-dire que si une description d est dans l'interprétation de $\neg\neg\varphi$, alors toute description plus précise $d' \leq d$ vérifie :

$$\begin{aligned} d' &\notin \llbracket \neg\varphi \rrbracket \\ \exists d'' \leq d' : d'' &\in \llbracket \varphi \rrbracket \end{aligned}$$

Ainsi, en partant d , il ne sera jamais possible de prouver $\neg\varphi$ en gagnant de l'information, et au contraire il est possible d'arriver à prouver φ en ajoutant suffisamment d'informations. Cependant, on ne possède pas suffisamment d'informations pour pouvoir prouver φ . Cette interprétation rejoint l'interprétation temporelle précédente, particulièrement dans le cas de l'exemple que l'on a développé, puisque l'évolution temporelle se confond avec le gain d'informations. Cependant, le temps n'apparaît pas toujours dans la description d'un système, et cette seconde interprétation nous semble meilleure. Dans l'acquisition progressive d'informations concernant l'état du système, une fois que l'on peut prouver $\neg\neg\varphi$, on pourra le faire jusqu'à la fin, et à partir d'un moment, c'est-à-dire une fois que l'on a suffisamment d'information, on pourra prouver φ .

Cette interprétation est naturelle pour l'exemple donné en début de section, où φ est de la forme $\psi \vee \neg\psi$. En effet, on peut montrer que toutes les descriptions vérifient $\neg\neg(\psi \vee \neg\psi)$, ce qui fait que, si l'on considère la description vide, on peut affirmer que le système vérifie soit ψ , soit $\neg\psi$, mais que l'on ne peut pas forcément dire laquelle des deux propositions est vérifiée (ce qui reviendrait à prouver la proposition $\psi \vee \neg\psi$).

Ainsi, on peut interpréter l'opération de double négation comme servant à exprimer le fait qu'une propriété est forcément vraie, mais que l'on ne possède pas suffisamment d'informations pour prouver ce fait.

8.1.3 Lien avec la logique classique

L'existence de la double-négation en tant qu'opérateur à part entière dans la logique intuitionniste (c'est-à-dire différent de l'identité) fait qu'elle joue un rôle crucial dans les rapports entre la logique classique et la logique intuitionniste. Il existe en effet de nombreuses méthodes

de traduction de formules logiques $\varphi \rightsquigarrow \widehat{\varphi}$ équivalentes classiquement ($\vdash_{\text{CL}} \varphi \leftrightarrow \widehat{\varphi}$) et telles que :

$$\vdash_{\text{CL}} \varphi \Leftrightarrow \vdash_{\text{IL}} \widehat{\varphi}$$

Un certain nombre de ces méthodes de traduction repose sur l'introduction de négations et de doubles négations (pour bénéficier en particulier de l'équivalence intuitionniste $\neg\neg\neg\varphi \leftrightarrow \neg\varphi$). Nous présentons en guise d'exemple la fonction de traduction $\varphi \rightarrow \varphi^\kappa$ due à Kolmogorov :

$$\begin{aligned} \psi^\kappa &= \neg\neg\psi & \psi &\in \Psi \\ (\neg\varphi)^\kappa &= \neg\varphi^\kappa \\ (\varphi \star \psi)^\kappa &= \neg\neg(\varphi^\kappa \star \psi^\kappa) & \star &\in \{\vee, \wedge, \rightarrow\} \end{aligned}$$

Une présentation de ce genre de traduction peut être trouvée dans [27]. On peut retenir que la relation entre les deux logiques basée sur ces transformations illustre l'interprétation que l'on a donnée de la double négation. Ainsi, une proposition φ prouvable classiquement correspond à une assertion de la forme " *φ est forcément vraie, il ne peut pas en être autrement, mais je ne peux pas forcément le prouver d'après mes observations*". Au contraire, si elle est prouvable de façon intuitionniste, elle découle des observations.

On voit donc que le fait que la logique OL est basée sur la logique intuitionniste traduit explicitement la prépondérance que l'on a donné aux observations comme éléments de base de la connaissance dans ce travail.

8.2 L'axiome 5

Nous allons maintenant nous intéresser à un axiome très présent et utilisé dans le domaine de l'étude logique de la connaissance. Il s'agit de l'axiome 5, qui est généralement intitulé *axiome d'introspection négative*. Sa signification en logique épistémique est qu'un agent sait les informations qu'il ne sait pas, ou plus précisément, dans le contexte de l'observation d'un système, s'il est impossible qu'un agent puisse savoir une certaine information φ , alors le même agent peut prouver cette impossibilité. Formellement, on l'exprime de la façon suivante :

$$\mathbf{5} : \neg K_i \varphi \rightarrow K_i \neg K_i \varphi$$

On trouve l'usage de cet axiome dans une grande part de la littérature concernant l'étude logique de systèmes multi-agents, comme par exemple dans [23]. On l'a vu précédemment, cet axiome n'est pas vérifié en général par les systèmes de représentations, puisqu'il n'est pas une conséquence de la logique OL. Ainsi, en général, l'inclusion suivante n'est pas vérifiée :

$$\llbracket \neg K_i \varphi \rrbracket_j \subseteq \llbracket K_i \neg K_i \varphi \rrbracket_j$$

La figure 8.2 en donne un contre-exemple. Elle représente une hiérarchie de représentations à deux indices (avec $1 < 2$) où l'on exhibe une propriété dont l'interprétation ne vérifie pas l'axiome 5.

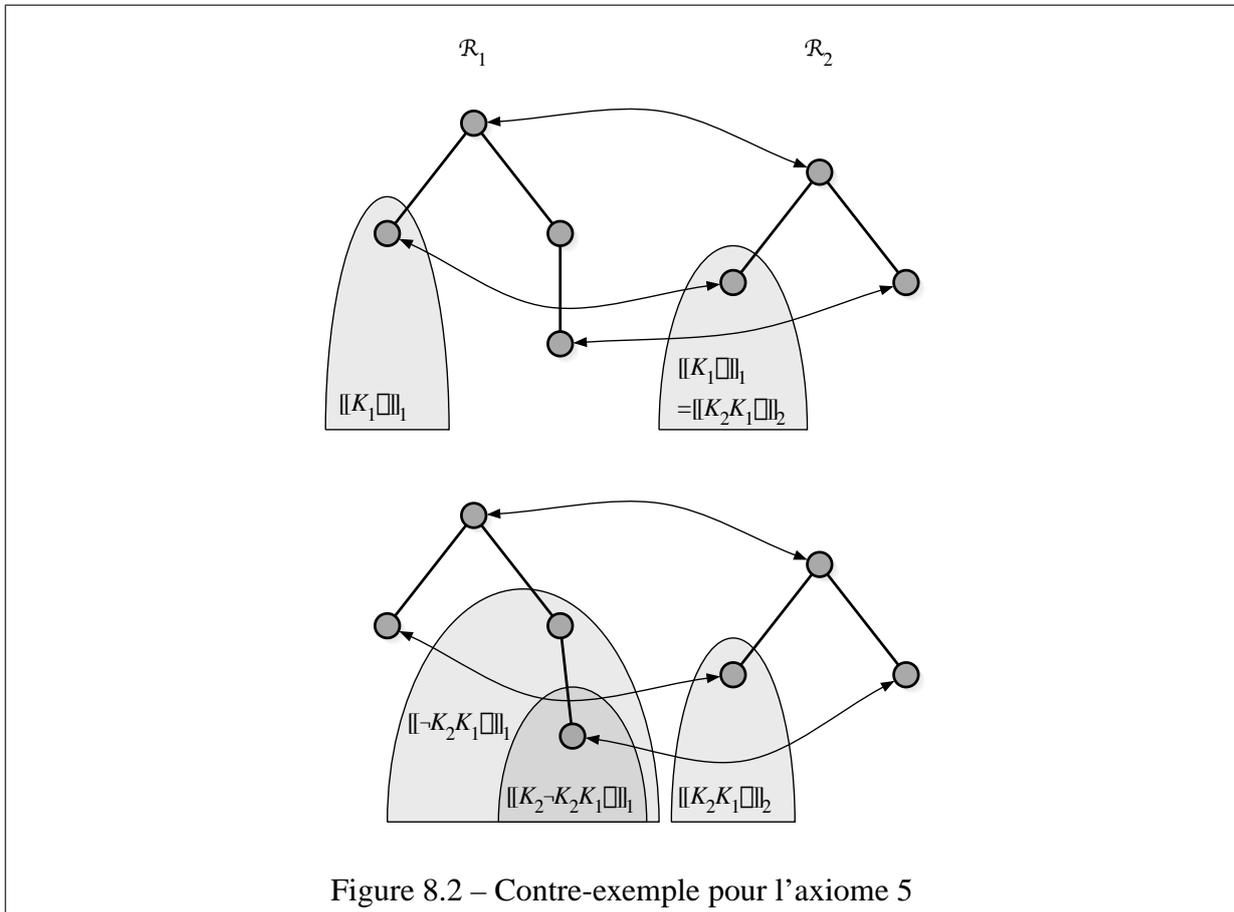


Figure 8.2 – Contre-exemple pour l'axiome 5

Cet exemple peut être développé pour illustrer la raison pour laquelle il est difficilement envisageable de pouvoir modifier la structure des hiérarchies de représentations de façon à vérifier l'axiome 5. En effet, la figure 8.2 peut être vue comme une représentation simplifiée d'une partie de pile ou face.

Dans ce jeu, supposons que l'on a deux joueurs, Alice et Bernard. Alice tient une pièce de monnaie dans sa main, et la lance. Le résultat est soit pile, soit face. Dans un premier temps, Alice regarde le résultat sans le communiquer à Bernard. C'est seulement dans un deuxième temps qu'il connaîtra le résultat. Une représentation plus complète du jeu est donnée dans la figure 8.3. La représentation de gauche correspond à Alice. Dans cette représentation, les descriptions P et F correspondent respectivement au résultat *Pile* ou *Face* de la pièce. À chacune de ces deux descriptions est associée une autre qui correspond au même résultat plus le fait qu'il a été communiqué à Bernard. Ainsi, la description Fc signifie que le résultat est *Face*, et

que Bernard le sait. La représentation de droite correspond à la connaissance de Bernard, et les seules descriptions présentes (en plus de \top) sont équivalentes aux descriptions Pc et Fc de la représentation d'Alice.

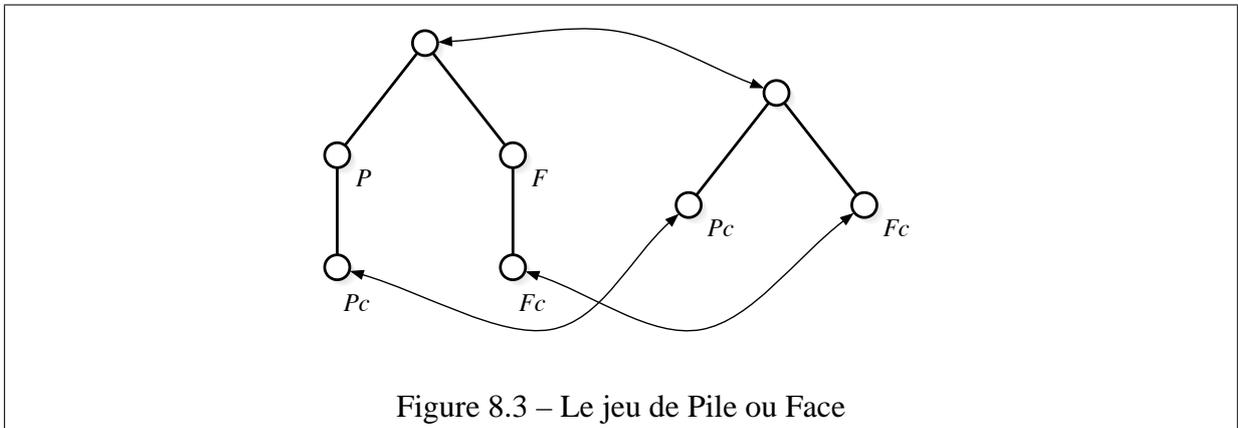


Figure 8.3 – Le jeu de Pile ou Face

D'un point de vue logique, on peut considérer que l'on a deux propositions atomiques : Pile dont l'interprétation est $\{P, Pc\}$, et Face ($\{F, Fc\}$). De simples calculs en se basant sur ces interprétations permet d'avoir :

$$\begin{aligned} \text{Pile} &\leftrightarrow \neg \text{Face} \leftrightarrow \neg \neg \text{Pile} \leftrightarrow \neg K_B \text{Face} \\ \text{Face} &\leftrightarrow \neg \text{Pile} \leftrightarrow \neg \neg \text{Face} \leftrightarrow \neg K_B \text{Pile} \\ \llbracket K_B \text{Pile} \rrbracket_A &= \{Pc\} \quad \llbracket K_B \text{Face} \rrbracket_A = \{Fc\} \end{aligned}$$

Reprenons maintenant le cours d'une partie de pile ou face. Tout d'abord, Alice lance la pièce et regarde le résultat. À ce moment, Bernard ne connaît pas le résultat. Supposons que le résultat soit *Pile*. Ainsi, la description de l'état du système est donnée par P . Or, d'après les calculs ci-dessus, on a $P \in \llbracket \text{Pile} \rrbracket_A$, donc la propriété $\neg K_B \text{Face}$ est vérifiée par le système. C'est prévisible, puisque si le résultat du lancer de la pièce est *Pile*, Alice n'a pas communiqué à Bernard le résultat *Face*. Mais la proposition $K_B \neg K_B \text{Face}$ n'est pas vérifiée, puisque Bernard n'a encore aucune information sur le résultat du lancer de la pièce. Si c'était le cas, on aurait alors $K_B \text{Pile}$ qui lui est équivalente.

Comme on peut le voir dans cet exemple, l'axiome 5 revient à autoriser l'acquisition d'informations sans qu'il y ait interaction avec ce que l'on observe, sans qu'il y ait une modification de l'état interne de l'agent. On peut relier cet axiome avec le fait que la règle du tiers-exclu n'est pas vérifiée, puisque dans les deux cas il est question de connaissance accessible ou non suivant que l'on a fait une observation ou non. Nous allons voir qu'il est possible de relier ces deux notions de façon plus directe.

8.2.1 Lien avec le tiers-exclu

La double négation permet d'exprimer un fait qui est inéluctable, mais que l'on ne peut pas prouver faute d'avoir suffisamment d'informations. L'axiome du tiers-exclu, quand à lui, exprime le fait que la double négation est équivalente à l'identité, étant équivalente à :

$$\vdash \neg\neg\varphi \rightarrow \varphi$$

Ainsi, avec cet axiome, il n'y a plus de différence entre ce qui est vrai (ce que l'on peut prouver), et ce qui est forcément, c'est-à-dire les faits que l'on ne peut pas forcément prouver, mais qui ne peuvent être faux, ceux dont la négation est absurde. Or, dans le cadre de l'observation, cela revient à dire que si l'on ne peut observer la négation d'un fait, alors l'observation du fait lui-même est inéluctable, et le système se comporte alors comme si cette observation avait eu lieu. C'est précisément l'interprétation de l'axiome 5.

Plus précisément, on peut montrer que dans la logique OL, l'axiome KV qui exprime la distributivité de la disjonction \vee par rapport aux opérateurs de modalité implique que l'axiome 5 est une conséquence du tiers exclu. En effet, de $\vdash K_i \varphi \vee \neg K_i \varphi$, on peut déduire la validité de $\vdash K_i (K_i \varphi \vee \neg K_i \varphi)$, et donc en utilisant KV, on obtient :

$$\vdash K_i \varphi \vee K_i \neg K_i \varphi$$

De ce point de vue, l'axiome 5 peut être vu comme une forme affaiblie du tiers exclu qui ne s'appliquerait qu'à la connaissance interne d'un agent.

Si l'axiome 5 était vérifié, cela signifierait qu'un agent pourrait gagner de l'information sans effectuer aucune observation du système. L'information gagnée proviendrait simplement du fait que de l'information acquise ailleurs dans le système impliquerait des impossibilités qui seraient connues de l'agent.

Dans l'approche que nous avons choisie, un tel déroulement est impossible, on ne peut pas gagner d'information sans effectuer une observation. Ainsi, l'axiome 5 ne peut être vérifié, et si l'on conserve le même principe de base, il ne peut pas non plus le devenir en enrichissant la structure.

8.3 La confiance réciproque entre agents

Nous allons maintenant nous intéresser à d'autres propriétés, concernant les relations entre les connaissances de différents agents. Une propriété de ce genre et vérifiée par les systèmes de représentation et la logique OL est l'axiome que l'on a intitulé T_2 dont l'intitulé est :

$$T_2 : K_i K_j \varphi \rightarrow K_j \varphi$$

Son interprétation est la suivante : si un agent \mathcal{A} sait qu'un autre agent \mathcal{B} sait qu'une propriété φ est vraie, alors c'est effectivement le cas, \mathcal{B} sait effectivement que φ est vraie. Cependant, on peut remarquer que cet axiome ne donne aucune indication sur ce que peut dire l'agent \mathcal{A} de la propriété φ .

Or, on peut imaginer que les agents se font confiance entre eux, et considèrent que les connaissances qu'ont les autres agents sont fiables, puisque chaque agent connaît la structure du système étudié, ainsi que les informations accessibles à chacun. Une telle condition pourrait s'exprimer de la façon suivante :

$$\mathbf{C} : K_i K_j \varphi \rightarrow K_i \varphi \quad (8.1)$$

On peut tout d'abord s'apercevoir que l'axiome \mathbf{C} n'est pas prouvable dans \mathbf{OL} . Pour s'en convaincre, remarquons que le seul axiome de cette logique mettant en relation plusieurs opérateurs modaux est \mathbf{T}_2 , et qu'il n'autorise pas de supprimer la modalité la plus à droite, comme on voudrait le faire avec \mathbf{C} . Cette considération permet d'exhiber un contre-exemple simple. En effet, supposons que \mathcal{I} contient au moins deux indices i et j , et que l'interprétation d'une proposition atomique ψ vérifie : $\llbracket \psi \rrbracket_i = \emptyset$ et $\llbracket \psi \rrbracket_j = \mathcal{R}_j$. On a dans ce cas $\vdash K_j \psi$ et donc en utilisant la règle \mathbf{Nec} $\vdash K_i K_j \psi$, ainsi que $\vdash \neg K_i \psi$. Dans ce cas, on ne peut pas avoir $\vdash K_i K_j \psi \rightarrow K_i \psi$, puisque cela entraînerait $\vdash \perp$.

Ce contre-exemple nous montre que l'axiome \mathbf{C} n'est pas vérifié ici du fait que l'on peut choisir les interprétations atomiques de façon indépendante d'une représentation à l'autre. Mais peut-être est-il possible de modifier la structure des systèmes de représentations pour que \mathbf{C} soit vérifié par l'interprétation en résultant. Malheureusement, la proposition suivante nous montre que l'ajout de cet axiome reviendrait à affaiblir considérablement la puissance expressive des systèmes de représentations ou de la structure que l'on considérerait.

Proposition 16

L'axiome \mathbf{C} combiné à la logique de l'observation \mathbf{OL} a comme conséquence :

$$\forall i, j, \forall \varphi, K_i \varphi \leftrightarrow K_j \varphi$$

Preuve En effet, on a :

$$\begin{aligned} & \vdash K_i (\varphi \rightarrow K_i \varphi) && \mathbf{L}_4 \\ \Rightarrow & \vdash K_j K_i (\varphi \rightarrow K_i \varphi) && \mathbf{Nec.} \\ \Rightarrow & \vdash K_j (\varphi \rightarrow K_i \varphi) && \mathbf{C} \\ \Rightarrow & \vdash K_j \varphi \rightarrow K_j K_i \varphi && \mathbf{K} \\ \Rightarrow & \vdash K_j \varphi \rightarrow K_i \varphi && \mathbf{T}_2 \end{aligned}$$

□

Ainsi, une logique contenant l'axiome de confiance **C** et la logique de l'observation **OL** ne serait pas intéressante, puisqu'alors, tous les opérateurs de modalité seraient équivalents. La présence éventuelle de cet axiome n'est donc pas à considérer, et nous devons nous pencher sur des axiomes moins forts.

8.4 La consistance globale

Une autre notion concernant les différents observateurs d'un système est celle de consistance. L'un des axiomes définissant la logique **OL** est l'axiome **D** dont l'intitulé est :

$$\mathbf{D} : K_i \varphi \rightarrow \neg K_i \neg \varphi$$

Cet axiome permet de s'assurer que les connaissances de chaque agent sont consistantes (une formulation de cet axiome, équivalente en présence de **K**, est $\neg K_i \perp$). Cependant, il pourrait être souhaitable que ce principe de consistance soit étendu à l'ensemble du système et des observateurs. Ainsi, supposons qu'un agent \mathcal{A}_i puisse prouver une propriété φ (on a donc $\vdash K_i \varphi$). Alors, dans quelles conditions un autres agent pourrait-il prouver le contraire ? S'il est possible qu'un autre agent puisse prouver la négation de φ (auquel cas $\vdash K_j \neg \varphi$), alors il en découle qu'il n'existe aucune façon de relier les informations des différents observateurs en général. Pour éviter cela et renforcer la logique que l'on a définie, nous proposons de définir un axiome de consistance globale **GD** :

$$\mathbf{GD} : K_i \varphi \rightarrow \neg K_j \neg \varphi$$

Cet axiome est plus faible que l'axiome de confiance **C**, puisque l'on avait montré que $\vdash K_i \varphi \rightarrow K_j \varphi$ qui, avec l'axiome **D**, implique $\vdash K_i \varphi \rightarrow \neg K_j \neg \varphi$. On peut remarquer tout d'abord que **GD** n'est pas une conséquence de la logique observationnelle **OL**, et l'exemple suivant montre que sa présence peut conduire à une inconsistance.

Exemple

Reprenons la structure présentée dans la section précédente. On avait deux indices i et j , ainsi qu'une proposition atomique ψ vérifiant $\llbracket \psi \rrbracket_i = \emptyset$ et $\llbracket \psi \rrbracket_j = \mathcal{R}_j$. Dans cette situation, on a $\vdash K_j \psi$ et $\vdash K_i \neg \psi$. Mais alors, l'adjonction de l'axiome **GD** permettrait de déduire $\vdash \neg K_i \neg \psi$ de $\vdash K_j \psi$ et donc on peut aboutir à $\vdash \perp$.

Cet exemple nous montre une limite du formalisme des systèmes de représentations tel qu'on l'a défini. Les relations qui existent entre les différentes observations sont très pauvres,

puisque la seule façon d'obtenir de telles relations est d'utiliser l'axiome \mathbf{T}_2 . Ainsi, une forme de consistance globale que l'on obtient avec \mathbf{OL} peut s'exprimer par :

$$K_i \varphi \rightarrow \neg K_j K_i \neg \varphi$$

Mais on ne peut pas en général relier $K_i \varphi$ et $K_j \varphi$. La raison fondamentale est la suivante : dans notre formalisme, on n'impose aux interprétations aucune condition sur les relations entre les ensembles $\llbracket \varphi \rrbracket_i$ et $\llbracket \varphi \rrbracket_j$. Ainsi, il n'y a aucun lien entre les deux, et donc entre les interprétations des formules $K_i \varphi$ et $K_j \varphi$.

Nous avons jusqu'à présent étudié la connaissance basée sur des observations partielles en définissant des structures algébriques adaptées pour modéliser ce type d'information, puis en déduisant leur comportement logique à partir de celles-ci. Nous avons ainsi défini la logique modale \mathbf{OL} . Mais nous nous sommes basé sur des structures très générales, et il en découle que les liens existant entre les différentes méthodes d'observation du système étudié sont très faibles, se limitant en fait à l'application de l'axiome \mathbf{T}_2 comme on vient de le voir.

Nous allons dans la suite renforcer les liens existant entre les différentes représentations en modifiant les structures utilisées. Pour cela, nous ajouterons les conditions supplémentaires entre les représentations, et nous introduirons une notion de comparaison entre celles-ci. Nous étudierons alors comment cela se traduit au niveau de la logique modélisée.

Troisième partie

Hiérarchies de représentations

Les systèmes de représentation que nous avons introduits dans la partie précédente sont une structure très générale pour modéliser l'étude d'un système dans le cas où existent plusieurs façons de l'observer. Cependant, comme nous l'avons vu, les relations existant entre ces différents points de vue sont très pauvres, et ne permettent pas de relier les informations qui leur sont associées de façon très intéressante.

Pour remédier à cette faiblesse, nous allons étudier une méthode permettant de renforcer cette structure, en définissant les *hiérarchies de représentations*. Pour cela, nous allons reprendre l'étude faite sur la possibilité de définir une représentation comme l'approximation d'une autre. Cela se traduira par la définition d'une propriété portant sur les fonctions de transformation entre ces deux représentations. Comme on le verra, cette propriété peut être vue comme une relation de pré-ordre entre les représentations. Aussi, cette discussion sera l'occasion de donner une définition formelle de la notion de précision ou d'expressivité d'une représentation évoquée dans l'introduction.

En utilisant cet ordre sur les représentations, nous allons imposer une condition précisant qu'il existe toujours un "moyen" de passer d'une représentation à une autre en passant par une représentation plus précise. C'est cette condition (plus une autre concernant la complétude des représentations) que l'on utilisera pour définir les hiérarchies de représentations. Avec cette nouvelle structure, les relations entre les différents points de vue sont plus fortes que dans le cas d'un système de représentation, puisqu'il existe alors toujours une représentation plus précise qu'une paire de représentations donnée.

Cela donnera de plus la possibilité de définir des outils flexibles pour manipuler les hiérarchies de représentations. Ainsi, les opérations de produit et de raffinage peuvent être étendues facilement au cas des hiérarchies. Cette dernière opération est en particulier rendue possible par le fait qu'il est toujours possible de passer d'une représentation à une autre, ce qui implique que l'ajout d'information structurelle à une hiérarchie de représentation peut toujours être définie pour une représentation donnée, la structure d'ordre permettant de l'étendre ensuite à l'ensemble de la hiérarchie (Pour des raisons de lisibilité, le chapitre concernant ces constructions a été reporté dans les annexes).

Enfin, nous pourrions aborder une discussion sur la façon dont l'information accessible pour une représentation donnée se transforme en changeant de représentation. En particulier, nous allons pouvoir étudier formellement la possibilité d'inclure l'un des postulats de départ (celui précisant qu'une information reste accessible si l'on passe à une représentation plus précise), et cela nous conduira à mettre en évidence deux propriétés importantes concernant la façon dont un système est modélisé : la localisation qui précise quelles sont les représentations suffisamment précises permettant de décider de la validité d'une propriété, et la consistance qui impose que les approximations entre représentations ne font que perdre de l'information.

Chapitre 9

Vers les hiérarchies de représentation

9.1 Un ordre sur les indices

Avant de définir les hiérarchies de représentation, nous allons nous intéresser plus précisément à l'ensemble des indices d'un système de représentations. En effet, la caractérisation que nous avons donnée de l'approximation d'une représentation par une autre peut être reprise et étendue au cas d'un système de représentations. Ainsi, étant donné un tel système $\mathcal{S} = \langle \{A, B\}, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$, la représentation \mathcal{R}_B est considérée comme une approximation de \mathcal{R}_A si la composition de fonctions $f_{B|A} \circ f_{A|B}$ est égale à l'identité de \mathcal{R}_B . On voit aisément que cette relation d'approximation entre représentations peut être étendue à un système dans son ensemble. Pour cela, définissons une relation binaire \triangleleft entre les indices de la façon suivante :

$$\forall i, j \in \mathcal{I}, i \triangleleft j \Leftrightarrow f_{j|i} \circ f_{i|j} = \text{id}_j$$

Ainsi, on voit qu'une représentation \mathcal{R}_j est une approximation de \mathcal{R}_i si et seulement si $i \triangleleft j$. Comme la notation le suggère, la relation que l'on vient de définir est un pré-ordre sur l'ensemble des indices \mathcal{I} . On a besoin pour cela de montrer que deux propriétés, la réflexivité et la transitivité, sont vérifiées. La réflexivité vient directement du fait que comme $f_{i|i} = \text{id}_i$, on a $f_{i|i} \circ f_{i|i} = \text{id}_i$. Pour montrer la transitivité, on procède de la façon suivante : supposons que l'on a trois indices i, j, k tels que $i \triangleleft j$ et $j \triangleleft k$. On a alors les inégalités suivantes :

$$\begin{aligned} \text{id}_k &= f_{k|k} \\ &\leq f_{k|i} \circ f_{i|k} \\ f_{k|i} \circ f_{i|k} &\leq f_{k|j} \circ (f_{j|i} \circ f_{i|j}) \circ f_{j|k} \\ &= f_{k|j} \circ f_{j|k} && \text{car } i \triangleleft j \\ &= \text{id}_k && \text{car } j \triangleleft k \end{aligned}$$

On a donc montré que $f_{k|i} \circ f_{i|k} = \text{id}_k$, et que $i \triangleleft j$ et $j \triangleleft k$ impliquent $i \triangleleft k$. La définition de ce pré-ordre sur les indices va permettre de structurer les systèmes de représentations en ajoutant

une dimension “verticale” aux représentations. Ainsi, de même qu’au sein d’une représentation \mathcal{R}_i , plus on descend dans l’ordre correspondant \leq_i plus on a des descriptions précises, dans le cas des indices de \mathcal{I} , plus l’on descend pour \triangleleft , plus on a des représentations expressives.

Nous allons utiliser cet ordre comme élément central de la définition de la prochaine structure algébrique. Cependant, nous allons affaiblir un peu son rôle, pour ne garder comme propriété que :

$$\forall i, j \in \mathcal{I}, i \triangleleft j \Rightarrow f_{j|i} \circ f_{i|j} = \text{id}_j \quad (9.1)$$

Cela vient du fait qu’au cours des manipulations qui pourront être faites sur ces structures, il pourra arriver que l’on ait l’égalité à droite de l’implication de façon “accidentelle”, c’est-à-dire sans que cela soit voulu, et donc en particulier sans qu’il n’y ait de répercussions sur l’ordre \triangleleft .

9.1.1 Implications sur la logique

L’existence d’un ordre sur les indices entraîne une relation dans la logique correspondante sur les opérateurs de modalité. En effet, soit deux indices tels que $i \triangleleft j$. On a par définition $f_{j|i} \circ f_{i|j} = \text{id}_j$. On peut alors comparer les interprétations des propositions $K_i K_j \varphi$ et $K_j \varphi$:

$$\begin{aligned} \llbracket K_i K_j \varphi \rrbracket_k &= \{ d \mid f_{j|i} \circ f_{i|k}(d) \in \llbracket \varphi \rrbracket_j \} \\ \llbracket K_j \varphi \rrbracket_k &= \{ d \mid f_{j|k}(d) \in \llbracket \varphi \rrbracket_j \} \end{aligned}$$

Tout d’abord, comme l’indique l’axiome \mathbf{T}_2 : $K_i K_j \varphi \rightarrow K_j \varphi$, l’inclusion $\llbracket K_i K_j \varphi \rrbracket_k \subseteq \llbracket K_j \varphi \rrbracket_k$ est vérifiée. Réciproquement, soit $d \in \llbracket K_j \varphi \rrbracket_k$, c’est-à-dire tel que $f_{j|k}(d) \in \llbracket \varphi \rrbracket_j$. On a :

$$f_{j|i} \circ f_{i|k}(d) \leq f_{j|i} \circ f_{i|j} \circ f_{j|k}(d) = f_{j|k}(d)$$

On en déduit donc que d appartient alors à $\llbracket K_i K_j \varphi \rrbracket_k$. Ainsi, on a montré que les deux ensembles $\llbracket K_i K_j \varphi \rrbracket_k$ et $\llbracket K_j \varphi \rrbracket_k$ sont égaux dans le cas où $i \triangleleft j$. Cela s’exprime sous forme logique par :

$$i \triangleleft j \Rightarrow \vdash K_i K_j \varphi \leftrightarrow K_j \varphi \quad (9.2)$$

L’ordre partiel \triangleleft , que l’on a défini comme une relation entre les fonctions de transformations entre les représentations, traduit le fait qu’une représentation est plus précise qu’une autre. Comme on le voit, en utilisant l’interprétation $\llbracket \cdot \rrbracket$, cet ordre se trouve traduit dans la logique par le fait que si une information φ est accessible dans une représentation \mathcal{R}_j (ce que l’on note $K_j \varphi$), alors ce fait est connu dans toute représentation plus précise \mathcal{R}_i avec $i \triangleleft j$.

9.2 Hiérarchies de représentations

Nous allons dans cette section définir la notion de *hiérarchie* de représentations, qui est une restriction de celle de système de représentation, mais choisie de façon adéquate pour être facile à manipuler. En particulier, nous voudrions avoir la possibilité de définir le produit et le raffinage sur cette structure.

9.2.1 Définition

Nous allons imposer une première condition sur le poset $\langle \mathcal{I}, \triangleleft \rangle$. En effet, comme nous l'avons indiqué au début de cette partie, nous voulons pouvoir considérer des structures telles qu'il est possible de construire un système de représentation en partant de représentations (ou de hiérarchies de représentations, comme nous le verrons) et en ajoutant des contraintes exprimant les relations que l'on exprime entre les différentes parties. Pour cela, il est important d'avoir un "moyen" de passer d'une représentation à une autre. Pour cela, nous allons imposer la condition suivante : pour tout couple de représentations \mathcal{R}_i et \mathcal{R}_j , il existe une troisième représentation \mathcal{R}_k approximée par les deux premières. Cette condition permettra par la suite d'assurer l'existence d'une représentation dans laquelle on peut formaliser les relations entre deux représentations quelconques. On peut l'écrire de la façon suivante :

$$\forall i, j \in \mathcal{I}, \exists k \in \mathcal{I} : k \triangleleft i \text{ et } k \triangleleft j \quad (9.3)$$

Cependant, nous n'imposons pas de condition supplémentaire, telle que l'existence d'un indice k maximum vérifiant cette condition, et encore moins son unicité.

Examinons maintenant comment définir et exprimer de façon simple un produit cartésien de deux représentations sous forme d'un système de représentations. À la base, on a deux représentations \mathcal{R}_A et \mathcal{R}_B , et on veut aboutir à un système en contenant trois : les deux représentations de départ \mathcal{R}_A et \mathcal{R}_B ainsi que leur produit $\mathcal{R}_{AB} = \mathcal{R}_A \otimes \mathcal{R}_B$, ainsi que les fonctions de transformation correspondant aux morphismes classiques d'un produit. On peut aisément caractériser les fonctions d'approximation (ainsi, on a $f_{A|AB}(\langle d_A, d_B \rangle) = d_A$). Mais comment définir la transformation inverse : si $d_A \in \mathcal{R}_A$, quelle est la valeur à donner à $f_{AB|A}(d_A)$? Il est clair que $f_{AB|A}(d_A)$ est de la forme $\langle d_A, \cdot \rangle$ mais alors, comme fournir en plus de d_A une description de l'état de la partie B constitue un apport d'information, on en déduit que $\forall d_B \in \mathcal{R}_B, \langle d_A, d_B \rangle \leq f_{AB|A}(d_A)$. La seule façon de pouvoir vérifier cette inégalité est d'imposer à \mathcal{R}_B l'existence d'une description maximale \top_B , ce qui permet de définir $f_{AB|A}(d_A) = \langle d_A, \top_B \rangle$. Par symétrie, il est logique d'imposer aussi à \mathcal{R}_A la présence d'une description maximale, et on en déduit qu'alors \mathcal{R}_{AB} possède elle aussi une description maxi-

male $\langle \top_A, \top_B \rangle$. Une condition nécessaire pour la définition d'un produit cartésien est donc la présence dans chaque représentation d'une description maximale \top .

Nous imposerons de plus deux conditions qui seront très utiles pour avoir une généralisation du raffinage. Il s'agit tout d'abord de la présence pour chaque représentation d'une opération de conjonction partielle telle qu'évoquée dans la section 4.1.2.0. Cela renforce le fait qu'une description peut être construite comme la mise en commun d'un ensemble de descriptions moins précises.

Nous allons enfin imposer une condition de fermeture sur les suites décroissantes des descriptions, qui stipule que toute suite décroissante bornée converge. Formellement, étant donné une suite croissante $\{x_i\}$ d'éléments d'une représentation \mathcal{R}_i , si l'on note $x \downarrow = \{y \mid y \leq_i x\}$, cette condition s'écrit :

$$\bigcap_j x_j \downarrow \neq \emptyset \Rightarrow \exists x_\infty \in \mathcal{R}_i : x_\infty \downarrow = \bigcap_j x_j \downarrow \quad (9.4)$$

Cette condition nous sera utile dans l'annexe D lorsque nous étendrons la notion de raffinage à l'aide d'opérateurs de clôture inférieure à une hiérarchie de représentation entière. On peut cependant remarquer que cette condition sur les représentations ne modifiera pas la logique. En effet, cette condition est vérifiée par tout treillis complet (puisque c'est une conséquence de la complétude pour l'ordre), et donc en particulier par les algèbres de Heyting complètes qui sont le modèle de base de la logique intuitionniste (voir la partie 2.2.0.0).

Nous pouvons maintenant définir une hiérarchie de représentations comme étant un système de représentations vérifiant toutes ces conditions. De façon formelle, on a donc :

Définition 1

Une hiérarchie de représentations est un système compositionnel de représentations

$$\mathcal{H} = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$$

où chaque représentation $\mathcal{R}_i = \langle R_i, \leq_i, \top_i, \wedge_i \rangle$ possède un élément maximum et une conjonction partielle, et tel que les équations 9.1, 9.3 et 9.4 soient vérifiées.

9.2.2 Approche équivalente

La définition que nous avons donné d'une hiérarchie de représentation est très proche de celle de systèmes de représentations, et utilise de façon similaire les fonctions $f_{i|j}$. Nous allons montrer qu'il est possible de définir ces structures en utilisant une autre approche, où l'on se contente de définir les relations entre représentations comparables.

Surjections de Galois

La structure que l'on impose à l'ensemble des indices d'une hiérarchie de représentations par le biais de l'équation 9.3 donne une grande importance à la notion d'approximation, et l'on peut étudier la présence des surjections de Galois dans ce contexte. Pour rappel, une surjection de Galois entre deux e.p.o. E et F est une paire de fonctions $\langle \alpha : E \rightarrow F, \gamma : F \rightarrow E \rangle$ telle que α est surjective, et que l'on a l'inégalité :

$$\forall x \in E, \forall y \in F, \alpha(x) \leq_F y \Leftrightarrow x \leq_E \gamma(y)$$

On peut reprendre le résultat donné en section 5.3.1 en l'exprimant avec les nouvelles notations de la façon suivante :

Proposition 17

Les fonctions de transformations entre deux représentations comparables pour \triangleleft forment une surjection de Galois : si l'on a deux indices tels que $i \triangleleft j$, alors la paire $\langle f_{j|i}, f_{i|j} \rangle$ est une surjection de Galois de \mathcal{R}_i vers \mathcal{R}_j .

Composition de surjections de Galois

De façon générale, si l'on a trois posets E, F, G et deux surjections de Galois $\langle \alpha_1, \gamma_1 \rangle : E \rightarrow F$ et $\langle \alpha_2, \gamma_2 \rangle : F \rightarrow G$, il est possible de les composer de manière à obtenir une surjection de E sur G en obtenant : $\langle \alpha_2 \circ \alpha_1, \gamma_1 \circ \gamma_2 \rangle$. En utilisant cette construction dans le cadre des hiérarchies de représentations, étant donnés des indices $i \triangleleft j \triangleleft k$, il est possible de définir deux surjections de Galois de \mathcal{R}_i vers \mathcal{R}_k . En effet, on a tout d'abord $\langle f_{k|i}, f_{i|k} \rangle$ puisque $i \triangleleft k$, et en utilisant les étapes de i vers j puis de j vers k , on a aussi $\langle f_{k|j} \circ f_{j|i}, f_{i|j} \circ f_{j|k} \rangle$. L'équation de compositionnalité 6.1 permet une première comparaison de ces fonctions, puisque par exemple $f_{k|i} \leq f_{k|j} \circ f_{j|i}$. On montre que dans le cas où les indices sont comparables, la relation est bien plus forte.

Proposition 18

Étant donné trois indices $i \triangleleft j \triangleleft k$, on a :

$$\begin{aligned} f_{i|k} &= f_{i|j} \circ f_{j|k} \\ f_{k|i} &= f_{k|j} \circ f_{j|i} \end{aligned}$$

Preuve Soit $y \in \mathcal{R}_k$, et $x = f_{i|j} \circ f_{j|k}(y)$. Par l'équation de composition 6.1, on a $f_{k|i}(x) \leq f_{k|j} \circ f_{j|i}(x)$, d'où :

$$x \leq f_{i|k} \circ f_{k|j} \circ f_{j|i}(x)$$

En remplaçant x par $f_{i|j} \circ f_{j|k}(y)$, et en simplifiant l'expression, on obtient :

$$f_{i|j} \circ f_{j|k}(y) \leq f_{i|k} \circ f_{k|j} \circ f_{j|i} \circ f_{i|j} \circ f_{j|k}(y) = f_{i|k}(y)$$

La première égalité en découle. Pour la seconde égalité, soit tout d'abord $x \in \mathcal{R}_i$ tel que $x = f_{i|k} \circ f_{k|i}(x)$. On a :

$$\begin{aligned} f_{k|i}(x) &= f_{k|i} \circ f_{i|k} \circ f_{k|i}(x) \\ &= f_{k|i} \circ f_{i|j} \circ f_{j|k} \circ f_{k|i}(x) \\ &= (f_{k|i} \circ f_{i|j}) \circ (f_{j|k} \circ f_{k|i})(x) \\ &\geq f_{k|j} \circ f_{j|i}(x) \end{aligned}$$

On en déduit donc que $f_{k|i}(x) = f_{k|j} \circ f_{j|i}(x)$ pour tout x vérifiant $x = f_{i|k} \circ f_{k|i}(x)$. Pour obtenir le cas général, il suffit de voir que $f_{k|i}(x) = f_{k|i}(f_{i|k} \circ f_{k|i}(x))$ et que $f_{i|k} \circ f_{k|i}(x) = f_{i|k} \circ (f_{k|i} \circ f_{i|k}) \circ f_{k|i}(x)$, ce qui permet de conclure. \square

Comme le montre cette proposition, il existe de fortes contraintes de composition sur les fonctions de transformations dans le cas de représentations comparables. Cela montre de plus qu'étant donné une hiérarchie de représentation, il est possible de définir une collection de surjections de Galois $\{\langle \alpha_{j|i}, \gamma_{i|j} \rangle \mid i \triangleleft j\}$ telle que si l'on a trois indices tels que $i \triangleleft j \triangleleft k$, alors $\langle \alpha_{k|i}, \gamma_{i|k} \rangle = \langle \alpha_{k|j}, \gamma_{j|k} \rangle \circ \langle \alpha_{j|i}, \gamma_{i|j} \rangle$. On pose pour cela, si $i \triangleleft j$, $\alpha_{j|i} = f_{j|i}$ et $\gamma_{i|j} = f_{i|j}$.

Il est possible de définir à l'inverse les $f_{i|k}$ en fonction de ces surjections de Galois, comme le montre la proposition qui suit :

Proposition 19

Soit i, j deux indices. On a :

$$\forall k, (k \triangleleft i \text{ et } k \triangleleft j) \Rightarrow f_{i|j} = \alpha_{i|k} \circ \gamma_{k|j}$$

Preuve On a en effet pour un tel k , les inégalités :

$$\begin{aligned} f_{i|j}(x) &\leq f_{i|k} \circ f_{k|j}(x) = \alpha_{i|k} \circ \gamma_{k|j}(x) \\ \alpha_{i|k} \circ \gamma_{k|j}(x) &= f_{i|k} \circ f_{k|j}(x) \leq f_{i|k} \circ f_{k|i} \circ f_{i|j}(x) = f_{i|j}(x) \end{aligned}$$

On en déduit l'égalité attendue. \square

Dans cette preuve, on se sert uniquement du fait que \mathcal{R}_i peut être vue comme incluse dans \mathcal{R}_k . Une des conséquences de cette proposition est que le choix de l'indice k ne joue aucun rôle. En effet, il suffit juste de prendre une représentation suffisamment fine pour contenir \mathcal{R}_i et \mathcal{R}_j , mais c'est la seule condition nécessaire.

Définition équivalente

On voit donc, puisque l'on peut traduire les fonctions de transformations en surjections de Galois et vice-versa, que l'on peut aussi bien définir les hiérarchies de représentations en

fonctions de surjections de Galois. En effet, la propriété importante ici est celle de fermeture de l'ensemble des indices exprimée par l'équation 9.3. On peut donc écrire :

Définition 2 (Hiérarchie de représentations, bis)

Une hiérarchie de représentations est un quadruplet :

$$\mathcal{H} = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i\}, \{ \langle \alpha_{j|i}, \gamma_{i|j} \rangle \mid i \triangleleft j \} \rangle$$

où chaque représentation $\mathcal{R}_i = \langle R_i, \leq_i, \top_i, \wedge_i \rangle$ possède un élément maximum et une conjonction partielle, tel que les conjonctions \wedge_i soient distributives par rapport aux fonctions $\alpha_{i|j}$ et $\gamma_{j|i}$, et tel que :

$$\begin{aligned} \forall i, j \in \mathcal{I}, \exists k \in \mathcal{I} : k \triangleleft i \text{ et } k \triangleleft j \\ \forall i \triangleleft j \triangleleft k, \langle \alpha_{k|i}, \gamma_{i|k} \rangle = \langle \alpha_{k|j}, \gamma_{j|k} \rangle \circ \langle \alpha_{j|i}, \gamma_{i|j} \rangle \end{aligned}$$

Les deux définitions d'une hiérarchie de représentations expriment deux aspects différents. Dans la première définition à l'aide de fonctions de transformation, on insiste sur l'aspect traduction d'une représentation à une autre, ce qui donne une vision horizontale de l'organisation de la hiérarchie. À l'opposé, la définition à base de surjection de Galois met en valeur une organisation verticale puisque ce qui est spécifié est la façon dont une représentation s'inclue dans une autre plus fine. Il est de plus intéressant de voir que l'inégalité de l'équation 6.1 de composition des fonctions de transformation se transforme en une égalité pour les surjections de Galois.

9.3 Comportement logique

Nous avons défini les hiérarchies de représentations en imposant des contraintes sur les systèmes de représentations compositionnels. La question qu'il faut se poser alors est de savoir si en faisant cela, nous n'avons pas modifié la logique, en validant des propriétés qui ne sont pas déductibles dans OL. Nous allons montrer dans cette partie que si l'ensemble d'indices \mathcal{I} est infini, alors il n'y a pas de modification de la logique modélisée. On obtient toujours OL. Cependant, si \mathcal{I} est fini, il y a une modification. En effet, dans ce cas, la condition imposée sur l'ordre \triangleleft impose qu'il existe un indice $\perp \in \mathcal{I}$ tel que $\forall i \in \mathcal{I}, \perp \triangleleft i$. Nous verrons que cela ajoute un axiome *structurel* à la logique.

9.3.1 Cas avec une infinité d'indices

Commençons par examiner le cas où \mathcal{I} est infini. Une propriété intéressante dans ce cas est que pour toute famille finie d'indices $\mathcal{J} \subseteq \mathcal{I}$, il est possible de trouver des hiérarchies

de représentations où ces indices ne sont pas comparables pour \triangleleft . Aussi, pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$, puisqu'elle comporte un nombre fini d'indices, il est possible de construire des hiérarchies de représentations qui, en se restreignant aux indices présents dans φ , se comportent comme un système de représentation.

Pour montrer cela, soit \mathcal{S} un système de représentations défini pour un ensemble fini d'indices \mathcal{J} . On peut le prolonger en une hiérarchie de représentations $\mathcal{H}(\mathcal{S}, \mathcal{I}, \perp)$ sur \mathcal{I} . Pour cela, nous avons besoin d'un indice qui sera plus petit que tous les autres. On identifie donc un indice $\perp \in \mathcal{I} \setminus \mathcal{J}$. Nous définissons \mathcal{R}_\perp par son ensemble de descriptions :

$$\{ \{ d_j \in \mathcal{R}_j \mid j \in \mathcal{J} \} \mid \forall j_1, j_2 \in \mathcal{J}, d_{j_1} \leq f_{j_1|j_2}(d_{j_2}) \}$$

Les autres éléments de \mathcal{R}_\perp s'en déduisent simplement. Par exemple, l'ordre sur \mathcal{R}_\perp est le produit point par point des ordres \leq_j , et $\top_\perp = \{ \top_j \mid j \in \mathcal{R}_j \}$. Pour les indices restant $i \in \mathcal{I} \setminus \mathcal{J}$, on pose :

$$\mathcal{R}_i = \langle \{ \top_i \}, \leq_{\top_i}, \top_i \rangle$$

La représentation la plus précise est le produit réduit des représentations de \mathcal{S} , les autres étant des représentations à un seul élément (l'ordre \leq_\top étant précisément défini comme l'ensemble $\langle \{ \top, \top \} \rangle$). Pour relier les différentes représentations avec \mathcal{R}_\perp , on peut vérifier que pour tout $i \neq \top$, la représentation \mathcal{R}_i est une approximation de \mathcal{R}_\perp . Aussi, on a $\forall i, \perp \triangleleft i$ (il s'agit bien d'une hiérarchie de représentations) et on peut définir une surjection de Galois $\langle \alpha_{i|\perp}, \gamma_{\perp|i} \rangle$ pour tout $i \in \mathcal{I} \setminus \{ \perp \}$.

On peut plus précisément s'intéresser aux $\langle \alpha_{j|\perp}, \gamma_{\perp|j} \rangle$ pour $j \in \mathcal{J}$. En effet, pour $i, j \in \mathcal{J}$ et $d_i \in \mathcal{R}_i$, l'ensemble $\{ f_{k|i}(d_i) \mid k \in \mathcal{J} \}$ est une description de \mathcal{R}_\perp , et c'est la plus grande contenant d_i . Ainsi, il s'agit de $\gamma_{\perp|i}(d_i)$. On en déduit que $\alpha_{j|\perp} \circ \gamma_{\perp|i}(d_i) = f_{j|i}(d_i)$ en prenant la $j^{\text{ème}}$ composante de l'ensemble précédent. On voit donc que la hiérarchie $\mathcal{H}(\mathcal{S}, \mathcal{I}, \perp)$ que l'on vient de définir de cette manière est bien un prolongement du système \mathcal{S} de départ.

Cette construction va nous être utile pour montrer la complétude des hiérarchies de représentations vis-à-vis de la logique **OL**, puisque pour une proposition φ donnée, si toutes les représentations la valident, alors on peut en déduire que tous les systèmes définis sur $\text{ind}(\varphi) \cup \{j\}$ (où $j \notin \text{ind}(\varphi)$) la valident aussi.

Théorème 8

Étant donné un ensemble \mathcal{I} d'indices, une proposition $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$ et un indice $j \in \mathcal{I} \setminus \text{ind}(\varphi)$, on a :

$$\models_{\mathcal{H}} \varphi \Rightarrow \vdash_{\mathbf{OL}, \text{ind}(\varphi) \cup \{j\}} \varphi$$

Preuve En effet, supposons que l'on a $\forall \mathcal{H}, \nu, \forall i, \llbracket \varphi \rrbracket_{\mathcal{H}, \nu, i} = \mathcal{R}_i$. Dans ce cas, soit un indice $j \notin \text{ind}(\varphi)$. Étant donné un système de représentations \mathcal{S} défini sur $\text{ind}(\varphi) \cup \{j\}$ et une interprétation atomique $\nu_{\mathcal{S}}$, en prolongeant \mathcal{S} en $\mathcal{H}(\mathcal{S}, \mathcal{I}, \perp)$ et $\nu_{\mathcal{S}}$ en une interprétation ν , on a par hypothèse que :

$$\forall i \in \text{ind}(\varphi) \cup \{j\}, \llbracket \varphi \rrbracket_{\mathcal{H}(\mathcal{S}, \mathcal{I}, \perp), \nu, i} = \mathcal{R}_i$$

Mais l'expression ensembliste de $\llbracket \varphi \rrbracket_{\mathcal{H}(\mathcal{S}, \mathcal{I}, \perp), \nu, i}$ ne comporte que des éléments issus du système de représentation \mathcal{S} de départ, ainsi que de $\nu_{\mathcal{S}}$, on en déduit donc que :

$$\forall i \in \text{ind}(\varphi) \cup \{j\}, \llbracket \varphi \rrbracket_{\mathcal{S}, \nu, i} = \mathcal{R}_i$$

En d'autres termes, on a $\mathcal{S}, \nu \models_{\mathcal{S}} \varphi$. Puisque c'est vérifié pour tous les systèmes de représentations, on en déduit que $\vdash_{\text{OL}, \text{ind}(\varphi) \cup \{i\}} \varphi$ \square

Nous pouvons maintenant utiliser ce résultat pour énoncer le théorème principal de caractérisation logique des hiérarchies de représentations pour un ensemble infini d'indices. Si l'on note $\models_{\mathcal{H}}$ pour exprimer la validité d'une formule pour toutes les hiérarchies de représentations pour \mathcal{I} fixé, on a :

Théorème 9

Soit \mathcal{I} un ensemble infini d'indices. Pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$, on a :

$$\models_{\mathcal{H}} \varphi \Leftrightarrow \vdash_{\text{OL}} \varphi$$

Preuve La correction, correspondant au sens \Leftarrow , vient du fait que les hiérarchies de représentations sont un type particulier de systèmes de représentations. Il faut maintenant montrer le sens de la complétude. Pour cela, étant donné un indice $j \in \mathcal{I} \setminus \text{ind}(\varphi)$, le théorème 8 nous assure que $\models_{\mathcal{H}} \varphi \Rightarrow \vdash_{\text{OL}, \text{ind}(\varphi) \cup \{j\}} \varphi$. L'application du théorème 7 permet alors de conclure. \square

Nous allons maintenant nous pencher sur le cas où l'ensemble d'indices \mathcal{I} est fini. Alors, le résultat précédent ne pourra plus s'appliquer directement, puisqu'il est possible d'exprimer au sein de la logique elle-même le fait qu'il existe une représentation plus précise que les autres.

9.3.2 Cas avec un ensemble fini d'indices

Pour étudier cette nouvelle situation, fixons un ensemble fini d'indices \mathcal{I} , et considérons une hiérarchie de représentations \mathcal{H} définie pour cet ensemble. Il est facile de prouver qu'il existe un indice $\perp \in \mathcal{I}$ tel que $\forall i, \perp \triangleleft i$. Dans ce cas, la hiérarchie de représentation valide toutes les propositions de la forme $K_i \varphi \rightarrow K_{\perp} K_i \varphi$:

$$\forall \varphi, \forall i \in \mathcal{I}, \mathcal{H} \models K_i \varphi \rightarrow K_{\perp} K_i \varphi$$

Nommons \mathbf{GD}_i la proposition $K_j \varphi \rightarrow K_i K_j \varphi$, où l'indice j peut varier, mais pas i . On voit donc que si une hiérarchie définie sur \mathcal{I} a comme représentation la plus précise i , alors elle valide l'axiome \mathbf{GD}_\perp . Puisqu'il n'est pas conséquence de la logique \mathbf{OL} , on en déduit que dans le cas où \mathcal{I} est fini, la logique modélisée par les hiérarchies de représentation contient au moins \mathbf{OL} et un axiome \mathbf{GD}_i pour un i donné.

Nous allons montrer que c'est suffisant. Pour cela, nous définissons le symbole $\models_{\mathcal{H},i}$ pour exprimer le fait qu'une propriété φ est validée par toutes les hiérarchies de représentation définies sur \mathcal{I} et de représentation minimale i . Avec cette notation, on peut énoncer le résultat de modélisation :

Théorème 10

Pour un ensemble fini \mathcal{I} d'indices, pour un indice $i \in \mathcal{I}$ et pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$, on a :

$$\models_{\mathcal{H},i} \varphi \Leftrightarrow \vdash_{\mathbf{OL}+\mathbf{GD}_i} \varphi$$

Preuve La correction découle du fait que les hiérarchies de représentations sont correctes vis-à-vis de \mathbf{OL} , et que l'on a $\models_{\mathcal{H},i} \mathbf{GD}_i$.

Pour prouver la complétude, nous allons adapter la preuve de complétude pour les systèmes de représentations (voir théorème 4 p.72). Par souci de lisibilité, nous considérerons que l'indice correspondant à la représentation la plus précise est \perp . On définit alors un système canonique \mathcal{C}_\perp , en spécifiant cette fois que les représentations \mathcal{R}_i sont constituées des ensembles premiers Γ_i pour la logique \mathbf{OL} à laquelle on a ajouté l'axiome $\mathbf{GD}_\perp : K_i \varphi \rightarrow K_\perp K_i \varphi$, et contenant les formules de la forme $\varphi \leftrightarrow K_i \varphi$. Le modèle canonique est un système de représentation, et on a :

$$\begin{aligned} \varphi \in \Gamma_i &\Rightarrow K_i \varphi \in \Gamma_i \\ &\Rightarrow K_\perp K_i \varphi \in \Gamma_i \\ &\Rightarrow K_i \varphi \in f_{\perp|i}(\Gamma_i) \\ &\Rightarrow \varphi \in f_{i|\perp} \circ f_{\perp|i}(\Gamma_i) \end{aligned}$$

Ainsi, $f_{i|\perp} \circ f_{\perp|i}$ est l'identité, ce qui fait que le modèle canonique que l'on vient de définir est une hiérarchie de représentations d'indice minimum \perp comme prévu.

En reprenant le raisonnement utilisé pour prouver le théorème 4, on montre que :

$$\models_{\mathcal{H},\perp} \varphi \Rightarrow \mathcal{C}_\perp \models_{\mathcal{H},\perp} \varphi \Rightarrow \vdash_{\mathbf{OL}+\mathbf{GD}_\perp} \varphi$$

On a donc le résultat espéré. □

Comme on l'a montré, l'ajout de ces conditions ne modifie pas la façon dont se comporte la connaissance (du moins dans le cas où il y a une infinité d'indices). Le fait que deux représentations données peuvent toujours être vues comme les approximations d'une troisième, il

est possible de manipuler ces structures facilement, puisque toute transformation impliquant plusieurs représentations (mais en nombre fini) peut être vu comme s'appliquant à une unique représentation. On pourra se référer à l'annexe D où sont présentées plusieurs constructions génériques pour la manipulation des hiérarchies de représentations. En particulier, nous montrons comme étendre les outils de base que sont les opérations de produit et de raffinage.

De plus, l'utilisation de hiérarchies de représentation va nous permettre de reprendre la discussion entreprise sur la façon dont est modélisée la connaissance dans le cadre d'observations partielles, et nous allons dans le chapitre suivant introduire une seconde manière de relier les propositions du langage $\mathcal{L}_{\mathcal{I},\Psi}$ et les ensembles de descriptions, ce qui va nous conduire à donner une nouvelle modélisation de la logique **IS4 + KV**.

Chapitre 10

Localisation et consistance

L'une des questions évoquées au début de ce travail était de savoir si une propriété prouvée pour un modèle donné devait être encore prouvable avec un modèle plus précis. Dans le formalisme que nous avons développé, il est possible d'avoir une traduction précise de ces deux notions, d'une part la comparaison entre la précision des représentations, incarnée par l'ordre partiel \triangleleft sur les indices, et d'autre part le fait qu'une propriété est prouvable pour une représentation donnée, avec les opérateurs de modalité K_i .

Nous pouvons donc exprimer de façon formelle et rigoureuse le postulat de base sous la forme de la proposition suivante :

$$\mathbf{Pos} : \forall \varphi, \forall i \triangleleft j, \vdash K_j \varphi \rightarrow K_i \varphi \quad (10.1)$$

Mais la logique OL nous indique que cette condition ne peut pas être appliquée en règle générale pour un système d'observations. En effet, on retrouve le même problème qu'avec l'axiome de confiance C puisque l'on en déduit pour tout φ et tous indices $i \triangleleft j \vdash K_i \varphi \rightarrow K_j \varphi$. En effet, on peut écrire :

$$\begin{aligned} \vdash K_j (\varphi \rightarrow K_j \varphi) \wedge \vdash K_j (\varphi \rightarrow K_j \varphi) &\rightarrow K_i (\varphi \rightarrow K_j \varphi) \\ \Rightarrow \vdash K_i (\varphi \rightarrow K_j \varphi) & \\ \Rightarrow \vdash K_i \varphi \rightarrow K_i K_j \varphi & \\ \Rightarrow \vdash K_i \varphi \rightarrow K_j \varphi & \end{aligned}$$

Ainsi, essayer d'ajouter le postulat de départ en modifiant la structure des systèmes ou des hiérarchies de représentations conduirait à rendre équivalentes toutes les représentations comparables pour l'ordre \triangleleft . Nous allons donc essayer de le reformuler pour qu'il soit possible de l'exprimer et l'utiliser dans le cadre des hiérarchies de représentations.

10.1 Localisation des propositions

Le postulat tel qu'il est exprimé dans l'équation 10.1 est trop fort. En effet, un point de vue peut être trop imprécis pour décider si une propriété est vraie ou non, il peut même ne fournir aucune information relative à cette proposition. Pour remédier à cela, nous allons définir, pour chaque proposition $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$ un ensemble $\text{Loc}(\varphi) \subseteq \mathcal{I}$ qui désigne l'ensemble des représentations correspondant à une observation suffisamment fine pour pouvoir être sûr de la véracité de la proposition φ . Cet ensemble correspondra à la "localisation" d'une propriété donnée.

Caractérisation de la fonction Loc

On peut identifier plusieurs propriétés que l'on utilisera pour définir précisément la fonction Loc . Tout d'abord, cette fonction va permettre de restreindre le champs d'application du postulat Pos en limitant l'ensemble des indices possibles pour chaque proposition. Ainsi, si pour une proposition φ donnée, on a un indice $j \in \text{Loc}(\varphi)$, cela signifie qu'il représente un point d'observation assez fin pour décider si φ est vraie. Mais alors, pour tout $i \triangleleft j$, la représentation \mathcal{R}_i est aussi suffisamment fine pour décider si φ est vraie ou non, ce qui implique que $j \in \text{Loc}(\varphi) \Rightarrow i \triangleleft j \Rightarrow i \in \text{Loc}(\varphi)$. On en déduit donc que les ensembles $\text{Loc}(\varphi)$ sont des idéaux de \mathcal{I} .

Maintenant, supposons que l'on a un indice i tel que $i \in \text{Loc}(\varphi) \cap \text{Loc}(\psi)$. Alors, cet indice correspond à une représentation où l'on peut décider de la validité des deux propositions φ et ψ . On peut alors aussi décider de la validité de $\varphi \vee \psi$, $\varphi \wedge \psi$ et $\varphi \rightarrow \psi$. Réciproquement, pour décider de la validité de chacune de ces trois propositions, il faut pouvoir décider de la validité des sous-termes. On pose donc :

$$\text{Loc}(\varphi \vee \psi) = \text{Loc}(\varphi \wedge \psi) = \text{Loc}(\varphi \rightarrow \psi) = \text{Loc}(\varphi) \cap \text{Loc}(\psi)$$

La proposition fautive \perp ne correspond à aucune information, et toute représentation est suffisamment fine pour décider de sa validité. Donc, on définit : $\text{Loc}(\perp) = \mathcal{I}$. Cela implique en particulier que $\text{Loc}(\neg\varphi) = \text{Loc}(\varphi)$. Pour décider de la validité d'une proposition de la forme $K_i \varphi$, il suffit de prendre une représentation contenant \mathcal{R}_i , puisque l'on se limite effectivement aux informations accessibles depuis cette représentation. Ainsi, $\text{Loc}(K_i \varphi) = \{i' \mid i' \triangleleft i\} = i \downarrow$. Pour les propositions atomiques, il faut fournir "à la main" les ensembles $\{\text{Loc}_\nu(\psi) \mid \psi \in \Psi\}$ que l'on suppose tous non vides. Avec les définitions ci-dessus, on peut définir la fonction Loc par induction sur la structure des termes sur l'ensemble de $\mathcal{L}_{\mathcal{I},\Psi}$. La définition de cette fonction est rappelée dans la figure 10.1.

L'usage de hiérarchies de représentations sera alors très utile dans le cadre de l'utilisation de la fonction Loc , puisque la propriété de fermeture de l'ensemble des indices (définie par

$$\begin{aligned}
\forall \psi \in \Psi, \text{Loc}(\psi) &= \text{Loc}_\nu(\psi) \\
\text{Loc}(\varphi \vee \psi) &= \text{Loc}(\varphi) \cap \text{Loc}(\psi) \\
\text{Loc}(\varphi \wedge \psi) &= \text{Loc}(\varphi) \cap \text{Loc}(\psi) \\
\text{Loc}(\varphi \rightarrow \psi) &= \text{Loc}(\varphi) \cap \text{Loc}(\psi) \\
\text{Loc}(K_i \varphi) &= i \downarrow \\
\text{Loc}(\perp) &= \mathcal{I}
\end{aligned}$$

Figure 10.1 – Définition inductive de Loc

l'équation 9.3 p.107) implique que l'on a :

$$\forall \varphi, \text{Loc}(\varphi) \neq \emptyset$$

Aussi, dans la suite, nous ne considèrerons que des hiérarchies de représentations afin de pleinement profiter de la notion de localisation des propriétés donnée par la fonction Loc.

Traduction du postulat

Grâce à cette fonction, nous allons pouvoir adapter le postulat de départ reliant les représentations comparables, en se restreignant à celles où une proposition donnée est localisée. On peut donc écrire une nouvelle définition du postulat sous la forme de la proposition :

$$\forall \varphi, \forall j \in \text{Loc}(\varphi), \forall i \triangleleft j, \vdash K_j \varphi \rightarrow K_i \varphi$$

Or, comme nous l'avons vu précédemment, cela implique en fait $\vdash K_i \varphi \rightarrow K_j \varphi$ en gardant les mêmes indices, puisque l'on peut reprendre la dérivation donnée page 117 en remarquant que $\text{Loc}(\varphi \rightarrow K_j \varphi) = \text{Loc}(\varphi) \cap j \downarrow = j \downarrow$. Ainsi, cette condition implique l'équivalence $\vdash K_j \varphi \leftrightarrow K_i \varphi$ pour $i \triangleleft j$ ce qui permet d'écrire, étant donné la fermeture de l'ensemble des indices, la nouvelle définition du postulat :

$$\mathbf{Pos}_2 : \forall \varphi, \forall i, j \in \text{Loc}(\varphi), \vdash K_i \varphi \leftrightarrow K_j \varphi \quad (10.2)$$

En terme d'interprétation vers des hiérarchies de représentations, ce principe peut être formulé de façon équivalente de la façon suivante :

$$\forall \varphi, \forall i, j \in \text{Loc}(\varphi), \forall d \in \mathcal{R}_i, d \in \llbracket \varphi \rrbracket_i \Leftrightarrow f_{j|i}(d) \in \llbracket \varphi \rrbracket_j$$

Dans un premier temps, nous allons supposer que cette équivalence est bien vérifiée, ce qui nous permettra de définir une nouvelle fonction d'interprétation $\{\llbracket \cdot \rrbracket\}$ tenant compte de la localisation des propositions. Ensuite, une fois cette interprétation définie, nous allons en déduire une condition nécessaire et suffisante pour que le postulat \mathbf{Pos}_2 soit vérifié.

Définition de l'interprétation $\{\cdot\}$

Tout d'abord, pour les propositions atomiques, il faut maintenant fournir plusieurs fonctions, une première : $\text{Loc}_\nu : \Psi \rightarrow \wp^\perp(\mathcal{I})$ donnant la localisation, et un ensemble de fonctions $\{\nu_i : \Psi \rightarrow \wp^\perp(\mathcal{R}_i)\}$ donnant l'interprétation proprement dite. Ces fonctions vérifient, en accord avec le postulat :

$$\forall \psi \in \Psi, \forall i, j \in \text{Loc}_\nu(\psi), \forall d \in \mathcal{R}_i, d \in \nu_i(\psi) \Leftrightarrow f_{j|i}(d) \in \nu_j(\psi) \quad (10.3)$$

Pour les connecteurs de la logique intuitionniste, pour les indices inclus dans la localisation des propositions, l'interprétation est la même que celle définie pour les systèmes de représentation. Par exemple :

$$\forall \varphi, \psi, \forall i \in \text{Loc}(\varphi \rightarrow \psi), \{\{\varphi \rightarrow \psi\}\}_i = \{d \mid \forall d' \leq d, d' \in \{\{\varphi\}\}_i \Rightarrow d' \in \{\{\psi\}\}_i\}$$

Il reste à étudier deux éléments pour arriver à la définition complète de $\{\cdot\}$: l'interprétation des formules de la forme $K_i \varphi$ et la signification de $\{\{\varphi\}\}_i$ pour $i \notin \text{Loc}(\varphi)$. Pour le premier cas, définissons tout d'abord $\{\{K_i \varphi\}\}_i = \{d \mid f_{j|i}(d) \in \{\{\varphi\}\}_j\}$ où j est un indice appartenant à $\text{Loc}(\varphi)$. Il est facile de voir que la définition de cet ensemble ne dépend pas de l'indice j choisi. En effet, si l'on a un autre indice $k \in \text{Loc}(\varphi)$, on a :

$$\begin{aligned} f_{j|i}(d) \in \{\{\varphi\}\}_j &\Rightarrow f_{k|j} \circ f_{j|i}(d) \in \{\{\varphi\}\}_k \\ &\Rightarrow f_{k|i}(d) \in \{\{\varphi\}\}_k \end{aligned}$$

On en déduit donc la définition pour tout $k \triangleleft i$ (avec toujours $j \in \text{Loc}(\varphi)$) :

$$\{\{K_i \varphi\}\}_k = \{d \mid f_{j|i} \circ f_{i|k}(d) \in \{\{\varphi\}\}_j\}$$

Pour les indices en dehors de la localisation d'une proposition, on se base sur le fait que si l'on a une proposition φ , alors :

$$\forall i, j \in \text{Loc}(\varphi), \{\{\varphi\}\}_j = \{\{K_i \varphi\}\}_j$$

Nous allons utiliser cette égalité, et la généraliser à l'ensemble des indices, en définissant $\{\{\varphi\}\}_i$ comme étant l'ensemble $\{d \mid f_{j|i}(d) \in \{\{\varphi\}\}_j\}$ pour $j \in \text{Loc}(\varphi)$. En d'autres termes, on a : $\{\{\varphi\}\}_i = \{\{K_j \varphi\}\}_i$ dans le cas où $i \notin \text{Loc}(\varphi)$, pour un indice quelconque $j \in \text{Loc}(\varphi)$.

On obtient donc la définition inductive de la fonction d'interprétation $\{\cdot\}$, rappelée dans la figure 10.2 (où $\text{Loc}(\varphi, \psi)$ désigne l'ensemble $\text{Loc}(\varphi) \cap \text{Loc}(\psi)$). Cette définition s'appuie sur le fait que le postulat exprimé par Pos_2 est vérifié. Mais ce n'est pas le cas en général. Nous allons donc dans la section suivante identifier une condition nécessaire et suffisante pour que ce soit le cas.

$$\begin{aligned}
\{\psi\}_i &= \nu_i(\psi) & \psi \in \Psi, i \in \text{Loc}_\nu(\psi) \\
\{\varphi \vee \psi\}_i &= \{\varphi\}_i \cup \{\psi\}_i & i \in \text{Loc}(\varphi, \psi) \\
\{\varphi \wedge \psi\}_i &= \{\varphi\}_i \cap \{\psi\}_i & i \in \text{Loc}(\varphi, \psi) \\
\{\varphi \rightarrow \psi\}_i &= \{d \mid \forall d' \leq d, d' \in \{\varphi\}_i \Rightarrow d' \in \{\psi\}_i\} & i \in \text{Loc}(\varphi, \psi) \\
\{K_j \varphi\}_i &= \{d \mid f_{k|j} \circ f_{j|i}(d) \in \{\varphi\}_k\} & k \in \text{Loc}(\varphi), i \in \text{Loc}(K_j \varphi) \\
\{\varphi\}_i &= \{d \mid f_{j|i}(d) \in \{\varphi\}_j\} & i \notin \text{Loc}(\varphi), j \in \text{Loc}(\varphi) \\
\{\perp\}_i &= \emptyset
\end{aligned}$$

Figure 10.2 – Définition de la fonction d'interprétation $\{\cdot\}$

10.2 Hiérarchies consistantes

Pour cela, nous allons définir et étudier une hiérarchie particulière, pour isoler une condition sur la structure de celle-ci, et nous montrons qu'elle est suffisante pour que la condition Pos_2 soit vérifiée.

Définition de la condition

Soit deux indices $i \triangleleft j$ et une description $d_0 \in \mathcal{R}_i$. On suppose que l'on a aussi deux propositions primitives ψ et ϑ . Nous allons définir leur interprétation de la façon suivante :

$$\begin{aligned}
\text{Loc}(\psi) &= \text{Loc}(\vartheta) = j \downarrow \\
\{\psi\}_j &= \{d \mid d \leq f_{j|i}(d_0)\} & \{\vartheta\}_j &= \{f_{j|i}(d) \mid d \leq d_0\}
\end{aligned}$$

On remarque que cette définition est suffisante, car les ensembles $\{\psi\}_k$ et $\{\vartheta\}_k$ s'en déduisent pour $k \triangleleft j$ en utilisant le postulat Pos_2 . Avec ces propositions atomiques, regardons la situation pour la proposition $\psi \rightarrow \vartheta$. On montre tout d'abord que $d_0 \in \{\psi \rightarrow \vartheta\}_i$. En effet, on a :

$$\forall d \leq d_0, d \in \{\vartheta\}_i = \{d' \mid f_{j|i}(d') \leq f_{j|i}(d_0)\}$$

Cela implique, puisque $d_0 \in \{\psi \rightarrow \vartheta\}_i \Leftrightarrow f_{j|i}(d_0) \in \{\psi \rightarrow \vartheta\}_j$, la proposition suivante :

$$\forall d \leq f_{j|i}(d_0), d \in \{\psi\}_j \Rightarrow d \in \{\vartheta\}_j$$

Or, il découle de $d \leq f_{j|i}(d_0)$ que $d \in \{\psi\}_j$. On en déduit donc :

$$\forall d \leq f_{j|i}(d_0), d \in \{\vartheta\}_j = \{f_{j|i}(d') \mid d' \leq d_0\}$$

Cette condition peut être exprimée autrement en disant que $\{f_{j|i}(d') \mid d' \leq d_0\}$ est un idéal de \mathcal{R}_j . On obtient donc de façon générale la condition que nous allons retenir et dont nous allons montrer qu'elle est suffisante pour que Pos_2 soit vérifié avec l'interprétation $\{[\cdot]\}$:

$$\forall i \triangleleft j, \forall d \in \mathcal{R}_i, \{f_{j|i}(d') \mid d' \leq d\} \in \wp^\downarrow(\mathcal{R}_j) \quad (10.4)$$

Définitions équivalentes

Cette condition stipule que si l'on a deux indices $i \triangleleft j$, l'image d'un idéal de \mathcal{R}_i par $f_{j|i}$ est un idéal de \mathcal{R}_j . La proposition suivante en donne encore une autre formulation :

Proposition 20

La condition 10.4 est équivalente à :

$$\forall i \triangleleft j, \forall d \in \mathcal{R}_i, \{f_{j|i}(d') \mid d' \leq d\} = \{d'' \mid d'' \leq f_{j|i}(d)\} \quad (10.5)$$

Preuve Pour montrer que 10.4 implique 10.5, on remarque tout d'abord que par croissance des $f_{j|i}$, on a l'inclusion $\{f_{j|i}(d') \mid d' \leq d\} \subseteq \{d'' \mid d'' \leq f_{j|i}(d)\}$. Réciproquement, si $d'' \leq f_{j|i}(d)$, alors on déduit de la condition 10.4 qu'il existe une description $d' \leq d$ telle que $d' = f_{j|i}(d'')$.

Dans l'autre sens, si 10.5 est vérifiée, alors il est clair qu'un ensemble de la forme :

$$\{f_{j|i}(d') \mid d' \leq d\}$$

est un idéal de \mathcal{R}_j . □

Écrit autrement en utilisant la notation $d \downarrow$, cette condition peut se noter aussi $f_{i|j}(d \downarrow) = (f_{i|j}(d)) \downarrow$.

Une autre façon d'exprimer cette condition est la suivante : si l'on a deux indices $i \triangleleft j$ et deux descriptions $d_i \in \mathcal{R}_i$ et $d_j \in \mathcal{R}_j$ telles que $d_j \leq f_{j|i}(d_i)$, alors il existe une autre description $d'_i \in \mathcal{R}_i$ telle que $d'_i \leq d_i$ et $f_{j|i}(d'_i) = d_j$. Mais alors, on peut caractériser un tel élément d'_i : puisqu'il est plus petit que d_i et que $f_{i|j}(d_j)$, la description $d_i \wedge f_{i|j}(d_j)$ fait forcément l'affaire. En reprenant la notation \frown vue page 41 ($d_1 \frown d_2$ signifie que $\langle d_1, d_2 \rangle$ appartient au domaine de définition de \wedge), on peut écrire :

Proposition 21

La condition 10.4 est équivalente à :

$$\forall i \triangleleft j, \forall d_i \in \mathcal{R}_i, \forall d_j \in \mathcal{R}_j, d_j \leq f_{j|i}(d_i) \Rightarrow d_i \frown f_{i|j}(d_j) \quad (10.6)$$

Preuve On a en effet :

$$\begin{aligned}
d_j \in \{f_{j|i}(d) \mid d \leq d_i\} &\Leftrightarrow \exists d'_i : d'_i \leq d_i \text{ et } f_{j|i}(d'_i) = d_j \\
&\Leftrightarrow \exists d'_i : d'_i \leq d_i \text{ et } d'_i \leq f_{i|j}(d_j) \text{ et } f_{j|i}(d'_i) = d_j \\
&\Leftrightarrow \exists d'_i : d'_i \leq d_i \wedge f_{i|j}(d_j) \text{ et } f_{j|i}(d'_i) = d_j \\
&\Leftrightarrow d_i \frown f_{i|j}(d_j)
\end{aligned}$$

□

Comme on le voit dans ces différentes formulations, la condition que l'on a exhibée sur les hiérarchies de représentations exprime le fait qu'une représentation plus approximative qu'une autre doit cependant "en recouvrir toute la largeur".

Preuve d'équivalence

Nous allons supposer maintenant que cette condition est vérifiée, et nous allons montrer qu'alors, en utilisant l'interprétation $\{\cdot\}$, la proposition \mathbf{Pos}_2 est vérifiée. Ainsi, on aura prouvé que la proposition 10.4 est une condition nécessaire et suffisante pour que \mathbf{Pos}_2 soit vérifiée elle aussi.

Proposition 22

Étant donné une hiérarchie de représentation \mathcal{H} vérifiant la condition 10.4, une fonction $\text{Loc}_\nu : \Psi \rightarrow \wp^\downarrow(\mathcal{I})$ et un ensemble de fonctions $\nu_i : \Psi \rightarrow \wp^\downarrow(\mathcal{R}_i)$ vérifiant la condition 10.3, la condition \mathbf{Pos}_2 :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I}, \Psi}, \forall i, j \in \text{Loc}(\varphi), \forall d \in \mathcal{R}_i, d \in \{\{\varphi\}\}_i \Leftrightarrow f_{j|i}(d) \in \{\{\varphi\}\}_j$$

est elle aussi vérifiée.

Preuve On procède par récurrence pour montrer ce résultat. Le cas de base, c'est-à-dire le cas des propositions atomiques, est réglé par le fait que Loc_ν et les ν_i vérifient la proposition 10.3. Supposons maintenant par récurrence que deux propositions φ et ψ vérifient la condition voulue.

Soient deux indices $i, j \in \text{Loc}(\varphi) \cap \text{Loc}(\psi)$. On a :

$$\begin{aligned}
d \in \{\{\varphi \vee \psi\}\}_i &\Leftrightarrow d \in \{\{\varphi\}\}_i \cup \{\{\psi\}\}_i \\
&\Leftrightarrow f_{j|i}(d) \in \{\{\varphi\}\}_j \cup \{\{\psi\}\}_j \\
&\Leftrightarrow f_{j|i}(d) \in \{\{\varphi \vee \psi\}\}_j
\end{aligned}$$

De la même façon, on montre que $d \in \{\{\varphi \wedge \psi\}_i\} \Leftrightarrow f_{j|i}(d) \in \{\{\varphi \wedge \psi\}_j\}$. Le cas de l'implication est un peu plus compliqué. Tout d'abord, en utilisant les définitions et l'hypothèse de récurrence, on peut écrire :

$$\begin{aligned} d \in \{\{\varphi \rightarrow \psi\}_i\} &\Leftrightarrow \forall d' \leq d, f_{j|i}(d') \in \{\{\varphi\}_j\} \Rightarrow f_{j|i}(d') \in \{\{\psi\}_j\} \\ f_{j|i}(d) \in \{\{\varphi \rightarrow \psi\}_j\} &\Leftrightarrow \forall d' \leq f_{j|i}(d), d' \in \{\{\varphi\}_j\} \Rightarrow d' \in \{\{\psi\}_j\} \end{aligned}$$

On peut alors conclure pour le cas de l'implication, puisque la proposition 20 nous assure que les ensembles $\{f_{j|i}(d') \mid d' \leq d\}$ et $\{d' \mid d' \leq f_{j|i}(d)\}$ sont égaux.

Le dernier cas, concernant $K_i \varphi$, est direct puisque par définition, on a :

$$\forall j \triangleleft i, d \in \{\{K_i \varphi\}_j\} \Leftrightarrow f_{i|j}(d) \in \{\{K_i \varphi\}_i\}$$

On en déduit donc, pour $k \triangleleft i$:

$$\begin{aligned} d \in \{\{K_i \varphi\}_j\} &\Leftrightarrow f_{i|j}(d) \in \{\{K_i \varphi\}_i\} \\ &\Leftrightarrow f_{i|k} \circ f_{k|i} \circ f_{i|j}(d) \in \{\{K_i \varphi\}_i\} \\ &\Leftrightarrow f_{i|k} \circ f_{k|j}(d) \in \{\{K_i \varphi\}_i\} \\ &\Leftrightarrow f_{k|j}(d) \in \{\{K_i \varphi\}_k\} \end{aligned}$$

On a utilisé pour cela les inégalités :

$$f_{i|k} \circ f_{k|i} \circ f_{i|j} = f_{i|j} \leq f_{i|k} \circ f_{k|j} \leq f_{i|k} \circ f_{k|i} \circ f_{i|j}$$

L'hypothèse de récurrence est donc bien vérifiée, ce qui permet de conclure. \square

La condition que nous avons exhibée nous permet donc, en utilisant l'interprétation $\{\{\cdot\}\}$, d'avoir le comportement entre des représentations comparables exprimé par la condition Pos_2 .

Définition 3

Une hiérarchie consistante de représentations $\mathcal{H} = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$ est une hiérarchie de représentation vérifiant la condition 10.4 que l'on rappelle :

$$\forall i \triangleleft j, \forall d \in \mathcal{R}_i, f_{j|i}(d) \downarrow = f_{j|i}(d \downarrow)$$

Les hiérarchies consistantes, avec l'interprétation $\{\{\cdot\}\}$, modélisent une logique différente de OL. C'est ce que nous allons étudier maintenant.

10.3 Logique des hiérarchies consistantes

Retour sur la consistance globale

L'utilisation conjointe de hiérarchies consistantes et de l'interprétation $\{\cdot\}$ permet d'obtenir une logique différente **OL**. Par exemple, l'axiome de consistance globale **GD** vu précédemment (voir 8.4 p.98) est maintenant vérifié. On rappelle son intitulé :

$$\mathbf{GD} : \forall \varphi, \forall i, j, \vdash K_i \varphi \rightarrow \neg K_j \neg \varphi$$

Pour montrer cela, on observe tout d'abord que $\text{Loc}(K_i \varphi \rightarrow \neg K_j \neg \varphi) = i \downarrow \cap j \downarrow$. Soit k un indice appartenant à cet ensemble, ainsi que $l \in \text{Loc}(\varphi)$. On a par définition :

$$\begin{aligned} \{[K_i \varphi \rightarrow \neg K_j \neg \varphi]\}_k &= \{\neg (K_i \varphi \wedge K_j \neg \varphi)\}_k \\ &= \{d \mid \forall d' \leq d, d' \notin \{[K_i \varphi]\}_k \cap \{[K_j \neg \varphi]\}_k\} \end{aligned}$$

Mais l'ensemble $\{[K_i \varphi]\}_k \cap \{[K_j \neg \varphi]\}_k$ est vide, comme le montre la suite d'implications suivante :

$$\begin{aligned} d' \in \{[K_i \varphi]\}_k &\Rightarrow f_{l|i} \circ f_{i|k}(d') \in \{[\varphi]\}_l \\ &\Rightarrow f_{l|k}(d') \in \{[\varphi]\}_l \\ &\Rightarrow \exists d'' \leq f_{l|j} \circ f_{j|k}(d') : d'' \in \{[\varphi]\}_l \\ &\Rightarrow f_{l|j} \circ f_{j|k}(d') \notin \{[\neg \varphi]\}_l \\ &\Rightarrow d' \notin \{[K_j \neg \varphi]\}_k \end{aligned}$$

Cela nous assure donc que l'on a $\{[K_i \varphi \rightarrow \neg K_j \neg \varphi]\}_k = \mathcal{R}_k$, et que de façon plus générale, l'axiome de consistance globale **GD** est vérifié par les hiérarchies consistantes pour l'interprétation $\{\cdot\}$.

Étude de la logique

Nous allons maintenant caractériser précisément la nouvelle logique que l'on obtient en utilisant les hiérarchies consistantes et l'interprétation $\{\cdot\}$. Tout d'abord, nous venons de voir que l'axiome de consistance globale **GD** est vérifié. On peut en fait montrer qu'un autre axiome plus puissant, en l'occurrence **T** : $K_i \varphi \rightarrow \varphi$, est lui aussi vérifié. En effet, on a pour un indice $k \in \text{Loc}(\varphi) \cap i \downarrow$:

$$\begin{aligned} \{[K_i \varphi \rightarrow \varphi]\}_k &= \{d \mid \forall d' \leq d, d' \in \{[K_i \varphi]\}_k \Rightarrow d' \in \{[\varphi]\}_k\} \\ &= \{d \mid \forall d' \leq d, f_{k|i} \circ f_{i|k}(d') \in \{[\varphi]\}_k \Rightarrow d' \in \{[\varphi]\}_k\} \\ &= \mathcal{R}_k \end{aligned}$$

puisque l'on a $k \triangleleft i \Rightarrow \forall d, d \leq f_{k|i} \circ f_{i|k}(d)$.

Il est cependant intéressant de remarquer que la logique que l'on obtient avec cette nouvelle interprétation ne contient pas la précédente, c'est-à-dire **OL**. En effet, l'axiome $L_4 : K_i(\varphi \rightarrow K_i \varphi)$ n'est plus vérifié, puisqu'en conjonction avec l'axiome **T**, il permet de déduire $\varphi \rightarrow K_i \varphi$, et donc en utilisant à nouveau **T**, $\varphi \leftrightarrow K_i \varphi$, lequel n'est pas vérifié puisqu'en général, $f_{k|i} \circ f_{i|k}(d) \neq d$.

L'axiome **4**, qui est plus faible que L_4 , est lui vérifié, grâce au fait que :

$$f_{k|i} \circ f_{i|k} \circ f_{k|i} \circ f_{i|k} = f_{k|i} \circ f_{i|k}$$

Pour résumer, la nouvelle logique que l'on considère vérifie les axiomes de la logique observationnelle **OL** à laquelle on a ôté L_4 et l'on a ajouté les axiomes **T** et **4**. La logique que l'on obtient de cette manière est exactement **IS4 + KV**, que l'on avait rencontré précédemment dans le cas de l'approximation d'une représentation par une autre (voir le théorème 2 p.62). Nous allons en fait montrer que **IS4 + KV** est exactement la logique modélisée par les hiérarchies consistantes et l'interprétation $\{\{\cdot\}\}$, dans le cas où l'ensemble d'indices \mathcal{I} est infini. Comme précédemment, dans le cas où \mathcal{I} est fini, il faut spécifier la représentation qui est la plus précise, et ajouter un axiome en conséquence. Pour exprimer le résultat du cas infini de façon rigoureuse, définissons la notation $\models_{\mathcal{CH}}$ pour représenter la validation pour ces structures (en utilisant les notations usuelles, et \mathcal{H} désignant une hiérarchie consistante) :

$$\forall \varphi, \models_{\mathcal{CH}} \varphi \Leftrightarrow \forall \mathcal{H}, \text{Loc}_\nu, \nu, \forall i \in \text{Loc}(\varphi), \{\{\varphi\}\}_i = \mathcal{R}_i$$

On peut alors écrire :

Théorème 11

*Pour tout ensemble infini d'indices \mathcal{I} , toute proposition $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$ est prouvable pour la logique **IS4 + KV** si et seulement si elle est valide pour les hiérarchies consistantes :*

$$\forall \varphi \in \mathcal{L}_{\mathcal{I},\Psi}, \vdash_{\mathbf{IS4+KV}} \varphi \Leftrightarrow \models_{\mathcal{CH}} \varphi$$

Preuve La preuve complète se trouve en annexe B.5. □

Si l'ensemble \mathcal{I} est fini, et que \perp est associé à la représentation la plus précise, on voit qu'il faut ajouter un axiome de la forme $\mathbf{GD}'_{\perp} : \varphi \rightarrow K_{\perp} \varphi$ pour obtenir la logique modélisée par ces hiérarchies consistantes. On montre en particulier dans l'annexe B.5 que si l'on désigne par $\models_{\mathcal{CH},\perp} \varphi$ la validité pour toutes les hiérarchies consistantes telle que la représentation \mathcal{R}_{\perp} est plus précise que toutes les autres, on a :

$$\forall \varphi, \models_{\mathcal{CH},\perp} \varphi \Leftrightarrow \vdash_{\mathbf{IS4+KV+GD}'_{\perp}} \varphi$$

L'utilisation de hiérarchies consistantes avec l'interprétation $\{\{\cdot\}\}$ permet donc de modéliser la logique **IS4 + KV**, qui a l'avantage par rapport à **OL** de permettre d'exprimer des relations plus fortes entre les différentes façons d'observer un système, assurant par exemple que des observations différentes ne seront jamais en contradiction l'une avec l'autre.

Nous allons cependant montrer dans la section suivante que ces structures sont moins intuitives ou pratiques à utiliser que les hiérarchies de représentations.

10.4 Utilisation de hiérarchies consistantes

Les exemples et sujets que nous allons développer dans cette section vont rester à un niveau moins formel que le reste de ce travail. En effet, il s'agit ici de la présentation d'arguments illustrant ce que nous estimons être des faiblesses des hiérarchies consistantes et de l'interprétation $\{\{\cdot\}\}$. Cependant, ces arguments sont subjectifs et ne sont que le reflet de notre opinion vis-à-vis des avantages et inconvénients des structures considérées.

10.4.1 La localisation des propositions

On peut tout d'abord s'interroger sur le sens de la fonction de localisation Loc . Cette fonction a été introduite pour traduire le fait qu'il fallait disposer d'un moyen pour identifier les observations ayant une "résolution" suffisante pour être sûr de la validité d'une proposition donnée. Ainsi, pour une proposition φ , l'ensemble $\text{Loc}(\varphi)$ désigne l'ensemble des représentations suffisamment précises pour pouvoir décider de la validité de φ .

Si la notion de résolution minimale peut être considérée comme légitime, la façon dont elle est formalisée dans le cadre de l'interprétation $\{\{\cdot\}\}$ fait qu'elle n'est pas présente dans la logique elle-même. Ainsi, pour chaque proposition φ et une interprétation donnée, il existe au moins un indice i tel que $i \in \text{Loc}(\varphi)$. On a alors $\forall j, \{\{\varphi\}\}_j = \{\{K_i \varphi\}\}_j$. Cette égalité dans les interprétations se traduit dans la logique par la proposition valide :

$$\models \varphi \leftrightarrow K_i \varphi$$

Il existe une proposition de ce type pour chaque $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$, et on peut écrire de façon générale :

$$\forall \varphi \in \mathcal{L}_{\mathcal{I},\Psi}, \exists i \in \mathcal{I} : \models \varphi \leftrightarrow K_i \varphi$$

Mais cette propriété, qui représente une caractéristique importante de la logique modélisée par une interprétation donnée, n'est pas exprimable dans cette logique. Elle reste extérieure à celle-ci : étant donné une formule φ , il n'existe aucun moyen à l'intérieur de la logique de déterminer un indice i tel que φ et $K_i \varphi$ soient équivalents. Pour que cela soit possible, il faudrait au moins

étendre le langage de façon à introduire une quantification sur les indices. Mais même alors, cela ne suffirait pas pour trouver, pour une proposition φ , un indice i correspondant.

Ainsi, certains aspects importants du raisonnement que l'on peut faire sur un système modélisé par une hiérarchie consistante obligent à se référer au modèle. Toute l'information que celui-ci donne sur le système étudié n'est pas accessible de façon interne.

10.4.2 La manipulation des hiérarchies consistantes

Nous allons maintenant donner quelques exemples montrant que la condition sur l'on impose sur les hiérarchies de représentations pour qu'elles soient consistantes entraîne une grande perte dans la souplesse de manipulation de ces structures. On reprends pour cela quelques-unes des constructions présentées dans l'annexe D.

Les partitions de représentations

Nous allons tout d'abord nous intéresser à la construction de partition décrite dans la partie D.4.3. Il est facile de montrer que cette construction ne peut être appliquée telle qu'elle à une hiérarchie consistante dans le cas général. Un contre-exemple peut être défini simplement. Considérons en effet une hiérarchie \mathcal{R} constituée d'une unique représentation composée de quatre éléments, une description vide \top et trois descriptions d_1, d_2, d_3 incomparables entre elles. Cette hiérarchie est consistante, comme le sont toutes les hiérarchies à un seul indice. Un exemple de partition à partir de \mathcal{R} est représenté dans la figure 10.3. La hiérarchie obtenue n'est pas consistante, puisque le couple $\langle d_2, d_1 \rangle$ n'appartient pas au domaine de définition de la conjonction \wedge comme l'imposerait l'équation 10.6 p.122 sachant que $d_1 = f_{\top|B}(t)$.

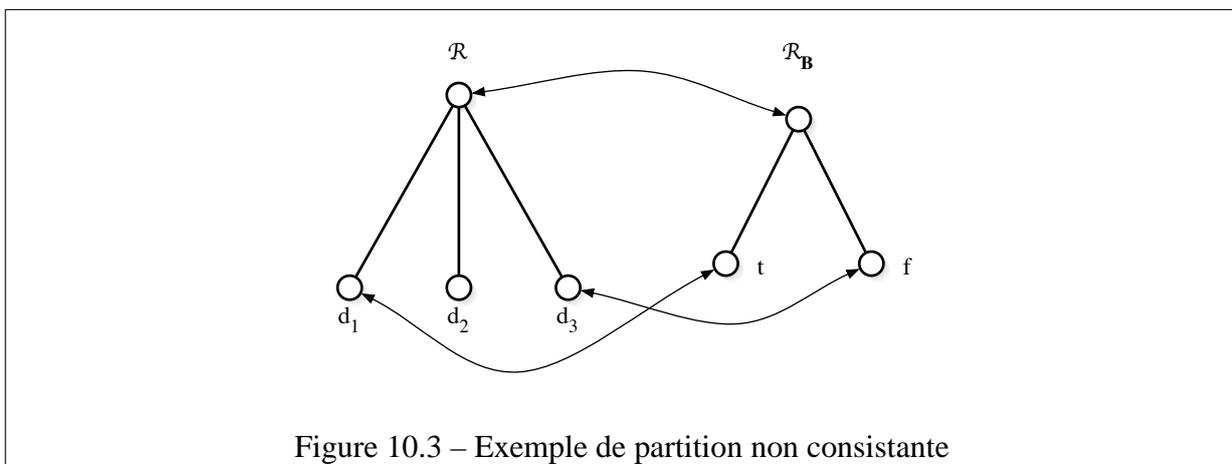


Figure 10.3 – Exemple de partition non consistante

On voit donc que la possibilité de partitionner une hiérarchie consistante dans le cas général implique, si l'on veut obtenir une hiérarchie consistante comme résultat, d'ajouter des

contraintes sur la structure des représentations. Il en résulterait donc un alourdissement supplémentaire de ces structures.

Le raffinage de hiérarchies

Comme on vient de le voir, l'opération de partitionnement est problématique dans le cadre des hiérarchies consistantes. Cette opération est un cas particulier de l'opération de raffinage. C'est à cette dernière que nous allons maintenant nous intéresser.

Nous allons étudier un aspect de l'utilisation d'un tel opérateur pour effectuer un raffinage. On suppose pour cela que l'on a deux représentations \mathcal{R}_i et \mathcal{R}_j avec $i \triangleleft j$, et deux fonctions de transformation α, γ de telle sorte que l'ensemble forme une hiérarchie consistante. Nous allons appliquer à la représentation \mathcal{R}_i un opérateur ϕ_i . En utilisant les définitions du raffinage, les représentations obtenues après résultats sont \mathcal{R}'_i et \mathcal{R}'_j où l'on a :

$$\begin{aligned} \forall d_i \in \mathcal{R}'_i, d_i &= \phi_i(d_i) \\ \forall d_j \in \mathcal{R}'_j, d_j &= \alpha \circ \phi_i \circ \gamma(d_j) \end{aligned}$$

Pour que la hiérarchie obtenue soit consistante, soient deux descriptions d_i et d_j telles que $d_j \leq \alpha'(d_i)$ (on rappelle que $\alpha' = \alpha$). La proposition 10.6 nous impose alors que la description $d_i \wedge' \gamma'(d_j) = \phi_i(d_i \wedge \phi_i \circ \gamma(d_j))$ est bien définie. Mais par consistance de la hiérarchie de départ, la seule information que l'on a est l'existence de la description $d_i \wedge \gamma(d_j)$.

Avec cette définition, on s'aperçoit que pour que la condition soit vérifiée dans le cas général (c'est-à-dire que la description $\phi_i(d_i \wedge \phi_i \circ \gamma(d_j))$ existe dans les conditions que l'on vient d'énoncer), il faut ajouter des conditions supplémentaires. En effet, il est facile de construire un exemple où elle n'est pas vérifiée. Considérons pour cela les hiérarchies représentées dans la figure 10.4. On part d'une hiérarchie consistante \mathcal{H} à laquelle on applique un opérateur ϕ qui a pour action de transformer toute description $d \leq x$ en $d \wedge y$.

Il est facile de s'apercevoir que la hiérarchie obtenue n'est plus consistante, puisque par exemple dans la hiérarchie de départ, $d_i \wedge \gamma(d_j)$ est bien définie, mais $d_i \wedge' \gamma'(d_j)$ ne l'est plus à l'arrivée.

On voit donc que l'opération de raffinage, qui était présentée dans le chapitre 4.2 et l'annexe D comme l'une des opérations les plus importantes avec le produit cartésien pour la manipulation de hiérarchies, nécessiterait d'être modifiée pour pouvoir être appliquée aux hiérarchies consistantes. En particulier, il faut appliquer des restrictions aux opérateurs de clôture que l'on souhaiterait utiliser dans ce cadre. Cependant, nous n'avons pu caractériser aucune restriction

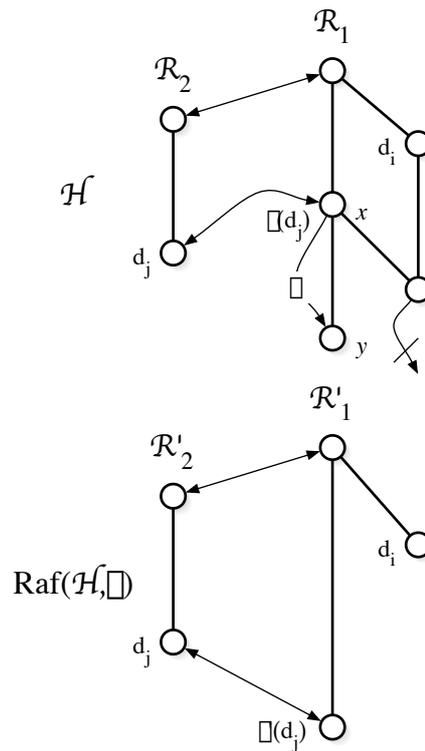


Figure 10.4 – Exemple de raffinage non consistant

“raisonnable”, ce qui fait que nous n’avons pas trouvé de classe d’opérateurs suffisamment intuitive et facile à définir, comme l’est la classe des opérateurs de clôture inférieure pour les hiérarchies de représentations.

De façon plus générale, nous voyons la nécessité de modifier les opérations de base comme un inconvénient à l’utilisation des hiérarchies consistantes. Alors que pour les hiérarchies de représentations, nous avons présenté un ensemble d’opérations générales, faciles à définir et permettant une grande souplesse dans la manipulation de ces structures, nous voyons que si l’on se restreint aux hiérarchies consistantes, on perd une grande partie de ces avantages.

Cependant, l’étude de cette restriction des structures de base a permis d’isoler un ensemble d’idées et de concepts importants relativement au but de départ qui est la modélisation et l’étude logique de la connaissance basée sur l’observation de systèmes. Dans la prochaine partie, nous allons mettre en parallèle ces deux formalismes, et présenter une méthode pour combiner les deux approches de façon à obtenir dans une certaine mesure le meilleur de celles-ci.

Quatrième partie

Comparaison

Jusqu'à présent, nous avons développé deux approches différentes de la modélisation de la connaissance dans le cadre d'observations partielles. Tout d'abord, nous avons défini les systèmes de représentation, ainsi que la logique **OL** qui lui est associée, cette logique pouvant être vue comme assez faible, puisque ne permettant pas d'exprimer de liens forts entre les différents points d'observation.

Ensuite, nous avons introduit les hiérarchies de représentations, qui sont une extension de la structure précédente, ainsi qu'une nouvelle manière d'interpréter les propositions de $L_{\mathcal{I},\Psi}$. La logique correspondante, que nous avons identifiée comme étant **IS4 + KV**, fournit plus de liens entre les différentes observations, mais le prix à payer est comme on vient de le voir une perte importante de la généralité et de la souplesse des structures manipulées.

Nous allons maintenant comparer plus en profondeur ces deux approches. En particulier, nous reviendrons sur la notion de localisation, qui permet de relier la prouvabilité d'une propriété à un ensemble de représentations. En modifiant cette notion pour mettre plus précisément en correspondance une proposition et une représentation où elle a été observée, nous pourrions identifier une catégorie de propositions qui sont utilisables de façon plus sûre que les autres, en ce sens que l'on peut les utiliser en faisant abstraction de leur représentation d'origine. Ainsi, l'utilisation de la logique **OL** et de ses limitations par rapport à **IS4 + KV** nous permettra d'obtenir une ébauche de théorie des informations communicables, en fournissant à la fois un critère de sélection des propositions et une méthode de transformation des propositions en général pour qu'elles remplissent ce critère.

Chapitre 11

Synthèse

Nous pouvons maintenant résumer les deux approches que l'on a décrites dans ces pages. On a d'une part la structure des hiérarchies de représentations, avec l'interprétation $\llbracket \cdot \rrbracket$, et d'autre part les hiérarchies consistantes avec $\{\cdot\}$. Le premier formalisme mène, comme on l'a vu, à une modélisation du comportement logique de l'information suivant la logique observationnelle **OL**. Elle est constituée de la logique intuitionniste à laquelle on a ajouté la règle de nécessité, ainsi que les axiomes et règles résumés dans la table 9.

$K_i(\varphi \rightarrow \psi) \rightarrow K_i\varphi \rightarrow K_i\psi$	K
$K_i\varphi \rightarrow \neg K_i\neg\varphi$	D
$K_i(\varphi \leftrightarrow K_i\varphi)$	L
$K_i K_j\varphi \rightarrow K_j\varphi$	T₂
$K_i(\varphi \vee \psi) \rightarrow K_i\varphi \vee K_i\psi$	KV
$\frac{\forall i, \vdash K_i\varphi}{\vdash \varphi}$	Univ
$\frac{\vdash \varphi}{\vdash K_i\varphi}$	Nec

Table 9 – Axiomes et règles de la logique **OL**

Ce formalisme peut être caractérisé par deux traits principaux : les hiérarchies de représentations sont une structure algébrique très générale et facile à manipuler, comme l'annexe D nous le montre. En contrepartie, on a vu au début de cette partie que la logique **OL** avait le défaut de fournir très peu de liens entre les informations correspondant aux différentes représentations.

Cette dernière constatation a amené dans le chapitre précédent à définir une nouvelle struc-

ture – les hiérarchies consistantes – et une nouvelle interprétation de cette structure sous forme logique. On obtient alors la logique **IS4 + KV** comme modélisation du comportement de l'information dans le cas de représentations partielles. L'un des principaux avantages de cette logique est la vérification de l'axiome **T** : $K_i \varphi \rightarrow \varphi$. Cet axiome permet en effet de prouver les axiomes de confiance entre agents et de consistance globale, dont on rappelle les énoncés :

$$\begin{array}{ll} K_i K_j \varphi \rightarrow K_i \varphi & \mathbf{C} \\ K_i \varphi \rightarrow \neg K_j \neg \varphi & \mathbf{GD} \end{array}$$

Malheureusement, les hiérarchies consistantes sont difficiles à manipuler, comme nous avons essayé de le montrer dans le chapitre précédent.

Nous allons maintenant présenter une adaptation de l'approche utilisant les hiérarchies de représentation permettant de regagner une partie de l'expressivité de la logique **IS4 + KV**, tout en conservant une grande souplesse d'utilisation. Pour cela, nous allons examiner à nouveau la notion de localisation des propositions, ce qui impliquera une légère transformation du langage logique utilisé, et plus précisément l'ajout explicite d'indices pour indiquer la localisation.

11.1 Nouvelle approche de la localisation

Nous allons tout d'abord revenir à la notion de localisation des propositions. Elle avait été définie en suivant l'idée qu'on ne pouvait décider de la validité d'une proposition que dans des représentations ayant une précision suffisante pour cela. Le traitement que l'on avait fait de cette notion avait conduit à la définition des hiérarchies consistantes et de la fonction d'interprétation notée $\{\cdot\}$ modélisant la logique **IS4 + KV**. La discussion qui a été faite dans le chapitre précédent sur certaines propriétés de ce nouveau cadre nous incite à croire qu'il ne s'agit pas du meilleur traitement que l'on peut faire de la notion de localisation. En particulier, cette localisation est présente de façon directe dans la structure algébrique, mais absente de la logique, ce qui la rend peu intuitive.

De plus, telle qu'elle est définie dans ce cadre, cette notion est trop générale. Pour illustrer ce fait, intéressons-nous à la signification de l'interprétation d'une proposition primitive $\psi \in \Psi$ suivant deux représentations différentes. D'après la définition d'une interprétation basée sur des hiérarchies de représentations, on a $\llbracket \psi \rrbracket_i = \nu_i(\psi)$. Or, pour deux indices distincts i et j , $\nu_i(\psi)$ et $\nu_j(\psi)$ ont des valeurs quelconques, et il n'existe aucune relation entre elles. Il est possible de déduire deux propriétés impliquées par cette constatation pour la localisation des propositions atomiques :

- Deux versions différentes d'une même proposition atomique n'ont aucun rapport entre elles si elles se rapportent à des représentations différentes. Ainsi, une proposition ψ

apparaissant dans le terme $K_i \psi$ ne correspond pas à la même propriété du système que dans le terme $K_j \psi$ si $i \neq j$.

- Lorsqu’une proposition atomique apparaît dans un terme φ , il faut qu’elle se rapporte à une représentation, puisque parler d’une proposition ψ en général n’a pas de sens, d’après la première remarque. Il faut donc se limiter aux termes du langage $\mathcal{L}_{\mathcal{I}, \Psi}$ où toutes les propositions atomiques sont localisées.

Pour formaliser plus précisément le premier point, nous allons introduire une notion de localisation d’un sous-terme et plus particulièrement des occurrences de propositions atomiques dans un terme. Ainsi, l’occurrence d’une proposition atomique ψ est localisée en i si elle est située à l’intérieur d’un sous-terme de la forme $K_i \varphi$, et que ce sous-terme est le plus petit de cette forme. Cette notion est illustrée dans les exemples suivants, où pour chaque proposition atomique ψ est notée avec sa localisation en indice entre crochets (et où \emptyset indique qu’elle n’apparaît pas à l’intérieur d’un K_i) :

$$\begin{aligned} & K_i \psi_{[i]} \quad K_i K_j \psi_{[j]} \wedge \psi_{[\emptyset]} \\ & K_i (\psi_{[i]} \rightarrow (K_j \psi_{[j]} \vee \psi_{[i]})) \rightarrow \psi_{[\emptyset]} \end{aligned}$$

Plutôt que de définir précisément cette notion, nous allons nous intéresser directement à celle de permutation localisée. Étant donné une permutation σ de l’ensemble des propositions atomiques Ψ , on définit la permutation de φ selon σ en i (que l’on note $\varphi[\sigma, i]$ comme le terme obtenu en partant de φ où chaque proposition atomique ψ localisée en i est remplacée par $\sigma(\psi)$). On définit cette transformation de façon formelle :

Définition 4

Étant donné un terme φ , un indice i et une permutation σ de Ψ , l’application localisée de σ en i sur φ donne le terme $\varphi[\sigma, i] = \text{Subs}(\varphi, \emptyset, \sigma, i)$ avec la fonction Subs définie ainsi :

$$\text{Subs}(\psi, j, \sigma, k) = \begin{cases} \psi & j \neq k \\ \sigma(\psi) & j = k \end{cases} \quad \psi \in \Psi$$

$$\text{Subs}(K_i \varphi, j, \sigma, k) = K_i \text{Subs}(\varphi, i, \sigma, k)$$

$$\text{Subs}(\varphi_1 \vee \varphi_2, j, \sigma, k) = \text{Subs}(\varphi_1, j, \sigma, k) \vee \text{Subs}(\varphi_2, j, \sigma, k)$$

$$\text{Subs}(\varphi_1 \wedge \varphi_2, j, \sigma, k) = \text{Subs}(\varphi_1, j, \sigma, k) \wedge \text{Subs}(\varphi_2, j, \sigma, k)$$

$$\text{Subs}(\varphi_1 \rightarrow \varphi_2, j, \sigma, k) = \text{Subs}(\varphi_1, j, \sigma, k) \rightarrow \text{Subs}(\varphi_2, j, \sigma, k)$$

$$\text{Subs}(\perp, j, \sigma, k) = \perp$$

Dans cette définition, on suppose que l’indice \emptyset , correspondant à l’absence de localisation, ne fait pas partie de l’ensemble d’indices \mathcal{I} .

Nous proposons de plus une notation sur les interprétations atomiques ν en définissant :

$$\nu[\sigma, i] = \langle j, \psi \rangle \mapsto \begin{cases} \nu_j(\psi) & i \neq j \\ \nu_i(\sigma(\psi)) & i = j \end{cases}$$

On peut alors énoncer la proposition illustrant l'indépendance des propositions atomiques de localisations différentes :

Proposition 23

Étant donné un terme φ , une permutation σ de Ψ et un indice i , on a pour toute hiérarchie de représentations \mathcal{H} , toute interprétation atomique ν et tout indice j les égalités :

$$\begin{aligned} \llbracket \varphi[\sigma, i] \rrbracket_{\mathcal{H}, \nu, j} &= \llbracket \varphi \rrbracket_{\mathcal{H}, \nu[\sigma, i], j} \\ \llbracket \text{Subs}(\varphi, j, \sigma, i) \rrbracket_{\mathcal{H}, \nu, j} &= \llbracket \varphi \rrbracket_{\mathcal{H}, \nu[\sigma, i], j} \end{aligned}$$

Preuve On prouve ce résultat par récurrence sur la taille des termes. Les cas \vee , \wedge , \rightarrow et \perp se traitent directement. Nous allons développer le cas des propositions atomiques et des opérateurs modaux. Pour les propositions atomiques, deux cas se présentent, suivant que i et j sont égaux ou non :

$$\begin{aligned} i \neq j &\Rightarrow \llbracket \text{Subs}(\psi, j, \sigma, i) \rrbracket_{\nu, j} = \llbracket \psi \rrbracket_{\nu, j} = \nu_j(\psi) = \nu[\sigma, i]_j(\psi) = \llbracket \psi \rrbracket_{\nu[\sigma, i], j} \\ i = j &\Rightarrow \llbracket \text{Subs}(\psi, j, \sigma, i) \rrbracket_{\nu, j} = \llbracket \sigma(\psi) \rrbracket_{\nu, j} = \nu_j(\sigma(\psi)) = \nu[\sigma, i]_j(\psi) = \llbracket \psi \rrbracket_{\nu[\sigma, i], j} \end{aligned}$$

Pour les opérateurs modaux, on peut écrire en utilisant l'hypothèse de récurrence :

$$\begin{aligned} \llbracket \text{Subs}(K_k \varphi, j, \sigma, i) \rrbracket_{\nu, j} &= \llbracket K_k \text{Subs}(\varphi, k, \sigma, i) \rrbracket_{\nu, j} \\ &= \{ d \mid f_{k|j}(d) \in \llbracket \text{Subs}(\varphi, k, \sigma, i) \rrbracket_{\nu, k} \} \\ &= \{ d \mid f_{k|j}(d) \in \llbracket \varphi \rrbracket_{\nu[\sigma, i], k} \} \\ &= \llbracket K_k \varphi \rrbracket_{\nu[\sigma, i], j} \end{aligned}$$

On a ainsi prouvé la seconde égalité de la proposition. Pour obtenir la première égalité, il suffit de reprendre les calculs précédents, en remarquant que si $j = \emptyset$, alors $j \neq i$. \square

Cette proposition permet d'énoncer un corollaire important, qui exhibe précisément l'indépendance de propositions atomiques de localisations différentes.

Corollaire 23.1

On a pour une proposition φ , une permutation σ de Ψ et un indice $i \in \mathcal{I}$ l'équivalence suivante :

$$\vdash_{\text{OL}} \varphi \quad \Leftrightarrow \quad \vdash_{\text{OL}} \varphi[\sigma, i]$$

L'une des conséquences de cette proposition est que l'évocation d'une proposition atomique non localisée dans une proposition logique n'a aucun sens, puisque comme on vient de le montrer, il y a autant de significations pour une proposition atomique donnée qu'il y a d'indices. On justifie donc le second point de la liste donnée p.136.

11.2 Fermeture des propositions

Nous allons donc maintenant nous restreindre à l'étude des propositions où toute occurrence d'une proposition atomique se trouve à l'intérieur d'un sous-terme de la forme $K_i \varphi$. De tels termes peuvent être qualifiés de fermés, puisque les propositions atomiques ne peuvent pas s'échapper. Il est facile de définir cette notion de fermeture.

Le critère de fermeture

La fermeture d'un terme peut être caractérisée en utilisant la notion de proposition atomique libre, proche de la notion de variable libre en λ -calcul. Dans ce dernier, une variable est libre si elle n'apparaît pas à l'intérieur d'un λ . Dans le cas présent, on peut de même parler d'une proposition libre si elle n'apparaît pas à l'intérieur d'un K_i . Cela suggère les définitions suivantes :

Définition 5

L'ensemble des propositions atomiques ayant au moins une occurrence libre dans terme φ est égale à $\text{fp}(\varphi)$, où la fonction fp est défini de façon inductive :

$$\begin{aligned} \text{fp}(\psi) &= \{\psi\} & \psi \in \Psi \\ \text{fp}(\varphi_1 \vee \varphi_2) &= \text{fp}(\varphi_1) \cup \text{fp}(\varphi_2) \\ \text{fp}(\varphi_1 \wedge \varphi_2) &= \text{fp}(\varphi_1) \cup \text{fp}(\varphi_2) \\ \text{fp}(\varphi_1 \rightarrow \varphi_2) &= \text{fp}(\varphi_1) \cup \text{fp}(\varphi_2) \\ \text{fp}(K_i \varphi) &= \emptyset \\ \text{fp}(\perp) &= \emptyset \end{aligned}$$

Définition 6

Un terme $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$ est fermé si et seulement si $\text{fp}(\varphi) = \emptyset$.

On obtient donc une première définition de la fermeture d'un terme, qui de plus correspond bien à l'idée que l'on a énoncée plus tôt, qui impose que toute occurrence d'une proposition atomique soit localisée (ce que l'on appelle ici fermée).

Il est possible de contraindre encore plus les termes de telle sorte que chaque proposition atomique apparaissant dans un terme se trouve accolée à un terme K_i indiquant directement sa localisation.

Le langage $LL_{\mathcal{I},\Psi}$

Cette seconde approche de la localisation peut être obtenue en imposant au langage d'utiliser non pas les $\psi \in \Psi$ comme formules atomiques, mais à la place des termes de la forme $K_i \psi$ pour $i \in \mathcal{I}$ et $\psi \in \Psi$. On peut définir exactement l'ensemble des propositions ayant cette forme :

Définition 7

Le langage $LL_{\mathcal{I},\Psi}$ est le plus petit ensemble E vérifiant :

$$\begin{aligned} & \{K_i \psi \mid i \in \mathcal{I}, \psi \in \Psi\} \cup \{\perp\} \subseteq E \\ & \forall \varphi \in E, i \in \mathcal{I}, K_i \varphi \in E \\ & \forall \varphi_1, \varphi_2 \in E, \{\varphi_1 \vee \varphi_2, \varphi_1 \wedge \varphi_2, \varphi_1 \rightarrow \varphi_2\} \subseteq E \end{aligned}$$

D'après les deux définitions de localisation que l'on vient de donner, il semble que l'appartenance à $LL_{\mathcal{I},\Psi}$ est plus forte que la fermeture. On a en effet :

Proposition 24

$$\forall \varphi \in LL_{\mathcal{I},\Psi}, \text{fp}(\varphi) = \emptyset$$

Cependant, du point de vue logique, ces notions sont similaires, puisque toute formule fermée est équivalente à une formule de $LL_{\mathcal{I},\Psi}$:

Proposition 25

$$\forall \varphi \in \mathcal{L}_{\mathcal{I},\Psi}, (\text{fp}(\varphi) = \emptyset \Rightarrow \exists \varphi' \in LL_{\mathcal{I},\Psi} : \vdash_{\text{OL}} \varphi \leftrightarrow \varphi')$$

Preuve Ce résultat vient du fait qu'il est possible, dans un terme fermé, de faire "descendre" les K_i jusqu'aux propositions atomiques. On a en effet :

$$\begin{aligned} & \vdash_{\text{OL}} K_i (\varphi_1 \vee \varphi_2) \leftrightarrow K_i \varphi_1 \vee K_i \varphi_2 \\ & \vdash_{\text{OL}} K_i (\varphi_1 \wedge \varphi_2) \leftrightarrow K_i \varphi_1 \wedge K_i \varphi_2 \\ & \vdash_{\text{OL}} K_i (\varphi_1 \rightarrow \varphi_2) \leftrightarrow K_i (K_i \varphi_1 \rightarrow K_i \varphi_2) \end{aligned}$$

On peut donc appliquer ce genre de transformations successivement pour, partant d'un terme fermé, aboutir à un élément de $LL_{\mathcal{I},\Psi}$, chacune des étapes préservant l'équivalence. \square

11.3 Les observations atomiques

L'utilisation du langage $LL_{\mathcal{I},\Psi}$ modifie la façon dont il faut considérer la logique **OL** pour la modélisation de connaissances basées sur des observations. En effet, jusqu'à présent, les propositions atomiques, éléments de Ψ , pouvaient être prises pour représenter les informations de base dont on dispose à propos du système, c'est-à-dire les résultats d'observations. Ces éléments étaient utilisés comme briques de base, et toute proposition logique concernant le système étudié apparaît alors comme une mise en relation d'observations.

Puisque l'on impose la localisation des observations, on modifie ces briques de base, celles-ci étant remplacées par des termes de la forme $K_i \psi$. Pour illustrer la façon dont cela change le comportement des propriétés impliquant des propositions atomiques, nous allons introduire une nouvelle notation suggérant que l'on a un nouvel ensemble de propositions primitives. Ainsi, définissons ψ_i comme étant une version abrégée pour $K_i \psi$.

Dans le cadre de la logique **OL**, il est possible d'énoncer de nombreuses propriétés vérifiées par ce que l'on peut donc considérer comme de nouvelles propositions atomiques. Ainsi, on peut exprimer des propositions les concernant ressemblant à l'axiome **T**, ce qui peut donner un statut de "réalité" à ces propositions :

$$\vdash K_i \psi_j \rightarrow \psi_j$$

On peut déduire en particulier que l'axiome de confiance entre agents est alors vérifié :

$$\vdash K_i K_j \psi_k \rightarrow K_i \psi_k$$

De plus, puisque la localisation des propositions de la forme ψ_i apparaît explicitement, on peut énoncer des propriétés quand à leur observabilité :

$$\vdash \psi_i \leftrightarrow K_i \psi_i$$

Pour une hiérarchie de représentations particulière et en utilisant l'ordre \triangleleft sur les indices, on peut même écrire :

$$\forall j \triangleleft i, \vdash \psi_i \leftrightarrow K_j \psi_i$$

Observations négatives des propriétés

Dans les propriétés que l'on vient d'énoncer, on relie la véracité d'une proposition ψ_i à son observation. Ainsi, si l'on a $K_j \psi_i$, alors ψ_i est vraie. Réciproquement, si ψ_i est vraie, alors cette propriété est observable dans au moins une représentation. Mais ces différents cas se rapportent à une situation où une proposition ψ_i est bien vérifiée par le système.

On peut aussi s'intéresser au cas où la propriété ψ_i est fautive, où sa négation est vraie. Supposons tout d'abord que l'on a $K_j \neg\psi_i$. Cela ne suffit pas à assurer que ψ_i n'est pas vérifiée

par le système. En effet, la logique OL ne permet pas de déduire $\neg\psi_i$ de $K_j \neg\psi_i$. Il est même possible que ψ_i soit vérifiée en même temps que $K_j \neg\psi_i$, puisque $\neg(\psi_i \wedge K_j \neg\psi_i)$ n'est pas non plus prouvable dans OL. En particulier, on peut construire une hiérarchie \mathcal{H} telle que $\mathcal{H} \models \psi_i \wedge K_j \neg\psi_i$, comme la figure 11.1 le montre.

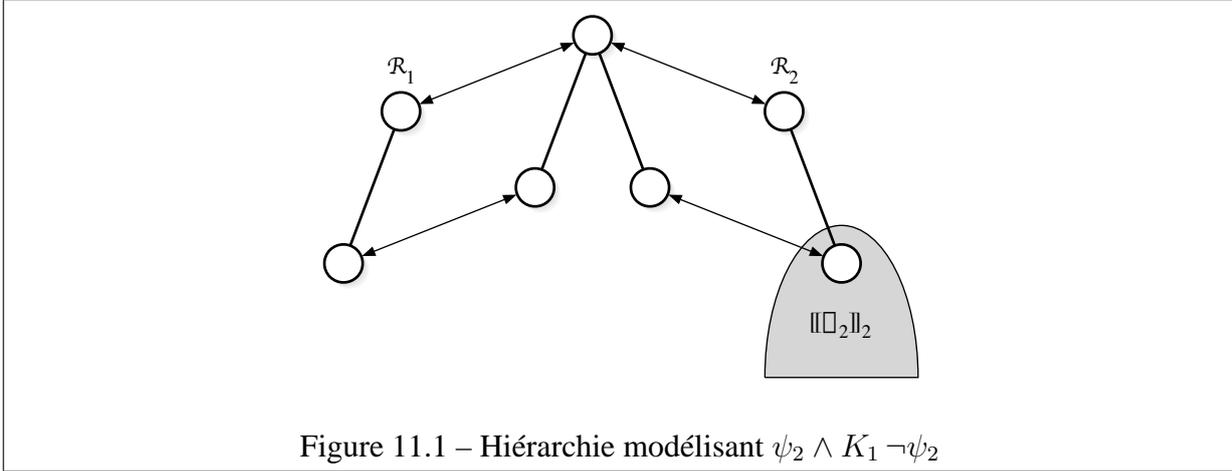


Figure 11.1 – Hiérarchie modélisant $\psi_2 \wedge K_1 \neg\psi_2$

Aussi, le terme $\neg\psi_i$ n'est pas suffisant pour s'assurer que la proposition ψ_i n'est pas vérifiée. Il faut de plus imposer que ce terme soit observé depuis la représentation où est localisée ψ_i , c'est-à-dire la représentation \mathcal{R}_i elle-même. Pour cette raison, le terme à utiliser pour désigner la négation d'une proposition ψ_i est $K_i \neg\psi_i$ et non simplement $\neg\psi_i$. On peut donc enrichir la notation ψ_i en définissant :

$$\psi_i^+ = K_i \psi \quad \psi_i^- = K_i \neg K_i \psi$$

Nous pensons que ces deux définitions permettent d'exprimer correctement l'idée d'une propriété qui est vérifiée ou non par le système. Pour illustrer cela, nous allons présenter quelques propriétés concernant cette notion.

Quelques propriétés concernant les ψ_i^+ et ψ_i^-

Tout d'abord, comme dans le cas précédent avec ψ_i , on précise le rapport entre l'observation d'une proposition de la forme ψ_i^p et sa validité (ainsi que son observabilité) :

$$\begin{aligned} \vdash K_j \psi_i^p &\rightarrow \psi_i^p & p \in \{+, -\} \\ \vdash \psi_i^p &\leftrightarrow K_i \psi_i^p & p \in \{+, -\} \end{aligned}$$

Ainsi, de telles propositions peuvent être vues comme des éléments de réalité. Si elles sont observées, alors elles sont prouvables. De plus, réciproquement, si elles sont prouvables, alors il est possible de les observer. La présence de la localisation en indice permet de plus la détermination d'une représentation dans laquelle elles sont observables, et cela de façon interne à la logique, sans avoir recours au modèle utilisé.

Les propositions ψ_i^+ et ψ_i^- sont définies comme correspondant au fait qu'une propriété ψ définie dans la représentation \mathcal{R}_i est vraie ou fausse. En général, en logique intuitionniste, une proposition φ et sa négation $\neg\varphi$ vérifient les propriétés suivantes : elles ne sont pas vraies en même temps ($\vdash_{\text{IL}} \neg(\varphi \wedge \neg\varphi)$) mais au moins l'une des deux est vraie, même si on ne peut pas dire laquelle ($\vdash_{\text{IL}} \neg\neg(\varphi \vee \neg\varphi)$).

Pour les propositions atomiques ψ_i^p , il est aussi possible d'énoncer des propriétés similaires mettant en relation les versions positives et négatives d'une même propriété ψ . Tout d'abord, elles ne peuvent pas être vérifiées en même temps :

$$\vdash \neg(\psi_i^+ \wedge \psi_i^-)$$

On peut de plus énoncer le fait que l'une des deux est forcément vérifiée par un système donné. Cependant, la proposition $\neg\neg(\psi_i^+ \vee \psi_i^-)$ est trop générale, et non prouvable dans **OL**. La raison pour cela est que ce résultat n'est pas vérifié dans toutes les représentations. Si l'on reprends la figure 11.1, l'interprétation de $K_2\psi$ vérifie $\llbracket \psi_2^+ \rrbracket_1 = \emptyset$ et $\llbracket \psi_2^- \rrbracket_2 = \emptyset$. On en déduit donc $\llbracket \psi_2^+ \vee \psi_2^- \rrbracket_1 = \emptyset$, ce qui montre la non prouvabilité de $\neg\neg(\psi_2^+ \vee \psi_2^-)$. La raison en est une fois de plus que toutes les représentations ne sont pas adaptées pour décider de la validité d'une proposition. Cependant, si l'on se place dans la représentation où est localisée la proposition ψ , on retrouve la propriété attendue. Ainsi, on peut écrire :

$$\vdash K_i \neg\neg(\psi_i^+ \vee \psi_i^-)$$

Tous ces aspects du comportement logique des termes de la forme ψ_i^+ et ψ_i^- nous montrent qu'ils constituent une formalisation adéquate de la notion de propriétés atomiques si l'on impose qu'elles soient localisées. Cela justifie de plus la restriction que l'on a imposé au langage en se limitant au sous-ensemble des propositions fermées, et même à $\text{LL}_{\mathcal{I},\Psi}$.

11.4 La connaissance accessible par observation

Pour reprendre l'un des postulats de départ, toutes les informations que l'on a sur l'état du système proviennent d'observations. Exprimées dans le langage **LL**, toutes ces observations se rapportent à une représentation précise, elles sont localisées. Ainsi, toute occurrence d'une proposition atomique ψ est associée à un indice i en apparaissant sous la forme $K_i\psi$.

L'observation d'une propriété depuis une représentation donnée n'est cependant pas la forme d'information la plus intéressante à manipuler. En effet, comme on l'a développé plus tôt, il n'y a pas assez de liens entre les différentes façons d'observer une même proposition.

Ainsi, dans le cas général, l'observation d'une propriété ou de sa négation peut ne pas correspondre à sa vérification effective par le système, et il est possible qu'une hiérarchie vérifie à

la fois $K_i \psi$ et $K_j \neg\psi$. Cela implique de plus que les agents ne peuvent rien communiquer entre eux, puisqu'ils peuvent être en désaccord même sur les informations les plus élémentaires.

La situation est bien plus favorable si l'on inclut la localisation des propositions. En forçant de plus l'utilisation de termes de la forme $K_i \neg K_i \psi$ pour exprimer la négation de $K_i \psi$, on obtient un ensemble de nouvelles propositions atomiques qui correspondent à des informations plus significatives. Comme on l'a vu, en utilisant les notations ψ_i^+ et ψ_i^- , ces propositions atomiques correspondent à des informations fiables sur l'état du système. Elles vérifient des propriétés calquées sur l'axiome **T** :

$$\vdash K_i \psi_j^p \rightarrow \psi_j^p$$

Cela implique qu'il s'agit d'une information communicable, puisque l'on peut prouver :

$$\vdash K_i K_j \psi_k^p \rightarrow K_i \psi_k^p$$

Et la consistance globale est vérifiée par ces propositions :

$$\vdash \neg(K_i \psi_k^+ \wedge K_j \psi_k^-)$$

Toutes ces propriétés ne s'appliquent pas en général à tous les termes de $LL_{\mathcal{I},\Psi}$. La distributivité de la disjonction et de la conjonction par rapport aux opérations modales font que tout terme formé à partir des ψ_i^p en appliquant les opérations \wedge et \vee constitue une information "fiable", en ce sens qu'elles vérifient les mêmes propriétés que les propositions atomiques.

C'est la présence au sein d'un terme de l'implication \rightarrow qui change le statut de l'information représentée par ce terme, puisqu'alors on ne peut pas prouver en général $\vdash K_i \varphi \rightarrow \varphi$. En conséquence, des propositions contenant des implications ne sont pas communicables dans le cas général et l'observation d'une telle propriété n'implique pas sa vérification. Dans la logique **OL**, on ne peut en effet pas prouver la proposition :

$$\neg(K_i(\varphi \rightarrow \psi) \wedge \neg(\varphi \rightarrow \psi))$$

Pour illustrer cela, la figure 11.2 montre un exemple de hiérarchie de représentations \mathcal{H} telle que :

$$\mathcal{H} \models K_1(K_2 \varphi \rightarrow K_2 \psi) \wedge \neg(K_2 \varphi \rightarrow K_2 \psi)$$

Comme la négation d'un terme est une forme d'implication, elle n'est pas communicable en général. C'est la raison pour laquelle dans le cas des propositions atomiques, on a défini ψ_i^- comme $K_i \neg K_i \psi$.

Nous allons dans le prochain chapitre étudier comment caractériser de façon plus générale des propositions ayant un comportement analogue aux termes de la forme ψ_i^p . Cela nous conduira à définir la notion d'*objectivité* d'une proposition.

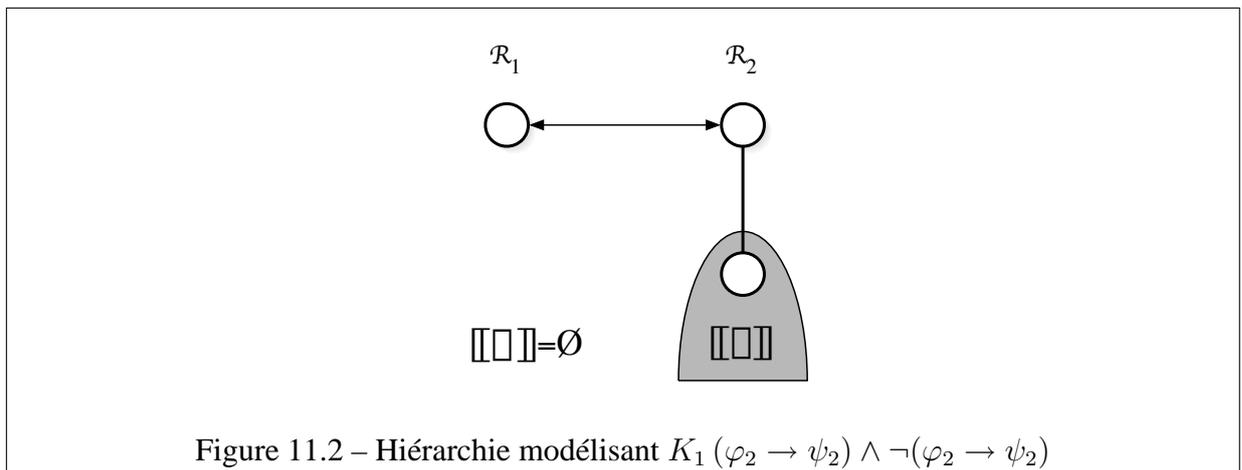


Figure 11.2 – Hiérarchie modélisant $K_1(\varphi_2 \rightarrow \psi_2) \wedge \neg(\varphi_2 \rightarrow \psi_2)$

Chapitre 12

Vers une théorie de la connaissance observable

En s'inspirant de l'étude faite précédemment, et de la comparaison entre les logiques OL et IS4 + KV, ainsi que des propriétés vérifiées par les termes de la forme ψ_i^p , il est possible d'identifier une propriété particulière que possèdent certains termes, une propriété qui traduit le fait que l'information est utilisable et échangeable.

Pour expliquer cela, on peut revenir à l'un des postulats de départ, qui affirme que toute connaissance provient d'une observation. De façon logique, cela légitime le fait que les éléments de base que l'on manipule sont des termes de la forme $K_i \varphi$, où i est l'indice correspondant au point d'observation utilisé, et φ est l'information en question.

Mais pour que cette information puisse être utilisable en général comme caractérisation de l'état du système, il faut qu'elle puisse être dissociée de sa représentation de départ. En effet, si l'on associe un agent à chacune des représentations, une information prouvée par un agent \mathcal{A}_i (correspondant à la représentation \mathcal{R}_i) ne peut être communiquée à un agent \mathcal{A}_j que si l'information telle qu'elle est formulée est exploitable par ce dernier. C'est à ce genre d'informations que nous allons nous intéresser maintenant.

12.1 Objectivité d'une proposition

Pour cela, nous allons définir un critère d'*objectivité*. Une proposition sera dite objective si elle ne dépend pas d'une représentation particulière. Une telle proposition peut être utilisée sans risque dans tout raisonnement concernant le système.

12.1.1 Quelques approches de l'objectivité

Il est possible de traduire cette notion de différentes façons au sein des logiques que l'on a rencontrées. On peut tout d'abord s'intéresser à $IS4 + KV$. Cette logique est apparue à deux reprises au cours de ce travail, la première fois dans le cadre de l'approximation d'une représentation par une autre, et dans le cadre des hiérarchies consistantes.

Dans le cadre de l'approximation d'une représentation par une autre, il existe une représentation plus précise que l'autre, ce qui implique que toute information peut être représentée dans celle-ci. Dans ce cas, toute information peut être vue comme objective, puisqu'il n'y a pas de problème de localisation des propositions, étant toutes exprimées dans la même représentation.

Cette situation se retrouve dans les hiérarchies consistantes, même si elle apparaît de façon moins directe. En effet, on peut considérer qu'il existe une représentation plus précise que toutes les autres. Une telle représentation peut être construite facilement à partir de la hiérarchie consistante, en prenant le produit réduit. D'après la fonction de localisation utilisée pour l'interprétation $\{\cdot\}$, on voit que là aussi, toute information se comporte comme si elle était exprimée dans cette représentation plus précise que toutes les autres.

Dans les deux cas, nous pensons que cette situation se traduit dans la logique par le fait que l'axiome $T : K_i \varphi \rightarrow \varphi$ est vérifié. Cet axiome peut en effet être traduit simplement par le fait que si une propriété est observée (soit $K_i \varphi$), alors elle est effectivement vérifiée par le système (soit $K_i \varphi \rightarrow \varphi$). Il implique de plus le fait que les agents peuvent se faire confiance entre eux, puisque l'axiome $C : K_i K_j \varphi \rightarrow K_i \varphi$ est alors vérifié.

La vérification de l'un ou l'autre de ces axiomes apparaît comme une bonne caractérisation de la notion d'objectivité que nous sommes en train de définir. De plus, ils ne sont pas vérifiés en général par la logique OL , ce qui légitime l'existence de cette notion dans le cadre des hiérarchies de représentations.

12.1.2 La notion de contexte

La notion d'objectivité dans le cas général, où l'on fait totalement abstraction de la structure, n'est a priori pas très intéressante à étudier, puisque qu'une grande part des affirmations que l'on peut faire sur l'état du système utilise une part d'information structurelle. Pour tenir compte de cette information, nous allons utiliser un ensemble éventuellement infini $\Gamma = \{\gamma_i\}$ de propositions appelé *contexte*. Alors, nous dirons que dans le contexte Γ , une proposition φ est vérifiée, ce que nous noterons $\Gamma \vdash \varphi$, si et seulement si :

$$\forall i \in \mathcal{I}, \forall d \in \mathcal{R}_i, d \in \llbracket \Gamma \rrbracket_i \Rightarrow d \in \llbracket \varphi \rrbracket_i$$

Dans cette définition, s'inspirant de la construction de conjonction, $\llbracket \Gamma \rrbracket_i$ représente l'ensemble $\bigcap_j \llbracket \gamma_j \rrbracket_i$. Si Γ est fini, la définition précédente est équivalente à la prouvabilité de $\vdash (\gamma_1 \wedge \dots \wedge \gamma_n) \rightarrow \varphi$.

On doit cependant ajouter une contrainte supplémentaire à la définition de contexte. En effet, son rôle est de représenter l'information structurelle associée à une hiérarchie de représentations donnée. Or, ce type d'information est indépendant de toute localisation, et doit être exprimée de façon équivalente dans toutes les représentations. On impose donc :

$$\forall i, j, \forall d \in \mathcal{R}_i, d \in \llbracket \Gamma \rrbracket_i \Leftrightarrow f_{j|i}(d) \in \llbracket \Gamma \rrbracket_j$$

Cette condition donne au contexte Γ le statut de connaissance commune, notion étudiée chez J. Halpern (voir par exemple [39]). Dans notre cas, cela signifie que l'on a pour toute suite finie d'indices i_1, \dots, i_n , l'équivalence :

$$\Gamma \Leftrightarrow K_{i_1} \dots K_{i_n} \Gamma$$

Cela permet d'assurer que l'information structurelle exprimée par Γ est connue, et peut être utilisée quelque soit le niveau d'imbrication d'opérateurs K_i apparaissant dans un raisonnement.

12.1.3 Définition de la notion d'objectivité

En utilisant la notion de contexte, nous pouvons maintenant nous intéresser à la définition formelle du critère d'objectivité. Comme nous l'avons vu, deux axiomes apparaissent en tant que candidats pour cette définition : l'axiome **T** d'une part, et l'axiome **C** d'autre part, qui est une version moins forte du premier. On peut donc écrire les deux conditions :

$$\forall i, \Gamma \vdash K_i \varphi \rightarrow \varphi \tag{12.1}$$

$$\forall i, j, \Gamma \vdash K_i K_j \varphi \rightarrow K_i \varphi \tag{12.2}$$

Tout d'abord, puisque l'axiome **T** est plus fort que **C**, on peut supposer que la condition 12.1 implique l'équation 12.2. Cependant, puisqu'à chaque fois, il y a quantification universelle sur les indices de \mathcal{I} , on peut montrer que les deux conditions sont en fait équivalentes. Pour montrer cela, écrivons-les sous forme ensembliste en utilisant l'interprétation $\llbracket \cdot \rrbracket$. On obtient :

$$\forall i, j, \forall d \in \llbracket \Gamma \rrbracket_j, f_{i|j}(d) \in \llbracket \varphi \rrbracket_i \Rightarrow d \in \llbracket \varphi \rrbracket_j \tag{12.3}$$

$$\forall i, j, k, \forall d \in \llbracket \Gamma \rrbracket_k, f_{j|i} \circ f_{i|k}(d) \in \llbracket \varphi \rrbracket_j \Rightarrow f_{i|k}(d) \in \llbracket \varphi \rrbracket_i \tag{12.4}$$

La condition 12.3 s'obtient à partir de 12.4 en posant $j = k$. Réciproquement, soit d tel que $d \in \llbracket \Gamma \rrbracket_k$ et $f_{i|j} \circ f_{j|k}(d) \in \llbracket \varphi \rrbracket_i$. Si φ vérifie la condition 12.3, alors comme par hypothèse sur

les contextes on a $d \in \llbracket \Gamma \rrbracket_k \Rightarrow f_{j|k}(d) \in \llbracket \Gamma \rrbracket_j$, on peut appliquer cette condition à $f_{j|k}(d)$, ce qui permet de déduire $f_{j|k}(d) \in \llbracket \varphi \rrbracket_j$, qui est le résultat espéré. On vient donc de montrer que les deux versions de la notion d'objectivité 12.1 et 12.2 sont équivalentes. On peut donc les utiliser comme définition.

Définition 8

Étant donné un contexte Γ , une proposition φ est objective pour Γ si et seulement si elle vérifie :

$$\forall i \in \mathcal{I}, \Gamma \vdash K_i \varphi \rightarrow \varphi$$

Dans la suite, nous noterons $\text{Obj}(\Gamma)$ l'ensemble des propositions objectives pour un contexte Γ donné. Il est possible, comme nous allons le voir, d'énoncer quelques propriétés simples vérifiées par cet ensemble.

12.1.4 Quelques propriétés de $\text{Obj}(\Gamma)$

À la manière de la définition du langage $\mathcal{L}_{\mathcal{I},\Psi}$, nous allons commencer par indiquer un certain nombre de types de propositions objectives, qui peuvent être vues comme des briques de bases, puis nous indiquerons dans quelle mesure il est possible de combiner des propositions pour que le résultat soit lui aussi objectif.

Les briques de base

On peut identifier trois types de propositions qui sont objectives pour un contexte donné Γ . Tout d'abord, tout ce qui est déductible du contexte est objectif, puisque si l'on a une proposition φ telle que $\Gamma \vdash \varphi$, alors pour tout i , on a :

$$\Gamma \vdash K_i \varphi \rightarrow \varphi$$

Ces éléments dépendent directement du contexte Γ . D'autres propriétés qui viennent de la définition de la logique **OL**, peuvent aussi être identifiées comme objectives quelque soit le contexte.

L'axiome **D** : $K_i \varphi \rightarrow \neg K_i \neg \varphi$, nous indique tout d'abord que la proposition absurde \perp est toujours objective, puisque en conjonction avec l'axiome **K**, l'une des conséquences de cet axiome est la proposition $\neg K_i \perp$, qui peut encore s'écrire sous la forme $K_i \perp \rightarrow \perp$, ce qui permet de conclure à l'objectivité de \perp .

Un autre axiome de **OL** permet lui aussi de fournir un ensemble de propositions objectives ne dépendant pas du contexte. Il s'agit de l'axiome **T₂** : $K_i K_j \varphi \rightarrow K_j \varphi$. Ce dernier nous montre en effet que toute proposition de la forme $K_i \varphi$ est objective. On retrouve ici les termes de la forme ψ_i^p que nous avons introduits dans le chapitre précédent.

On peut donc écrire, pour résumer :

$$\begin{aligned} \forall \varphi, \Gamma \vdash \varphi &\Rightarrow \varphi \in \text{Obj}(\Gamma) \\ \perp &\in \text{Obj}(\Gamma) \\ \forall \varphi, i, K_i \varphi &\in \text{Obj}(\Gamma) \end{aligned}$$

Construction récursive

Le langage des termes $\mathcal{L}_{\mathcal{I}, \Psi}$ est constitué, en plus des propositions de base, de cinq constructeurs. Le cas de la proposition absurde \perp et des opérateurs modaux K_i ont été traités en tant que briques de bases. Il reste les connecteurs binaires \vee , \wedge et \rightarrow .

Puisque l'on a $\vdash K_i(\varphi \wedge \psi) \rightarrow K_i \varphi \wedge K_i \psi$, on montre facilement que la conjonction de deux propositions objectives est aussi objective :

$$\left. \begin{array}{l} \Gamma \vdash K_i \varphi \rightarrow \varphi \\ \Gamma \vdash K_i \psi \rightarrow \psi \end{array} \right\} \Rightarrow \Gamma \vdash K_i \varphi \wedge K_i \psi \rightarrow \varphi \wedge \psi \Rightarrow \Gamma \vdash K_i(\varphi \wedge \psi) \rightarrow \varphi \wedge \psi$$

L'axiome **KV** : $K_i(\varphi \vee \psi) \rightarrow K_i \varphi \vee K_i \psi$ permet d'appliquer le même raisonnement à une disjonction de propositions objectives. On peut donc écrire :

$$\begin{aligned} \forall \varphi, \psi \in \text{Obj}(\Gamma), \varphi \vee \psi &\in \text{Obj}(\Gamma) \\ \forall \varphi, \psi \in \text{Obj}(\Gamma), \varphi \wedge \psi &\in \text{Obj}(\Gamma) \end{aligned}$$

Le cas de l'implication

Le raisonnement que l'on vient de faire pour la disjonction et la conjonction ne s'applique pas à l'implication. Nous en avons d'ailleurs déjà donné un contre-exemple dans la figure 11.2 p.145, où la représentation vérifie :

$$\mathcal{H} \models K_1(K_2 \varphi \rightarrow K_2 \psi) \wedge \neg(K_2 \varphi \rightarrow K_2 \psi)$$

On l'a vu, les sous-propositions $K_2 \varphi$ et $K_2 \psi$ sont objectives, aussi on a montré que l'implication de propositions objectives n'est pas objective en général.

Une courte étude de la preuve possible de l'objectivité d'un terme $\varphi \rightarrow \psi$ nous donne une indication de la raison pour laquelle elle n'est pas toujours possible. En utilisant l'axiome **K**, on obtient :

$$\Gamma, K_i(\varphi \rightarrow \psi) \vdash K_i \varphi \rightarrow K_i \psi$$

Tout d'abord, si le second sous-terme ψ est objectif, on obtient :

$$\Gamma, K_i(\varphi \rightarrow \psi) \vdash K_i \varphi \rightarrow \psi$$

Dans ce cas, pour arriver à l'objectivité de $\varphi \rightarrow \psi$, une condition suffisante est que φ vérifie $\forall i, \Gamma \vdash \varphi \rightarrow K_i \varphi$, qui est la réciproque du critère d'objectivité. La raison est donc intuitivement une question de variance pour le membre gauche d'une implication. Le membre droit est covariant, donc l'hypothèse d'objectivité peut s'appliquer, mais le membre gauche est contravariant, ce qui fait qu'il faut qu'il vérifie la réciproque.

Clôture déductive

Un autre aspect de l'implication qui caractérise l'ensemble $\text{Obj}(\Gamma)$ est le fait qu'il n'est pas clos par déduction. Pour montrer cela, on remarque que la logique intuitionniste, et donc la logique de l'observation, permettent de prouver :

$$\forall \varphi, \vdash_{\text{OL}} \perp \rightarrow \varphi$$

Cela signifie que toute proposition de la forme $\perp \rightarrow \varphi$ est objective. Comme \perp est elle-même objective, et puisqu'il existe des propositions qui ne le sont pas, on en déduit que l'ensemble des propositions objectives n'est pas clos pour la déduction.

Ce survol des propriétés vérifiées par les ensembles $\text{Obj}(\Gamma)$ montre que la notion d'objectivité est stable par toute opération covariante (comme la conjonction et la disjonction), et que toute proposition de la forme $K_i \varphi$ est objective, ce que l'on peut interpréter comme le fait que si l'on fournit suffisamment d'informations (comme la représentation utilisée pour valider la proposition φ), alors cette information est objective, c'est-à-dire qu'elle aura la même interprétation dans toutes les représentations.

Nous pensons donc être parvenu à un critère de sélection des propositions particulièrement intéressant, et nous allons reprendre l'étude de la logique OL en tenant compte de celui-ci.

12.2 L'objectivité dans le cadre de la logique de l'observation

Dans les chapitres précédents, nous avons critiqué le manque de puissance expressive de la logique de l'observation OL, en la comparant notamment à la logique modale intuitionniste IS4 + KV. La discussion s'était en particulier portée sur la vérification ou non de l'axiome T : $K_i \varphi \rightarrow \varphi$. Par la suite, ce dernier a d'ailleurs été retenu comme définition de la notion d'objectivité pour une proposition.

À la lumière de cette notion, nous pouvons comparer les deux logiques et leur adéquation pour la représentation de connaissances observables. Pour $IS4 + KV$, toutes les propositions sont objectives, c'est-à-dire qu'elles peuvent être utilisées telles qu'elles pour spécifier un ensemble d'états du système, et ne sont pas rattachées à une représentation particulière. Mais comme nous l'avons vu plus haut, cela vient du fait que pour les structures modélisant cette logique, la réalité, ou en tout cas une représentation la plus complète possible de celle-ci, est présente dans ces structures. Il y a donc une référence absolue à laquelle se rapportent toutes les descriptions. Or, cela est une contrainte très forte, puisqu'elle impose un vocabulaire commun et absolu à tous les agents pouvant observer le système. Cela ne correspond pas à l'approche que nous avons donné de la formalisation d'un système, où l'on commence par assembler diverses sous-parties (en utilisant le produit de hiérarchies de représentations) avant d'établir des relations entre ces parties (en appliquant des raffinages).

Le fait que dans la logique OL toutes les propositions ne sont pas objectives montre au contraire que les hiérarchies de représentations sont une structure algébrique possédant la souplesse voulue pour la modélisation d'un système. En particulier, on peut associer à chaque représentation un langage particulier. Cet aspect des hiérarchies est apparu dans les discussions précédentes, lorsque nous avons remarqué qu'il n'existe absolument aucun lien entre les différentes interprétations $\{\llbracket \psi \rrbracket_i\}$ d'une même proposition atomique ψ . La contre-partie de cette souplesse est naturellement le fait que toutes les propositions ne sont pas objectives, c'est-à-dire qu'elles ne peuvent pas toutes être utilisées hors contexte, en ne connaissant pas la représentation où elles sont valides.

Une analyse plus fine de la notion d'objectivité montre cependant qu'elle ne constitue pas un argument indiquant une faiblesse de la logique OL par rapport à $IS4+KV$. On peut pour cela se pencher sur l'information structurelle relative au système. Ce type d'information est disponible quelque soit la représentation choisie, aussi le langage relatif à chaque représentation est connu partout. Ainsi, si une proposition φ valide dans une représentation \mathcal{R}_i n'est pas objective, il suffit d'indiquer explicitement sa localisation en utilisant le terme $K_i \varphi$ pour qu'elle le devienne. On retrouve donc le fait que toutes les propositions de la forme $K_i \varphi$ sont objectives. En utilisant un vocabulaire d'expériences physiques, cela revient à changer une affirmation du type "on mesure une valeur de 15 volts" en "en effectuant une mesure suivant la procédure i , on trouve une valeur de 15 volts". Puisque les protocoles expérimentaux sont connus de tous, l'intérêt de la seconde formulation est évident.

Ainsi, on voit que le fait que toutes les propositions ne sont pas objectives ne fait que traduire la souplesse d'utilisation des hiérarchies de représentations, en n'imposant pas des conditions *a priori* qui devraient être vérifiées par toutes les représentations. De plus, puisque toute proposition peut facilement être transformée en une proposition objective, cela n'entraîne pas pour

autant une réduction de l'expressivité de cette modélisation. Au contraire, nous avons identifié un critère permettant de s'assurer qu'une propriété est correctement formulée ou non, en restant dans un cadre général. Ce critère suggère qu'il faut fournir suffisamment d'indications sur le point de vue utilisé pour énoncer la proposition considérée pour que celle-ci puisse être reprise telle qu'elle n'importe où.

Nous pensons donc être arrivés à un compromis raisonnable pour définir des informations communicables. Si toutes les propositions ne sont pas objectives, c'est parce que l'on n'a pas imposé de conditions fortes sur les représentations, et puisqu'il est facile de rendre une proposition objective, cela signifie que cette absence de conditions permet néanmoins de pouvoir exprimer toutes les propriétés que l'on souhaite sur le système.

Chapitre 13

Conclusion

L'objectif de ce travail était de fournir une base logique et formelle aux notions d'observation et de connaissance acquise par observation. Partant du postulat que toute connaissance que l'on peut avoir du monde qui nous entoure ou d'un sujet d'étude particulier provient uniquement des informations que l'on obtient en interagissant avec celui-ci, nous avons défini une structure théorique précise rendant compte de cette approche.

Contributions

Nous sommes partis de considérations très générales sur la notion de connaissance partielle. Cela nous a conduit à la définition d'une structure algébrique nommée *représentation* permettant de modéliser la connaissance accessible depuis une méthode d'observation. Nous avons alors enrichi cette structure en autorisant la présence de plusieurs méthodes d'observations, et en définissant un moyen algébrique simple de relier les représentations correspondantes. Ces considérations ont finalement conduit à la structure que nous avons appelé *hiérarchies de représentations*.

Ces structures ont, comme nous l'avons montré, une grande souplesse d'utilisation, et les outils que nous avons introduits pour les manipuler montrent qu'il est possible de les utiliser pour modéliser une grande variété de systèmes, puisque la méthodologie présentée pour cela (la constitution d'un produit de représentations auquel on applique ensuite un ensemble de raffinement) est très générale.

Nous avons alors effectué une étude logique du comportement de la connaissance dans ce contexte. Pour cela, nous avons défini une interprétation du langage propositionnel $\mathcal{L}_{\mathcal{I},\Psi}$ vers les hiérarchies de représentations, en introduisant aussi peu d'hypothèses supplémentaires que possible. La seule dont nous nous sommes servis est celle stipulant que la validité d'une propriété est préservée en cas d'ajout d'information. On impose ici une notion de validité assez forte,

qui exprime le fait que l'on se base sur les observations et rien d'autre pour prouver une propriété. Cela exclut totalement toute forme de raisonnement a priori ou par défaut. L'étude des propriétés de l'interprétation ainsi définie nous conduit à considérer une logique que nous qualifions d'observationnelle (notée **OL**). Cette logique donne lieu à de nombreux commentaires intéressants quant au comportement de la connaissance telle qu'elle est modélisée.

Tout d'abord, elle est basée sur la logique intuitionniste et non la logique classique. Cela signifie que l'on exclut tout raisonnement par l'absurde, et que l'on privilégie au contraire les preuves constructives, dont les briques de base sont des résultats d'observations. En particulier, le principe du tiers-exclus n'est pas vérifié, ce qui est l'une des caractéristiques fondamentales de la logique intuitionniste. Si une propriété n'a pas été observée, elle n'est pas fausse pour autant. On retrouve ce fait dans le comportement des opérateurs modaux qui ne vérifient pas l'axiome 5.

Un autre point qui a été développé est celui des liens qui existent entre la connaissance accessible depuis les différents points d'observation. En effet, dans la structure des hiérarchies de représentations, ils sont formalisés par un ensemble de fonctions qui correspondent à un moyen de transformer les descriptions d'une représentation vers une autre. Cela est possible puisqu'à chaque fois, il s'agit du même système qui est représenté. À partir de cela, il est raisonnable d'espérer pouvoir relier de façon simple les connaissances liées aux diverses représentations. Or, une première étude de cette question montre que la logique **OL** ne permet d'exprimer que des relations très pauvres à ce sujet. Pour illustrer cela, nous pouvons mettre en parallèle les deux axiomes suivants :

$$\begin{array}{ll} K_i K_j \varphi \rightarrow K_j \varphi & \mathbf{T}_2 \\ K_i K_j \varphi \rightarrow K_i \varphi & \mathbf{C} \end{array}$$

Le premier stipule que si une personne A peut prouver qu'une autre personne B peut prouver une propriété, alors B peut effectivement prouver la propriété en question. Il s'agit d'un axiome vérifié par la logique **OL**. Cependant, la non-vérification de l'axiome **C** par **OL** précise que la connaissance de ce fait par A ne lui suffit pas pour prouver la propriété.

Ces dernières considérations ont conduit à une longue discussion où nous avons développé différentes approches pour identifier plus précisément ce que l'on peut exprimer sur les relations logiques entre observations. Deux notions sont ressorties de ces discussions. Tout d'abord, celle de localisation, où il est apparu que puisque toute connaissance provient d'observations, il est possible de leur associer une information précisant explicitement la représentation utilisée pour exprimer cette information. Mais nous avons vu de plus que les propositions atomiques qui apparaissent dans les propositions ont un sens qui dépend uniquement de la représentation considérée, ce qui explique que l'axiome **C** précédent n'est pas vérifié : l'énoncé d'une proposition peut dépendre d'une localisation précise, et ne peut donc pas être utilisé sans modification

autre part.

Nous avons donc défini le concept d’objectivité pour une propriété, qui précise si une propriété donnée dépend d’une représentation ou non, ou encore si elle peut être utilisée sans reformulation quelque en soit l’endroit. En d’autres termes, cette notion précise si une propriété vérifie l’axiome **C** ou non. Il s’agit donc d’un critère d’autonomie, qui permet de s’assurer qu’une proposition contient suffisamment d’information sur son contexte d’utilisation. Plus généralement, on peut parler alors de propositions exprimables, puisque ce sont les seules qui permettent d’affirmer quelque chose sur l’état du système en ne connaissant que la structure de la hiérarchie de représentation, ce sont les seules que l’on peut énoncer directement en restant extérieur au système, sans se référer à une représentation particulière.

De plus, un autre aspect important de cette notion est que dans la formulation que nous avons donnée de ce critère, il apparaît qu’il est équivalent à la vérification de l’axiome **T** : $K_i \varphi \rightarrow \varphi$. Ainsi, ce qui au départ a été défini comme un critère d’indépendance vis-à-vis des représentations revient à un critère permettant de dire si une propriété observée est réellement vérifiée par le système ou non, si une propriété est “réelle” ou non.

Comparaison avec d’autres travaux

Il peut être intéressant de mettre ce travail en parallèle avec d’autres approches de la modélisation de la connaissance. Nous citerons deux d’entre elles, qui toutes deux ont servi comme modèle et source d’inspiration. En premier lieu, on peut citer les recherches de Jon Barwise sur l’analyse de flux d’information [3]. Il convient de remarquer à nouveau que ces recherches ont de nombreux points communs avec notre travail. En particulier, sa structure se rapproche de celle de [3], puisque dans les deux cas, le point de départ est la représentation de connaissances d’un point de vue donnée, puis le formalisme est étendu pour tenir compte de plusieurs points de vue, et on utilise ce formalisme pour définir une logique permettant d’analyser le comportement des systèmes ainsi modélisés.

Cependant, les approches utilisées sont différentes. J. Barwise et J. Seligman se concentrent sur les relations existant entre les logiques “locales” (c’est-à-dire la façon de définir le comportement du système, depuis un point de vue donnée) et ni les différentes représentations ni l’aspect approximation n’apparaissent explicitement dans celles-ci. Au contraire, nous nous sommes attachés sur ces deux notions, et avons défini une logique modale pour avoir ces aspects à l’intérieur même de la logique.

On peut de plus mettre ce travail en relation avec différents travaux centrés sur le raisonnement à propos de la connaissance [39, 23, 38, 49]. Dans ce contexte, l’étude se base aussi sur l’utilisation de logiques modales, où les opérateurs de modalité correspondent à la base aux différents agents observant le système, comme c’est l’usage en logique épistémique. Cependant,

comme nous l'avons remarqué précédemment, nous nous sommes basé sur une approche plus restrictive de la connaissance, puisque nous sommes partis du postulat qu'elle provient uniquement d'observations. Cela a conduit comme nous l'avons vu à un abandon d'axiome d'introspection négative $\mathbf{5}$, lequel abandon semble définitif comme expliqué dans la section 8.2. Dans le même ordre d'idées, la logique que nous avons définie ne vérifie pas l'axiome \mathbf{T} spécifiant que les connaissances observées peuvent être appliquées directement au système. Cet axiome a été remplacé par trois axiomes, la paire \mathbf{L}_4 et \mathbf{L}_T d'une part qui jouent un rôle équivalent $\mathbf{4}$ et \mathbf{T} mais restreints à un unique agent, et \mathbf{T}_2 qui joue le rôle de \mathbf{T} pour les propositions de la forme $K_i \varphi$ uniquement.

Ainsi, en analysant précisément le cadre utilisé pour obtenir des informations sur le système, nous avons obtenu une caractérisation de la logique correspondante, en supprimant l'axiome $\mathbf{5}$ et en restreignant fortement le champ d'application de \mathbf{T} .

Ce travail a permis d'élaborer un formalisme précis pour appréhender les problèmes d'accessibilité de la connaissance. Il s'agit à notre connaissance de la première étude de ce sujet. Un de ses intérêts principaux, en plus de la facilité d'utilisation de ce formalisme, est le fait que partant de postulats simples, nous sommes arrivés à une caractérisation logique assez riche de ce que peut être la connaissance observable, reflétant bien en particulier les limites qui en découlent. Nous avons montré que tout ce que l'on observe ne correspond pas nécessairement à la réalité, et que dans la démarche scientifique, on ne peut pas prouver en général la causalité effective d'une relation de cause à effet que l'on observe.

Perspectives

Parmi les différentes pistes de développement qui peuvent venir à l'esprit, on peut tout d'abord citer diverses façons d'enrichir le formalisme et la logique que nous avons développés. On peut revenir pour cela à certaines directions évoquées plus tôt.

Tout d'abord, l'ajout le plus évident que l'on peut faire \mathbf{OL} est celui des quantificateurs universels et existentiels, permettant d'obtenir une logique du premier ordre. Comme mentionné précédemment, cela soulève la question de la définition du domaine de valeurs que pourraient prendre les variables propositionnelles, qui en particulier peut dépendre de la représentation utilisée ou non. De plus, de même que l'on peut quantifier les propositions logiques, on pourrait aussi envisager d'inclure la possibilité de quantifier les indices de représentations.

Une autre piste de développement de la logique \mathbf{OL} est l'ajout d'autres opérateurs épistémiques. Une source d'inspiration possible à ce sujet est donnée par les opérateurs \mathbf{E} (la connais-

sance partagée citée plus haut) et \mathbf{C} décrits dans [23]. Ils sont définis de la façon suivante :

$$\mathbf{E} \varphi \Leftrightarrow \forall i \in \mathcal{I}, K_i \varphi$$

$$\mathbf{C} \varphi \Leftrightarrow \mathbf{E}(\varphi \wedge \mathbf{C} \varphi)$$

L'opérateur \mathbf{C} en particulier a été évoqué dans la définition de la notion d'objectivité, puisque la contrainte que l'on a imposée aux contextes Γ représentant l'information structurelle revenant à imposer qu'ils devaient être équivalent à $\mathbf{C} \Gamma$.

En utilisant ces opérateurs et des quantifications sur les indices, on peut espérer caractériser la logique pour les hiérarchies de représentation dans le cas où l'ensemble des indices est fini. Ainsi, il pourrait être intéressant d'étudier des axiomes de la forme :

$$\bigvee_i \mathbf{C} \bigwedge_j \forall x. K_j x \rightarrow K_i K_j x$$

qui expriment le fait qu'une représentation est connue communément pour être plus précise que les autres. De même, une hiérarchie de représentation pourrait être décrite à l'intérieur de la logique par :

$$\forall i, j, \exists k : \mathbf{C} \forall \varphi. (K_i \varphi \rightarrow K_k K_i \varphi) \wedge (K_j \varphi \rightarrow K_k K_j \varphi)$$

Un autre développement possible est celui d'une formalisation en calcul des séquents adéquate pour obtenir un théorème d'élimination des coupures. Cela permettrait en particulier d'étudier plus précisément des notions comme l'objectivité, la localisation, etc.

Une autre suite que l'on peut donner à ces travaux est leur application à des problèmes concrets. Ainsi, nous pouvons utiliser ce formalisme dans les différents domaines évoqués dans le survol que nous avons fait plus tôt des disciplines utilisant les notions d'abstraction et d'approximation.

De plus, de par son caractère très général et théorique, on peut envisager d'utiliser les notions provenant des hiérarchies de représentations à certains problèmes de physique mettant en jeu les concepts d'information, d'échange et de mesure. On peut aussi envisager dans le même registre une étude générale des phénomènes d'échanges d'informations entre systèmes dissociés. C'est une problématique très riche, puisque l'on peut y inclure des systèmes distribués en informatique comme certains aspects des interactions entre systèmes physiques.

Enfin, un dernier point que l'on peut soulever est celui des probabilités et des mesures d'informations. La plupart des théories essayant de formaliser un moyen de quantifier l'information se basent en général sur une approche probabiliste des descriptions des états d'un système, entraînant souvent des problèmes délicats comme l'existence de probabilités *a priori*. Dans le cadre des hiérarchies de représentations, nous ne manipulons que des données sûres, issues directement d'observations. Une tentative de définition d'une mesure de la quantité d'informations en utilisant ces structures pourrait être une direction de recherche intéressante.

Bibliographie

- [1] Samson Abramsky. *Domain Theory and the Logic of Observable Properties*. PhD thesis, Queen Mary College, University of London, 1987.
- [2] Michael Barr. **-Autonomous categories*. Number 752 in Lecture Notes in Mathematics. Springer-Verlag, 1979.
- [3] Jon Barwise and Jerry Seligman. *Information Flow, the logic of distributed systems*. Number 44 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1997.
- [4] Gavin M. Bierman and Valeria de Paiva. Intuitionistic necessity revisited. Technical Report CSR-96-10, School of Computer Science, University of Birmingham, 1996.
- [5] Gavin M. Bierman and Valeria de Paiva. On an intuitionistic modal logic. *Studia Logica*, 65 :383–416, 2000.
- [6] Garrett Birkhoff. *Lattice Theory*. Colloquim Publications. American Mathematical Society, 3rd edition, 1967.
- [7] Luitzen E. J. Brouwer. *Over de Grondslagen der Wiskunde*. Thesis, Amsterdam, 1907.
- [8] Luitzen E. J. Brouwer. De onbetrouwbaarheid der logisches principes. *Tijdschrift voor wijsbegeerte*, 2 :152–158, 1908.
- [9] Luitzen E. J. Brouwer. *Collected Works I. (A. Heyting ed.)*. North-Holland, Amsterdam, 1975.
- [10] Olivier Brunet. Classifications et treillis. In *Actes 5e Rencontres nationales sur de jeunes chercheurs en intelligence artificielle (RJCIA'00)*, pages 29–38, 2000.
- [11] Olivier Brunet. Lattice approach to classifications. In *Proc. 12th European summer school on logic, language and information (ESSLLI'00)*, pages 34–44, Birmingham, UK, 2000.
- [12] Olivier Brunet. A model for knowledge representation in distributed systems. In *Proceedings of the KI2001 workshop Modal Logic in AI*, pages 3–12, Wien, OS, 2001.
- [13] Olivier Brunet. A modal logic for observation-based knowledge representation. In *Proceedings of the IMLA'02 workshop (Intuitionnistic Modal Logic and Applications)*, Copenhagen, DK, 2002.

- [14] Brian Chellas. *Modal Logic, an Introduction*. Cambridge University Press, 1980.
- [15] Patrick Cousot. Abstract interpretation based formal methods and future challenges, invited paper. In R. Wilhelm, editor, *Informatics — 10 Years Back, 10 Years Ahead*, volume 2000 of *Lecture Notes in Computer Science*, pages 138–156. Springer-Verlag, 2001.
- [16] Patrick Cousot and Radhia Cousot. Abstract interpretation : a unified lattice model for static analysis of program by construction of approximate fixpoints. In *Conference records of the 4th ACM Symposium on Principles Of Programming Languages (POPL'77)*, Los Angeles, California, U.S.A., 1977.
- [17] Patrick Cousot and Radhia Cousot. Systematic design of program analysis framework. In *Conference records of the Sixth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'79)*, 1979.
- [18] Roberto Cremonini, Kim Marriott, and Harald Søndergaard. A framework for abstraction basec on abstract interpretation. In *Working Notes of AAAI-90 Workshop on Automatic Generation of Approximations and Abstractions*, pages 233–244, 1990.
- [19] Brian A. Davey and Hilary A. Priestley. *Introduction to Lattices and Order*. Cambridge Mathematical Textbooks, 1990.
- [20] Rowan Davies and Frank Pfenning. A modal analysis of staged computation. In *Proceedings of Symposium on Principles of Programming Languages*, pages 258–270, 1996.
- [21] Jean-Marc Deshouillers, Herman te Riele, and Yannick Saouter. New experimental results concerning the goldbach conjecture. *Algorithmic Number Theory*, pages 204 – 215, 1998.
- [22] Fred Dretske. *Knowledge and the Flow of Information*. The MIT Press, 1981.
- [23] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. The MIT Press, 1995.
- [24] Matt Fairtlough and Michael Mendler. An intuitionnistic modal logic with applications to the formal verification of hardware. *Lectures Notes in Computer Science*, 933, 1995.
- [25] Gisèle Fischer Servi. Axiomatizations for some intuitionistic modal logics. *Rend. Sem. Mat. Univers. Politecn. Torino*, 42 :179–194, 1984.
- [26] Melvin Fitting. *Proof Methods for Modal and Intuitionnistic Logics*, volume 169. D. Reidel Publishing, 1983.
- [27] Jean Gallier. Constructive logics. part i : A tutorial on proof systems and typed λ -calculi. *Theoretical Computer Science*, 110(2) :249 – 339, 1993.
- [28] Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis*. Springer Verlag, 1999.
- [29] Gerhard Gentzen. Investigations into logical deduction. In M. E. Szabo, editor, *The collected papers of Gerhard Gentzen*, pages 68 – 128. North Holland, 1969.

- [30] Jean-Yves Girard. *Proof Theory and Logical Complexity*. Bibliopolis, 1987.
- [31] Fausto Giunchiglia, Adolfo Villafiorita, and Toby Walsh. Theories of abstraction. *AI Communications*, 10(3-4) :167–176, 1997.
- [32] Fausto Giunchiglia and Toby Walsh. A theory of abstraction. *Artificial Intelligence*, 56(2-3) :323–390, October 1992.
- [33] Kurt Gödel. Zum intuitionistischen aussagenkalkül. *Ergebnisse eines mathematischen Kolloquiums*, 4 :42, 1932.
- [34] Jean Goubault-Larrecq. On computational interpretations of the modal logic $S4$ i-iii. Technical Report 1996-33 :35, Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe, 1996.
- [35] Jean Goubault-Larrecq and Eric Goubault. Order-theoretic, geometric and combinatorial models of intuitionistic $S4$ proofs. In *Intuitionistic Modal Logics and Applications (IMLA'99)*, 1999.
- [36] Jean Goubault-Larrecq and Eric Goubault. On the geometry of intuitionistic $S4$ proofs. Technical Report LSV-01-8, ENS Cachan, Cachan, France, 2001.
- [37] Vineet Gupta. *Chu Spaces : A Model of Concurrency*. PhD thesis, Stanford University, September 1994.
- [38] Joseph Y. Halpern. Reasoning about knowledge : A survey. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4, pages 1 – 34. Clarendon Press, 1995.
- [39] Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3) :549–587, 1990.
- [40] Gary Hardegree. An approach to the logic of natural kinds. *Pacific Phil. Quarterly*, 63 :122–132, 1983.
- [41] Arend Heyting. *Intuitionism : An Introduction*. North-Holland, 1956.
- [42] Chris Hillman. What is a concept ?, 1998.
- [43] Jikko Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.
- [44] George E. Hughes and Maxwell J. Cresswell. *An Introduction to Modal Logic*. Methuen and Co LTD., 1968.
- [45] George E. Hughes and Maxwell J. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.
- [46] Craig A. Knoblock. Learning abstraction hierarchies for problem solving. In T. Dietterich and W. Swartout, editors, *Proceedings of the Eighth National Conference on Artificial Intelligence*, Menlo Park, California, 1990. AAAI Press / MIT Press.

- [47] Craig A. Knoblock. Search reduction in hierarchical problem solving. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, pages 686–691, Anaheim, California, USA, 1991. AAAI Press / MIT Press.
- [48] Yves Lafont and Thomas Streicher. Games semantics for linear logic. In *Proc. 6th Annual IEEE Symp. on Logic in Computer Science*, pages 43–49, July 1991.
- [49] Hector J. Levesque and Gerhard Lakemeyer. *The Logic of Knowledge Bases*. The MIT Press, 2000.
- [50] Clarence I. Lewis. *A Survey of Symbolic Logic*. University of California, 1918.
- [51] P. Pandurang Nayak and Alan Levy. A semantic theory of abstractions. In Chris Mellish, editor, *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 196–203. Morgan Kaufmann, 1995.
- [52] David A. Plaisted. Theorem proving with abstraction. *Artificial Intelligence*, 16 :47–108, 1981.
- [53] Vaughan Pratt. The duality of time and information. In *Proc. Conf. on Concurrency (CONCUR'92)*, volume 630 of *Lecture Notes in Computer Science*, pages 237–253. Springer-Verlag, 1992.
- [54] Vaughan Pratt. Types as processes, via chu spaces, 1997.
- [55] Vaughan Pratt. Chu spaces as a semantic bridge between linear logic and mathematics, 1998.
- [56] Earl D. Sacerdoti. Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5 :115–135, 1974.
- [57] Lorenza Saitta and Jean-Daniel Zucker. Semantic abstraction for concept representation and learning. In *Symposium on Abstraction, Reformulation and Approximation (SARA'98)*, 1998.
- [58] Alex Simpson. *The Proof Theory and Semantics of Intuitionistic Modal Logics*. PhD thesis, University of Edinburgh, 1993.
- [59] Mike Smyth. Powerdomains and predicate transformers : a topological view. In J. Diaz, editor, *Automata, Languages and Programming*, volume 154 of *Lecture Notes in Computer Science*, pages 662–675. Springer-Verlag, 1983.
- [60] Colin P. Stirling. Modal logics for communicating systems. *Theoretical Computer Science*, 49 :311–347, 1987.
- [61] William W. Tait. Normal derivability in classical logic. In J. Barwise, editor, *The Syntax and Semantics of Infinitary Languages*. Springer Verlag, 1968.
- [62] Dirk van Dalen. Intuitionistic logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume III, pages 225–340. Reidel, 1986.

Annexe A

Index des axiomes

- **4** : *Introspection positive* p. 16, 62, 67, 125
$$K_i \varphi \rightarrow K_i K_i \varphi$$
- **5** : *Introspection négative* p. 16, 67, 93
$$\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$$
- **C** : *Confiance* p. 96, 98, 136, 149
$$K_i K_j \varphi \rightarrow K_i \varphi$$
- **D** : *Consistance* p. 16, 66, 98, 135, 150
$$K_i \varphi \rightarrow \neg K_i \neg \varphi$$
- **GD** : *Consistance globale* p. 98, 125, 136
$$K_i \varphi \rightarrow \neg K_j \neg \varphi$$
- **K** : *Distribution* p. 16, 62, 65, 135, 150
$$K_i (\varphi \rightarrow \psi) \rightarrow K_i \varphi \rightarrow K_i \psi$$
- **KV** : p. 62, 65, 135, 151
$$K_i (\varphi \vee \psi) \rightarrow K_i \varphi \vee K_i \psi$$
- **L** : *Localisation* p. 68, 135
$$K_i (\varphi \leftrightarrow K_i \varphi)$$
- **L₄** : p. 68, 125
$$K_i (\varphi \rightarrow K_i \varphi)$$
- **L_T** : p. 68
$$K_i (K_i \varphi \rightarrow \varphi)$$
- **T** : *Connaissance* p. 16, 62, 66, 125, 136, 141, 144, 148, 149, 152
$$K_i \varphi \rightarrow \varphi$$
- **T₁** : p. 66
$$K_i K_i \varphi \rightarrow K_i \varphi$$
- **T₂** : p. 72, 96, 135, 150
$$K_i K_j \varphi \rightarrow K_j \varphi$$

Annexe B

Preuves de correction et de complétude

B.1 Représentations

B.1.1 Correction

Pour prouver le résultat de correction, on montre que si un ensemble fini de formules $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ implique une formule φ , alors l'interprétation de celle-ci contient toujours l'intersection des interprétations des formules de Γ .

Proposition 26

Pour tout séquent de la logique intuitionniste $\gamma_1, \dots, \gamma_n \vdash \varphi$, pour toute interprétation vers une représentation $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu}$, on a :

$$\bigcap_i \llbracket \gamma_i \rrbracket_{\mathcal{R}, \nu} \subseteq \llbracket \varphi \rrbracket_{\mathcal{R}, \nu}$$

Preuve On montre ce résultat par induction sur la taille de la preuve. Pour cela, on regarde la dernière règle utilisée. Chaque cas se traite directement. Regardons par exemple le cas $\rightarrow L$. En notant $\Gamma = \gamma_1, \dots, \gamma_n$ et $\llbracket \Gamma \rrbracket = \bigcap \llbracket \gamma_i \rrbracket$, supposons donc que la dernière règle de la preuve est de la forme :

$$\frac{\Gamma \vdash \varphi \quad \Gamma, \psi \vdash \vartheta}{\Gamma, \varphi \rightarrow \psi \vdash \vartheta}$$

Soit $d \in \llbracket \Gamma \rrbracket \cap \llbracket \varphi \rightarrow \psi \rrbracket$. On veut montrer que $d \in \llbracket \vartheta \rrbracket$. Mais $d \in \llbracket \varphi \rightarrow \psi \rrbracket$ implique que $\forall d' \leq d, d' \in \llbracket \varphi \rrbracket \rightarrow d' \in \llbracket \psi \rrbracket$. Si l'on regarde en particulier le cas $d' = d$, cela se réécrit $d \in \llbracket \varphi \rrbracket \Rightarrow d \in \llbracket \psi \rrbracket$. Or, par récurrence, on a le séquent valide $\Gamma \vdash \varphi$, et donc comme $d \in \llbracket \Gamma \rrbracket$, on en déduit $d \in \llbracket \varphi \rrbracket$, puis $d \in \llbracket \psi \rrbracket$. Maintenant, par récurrence à nouveau, mais en considérant $\Gamma, \psi \vdash \vartheta$, on a $d \in \llbracket \Gamma \rrbracket \cap \llbracket \psi \rrbracket$, et donc $d \in \llbracket \vartheta \rrbracket$ ce qui permet de conclure. Les autres cas se font de manière similaire. \square

Corollaire 26.1

Pour toute proposition $\varphi \in \mathcal{L}_\Psi$ et toute interprétation $\llbracket \cdot \rrbracket_{\mathcal{R}, \nu}$, on a :

$$\vdash_{\mathbf{IL}} \varphi \Rightarrow \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} = \mathcal{R}$$

Preuve On applique la proposition 26 pour le séquent $\emptyset \vdash \varphi$. □

B.1.2 Complétude

Pour montrer que les propositions de \mathcal{L}_Ψ sont prouvables pour la logique intuitionniste si et seulement si elles sont valides pour les représentations, on va utiliser la méthode classique du modèle canonique [26, 62]. Pour cela, on va définir une représentation canonique \mathcal{R}_c pour laquelle sont vraies les seules formules prouvables dans **IL**. Pour cela, nous allons introduire la notion d'ensemble premier :

Definition 5 (Ensemble premier)

Un ensemble premier (ou saturé, prime set en anglais), pour un langage dénombrable \mathcal{L} et une relation de déduction pour ce langage \vdash , est un sous-ensemble Γ de \mathcal{L} clos par déduction ($\forall \varphi, \Gamma \vdash \varphi \Rightarrow \varphi \in \Gamma$), consistant ($\Gamma \not\vdash \perp$) et tel que si $\Gamma \vdash \varphi \vee \psi$, alors soit $\Gamma \vdash \varphi$, soit $\Gamma \vdash \psi$.

Une propriété fondamentale des ensembles premiers est la suivante :

Proposition 27

Pour tout sous-ensemble consistant Γ de \mathcal{L} et pour toute proposition $\varphi \in \mathcal{L}$ telle que $\Gamma \not\vdash \varphi$, il existe un ensemble premier Δ tel que $\Gamma \subseteq \Delta$, et $\Delta \not\vdash \varphi$.

Preuve On donne ici la preuve décrite dans [62] p. 252. Comme \mathcal{L} est dénombrable, on en numérote les formules. On va construire une suite croissante d'ensembles $\{\Gamma_k\}$ de la façon suivante : $\Gamma_0 = \Gamma$ et par récurrence, si $\Gamma_0, \dots, \Gamma_k$ ont déjà été construits, on considère la disjonction $\psi_1 \vee \psi_2$ ayant le plus petit numéro parmi celles non encore traitées et telles que $\Gamma_k \vdash \psi_1 \vee \psi_2$. Il est impossible que l'on ait $\Gamma_k, \psi_i \vdash \varphi$ pour les deux valeurs de i , et on définit donc $\Gamma_{k+1} = \Gamma_k \cup \{\psi_i\}$ pour i tel que $\Gamma_k, \psi_i \not\vdash \varphi$.

On définit $\Gamma_\infty = \bigcup \Gamma_k$ et l'on va montrer qu'il s'agit d'un ensemble premier vérifiant les propriétés recherchées. On montre cela en considérant les trois points suivants :

- Tout d'abord, on a bien $\Gamma_\infty \not\vdash \varphi$, car dans le cas contraire il existerait un k tel que $\Gamma_k \vdash \varphi$, ce qui est impossible. En particulier, cela implique que Γ_∞ est consistant.
- Si $\Gamma_\infty \vdash \psi_1 \vee \psi_2$, alors il existe un k tel que $\forall k' \geq k, \Gamma_{k'} \vdash \psi_1 \vee \psi_2$. Aussi, il existe un indice k'' tel que la disjonction $\psi_1 \vee \psi_2$ est bien traitée à cette étape. On en déduit que soit $\psi_i \in \Gamma_{k''+1}$ pour l'une des valeurs de i .

- Pour tout ψ , on a $\Gamma_\infty \vdash \psi \Rightarrow \psi \in \Gamma_\infty$. En effet, $\Gamma_\infty \vdash \psi$ implique que $\Gamma_\infty \vdash \psi \vee \psi$, et le point précédent nous permet de conclure que $\psi \in \Gamma_\infty$.

Ainsi, Γ_∞ est bien un ensemble premier, contenant Γ et vérifiant $\Gamma_\infty \not\vdash \varphi$. \square

On définit la représentation canonique comme étant composée d'ensembles premiers :

Definition 6 (Représentation canonique)

On définit la représentation canonique \mathcal{R}_c comme étant l'ensemble des ensembles premiers de \mathcal{L}_Ψ pour \mathbf{IL} , ordonné par l'inclusion inverse (c'est-à-dire $\Gamma \leq \Delta \Leftrightarrow \Delta \subseteq \Gamma$).

En posant $\nu_c(\psi) = \{\Gamma \mid \psi \in \Gamma\}$, on définit l'interprétation canonique $\llbracket \cdot \rrbracket_c$ comme étant l'interprétation $\llbracket \cdot \rrbracket_{\mathcal{R}_c, \nu_c}$. La proposition centrale pour prouver la complétude de l'interprétation est la suivante :

Proposition 28

Pour toute formule φ et tout ensemble premier Γ , on a :

$$\Gamma \vdash_{\mathbf{IL}} \varphi \Leftrightarrow \Gamma \in \llbracket \varphi \rrbracket_c$$

Preuve La preuve de ce résultat se fait par récurrence sur la structure du terme considéré. Pour le cas de base, on a par définition $\llbracket \psi \rrbracket_c = \{\Gamma \mid \Gamma \vdash_{\mathbf{IL}} \psi\}$. Dans le cas d'une conjonction $\varphi \vee \psi$, on a par définition des ensembles premiers $\Gamma \vdash \varphi \vee \psi$ si et seulement si $\Gamma \vdash \varphi$ ou $\Gamma \vdash \psi$. On a donc par récurrence :

$$\Gamma \vdash \varphi \vee \psi \Leftrightarrow \Gamma \vdash \varphi \text{ ou } \Gamma \vdash \psi \Leftrightarrow \Gamma \in \llbracket \varphi \rrbracket_c \cup \llbracket \psi \rrbracket_c \Leftrightarrow \Gamma \in \llbracket \varphi \vee \psi \rrbracket_c$$

La conjonction se traite de façon similaire en remarquant que cette fois, $\llbracket \varphi \wedge \psi \rrbracket_c = \llbracket \varphi \rrbracket_c \cap \llbracket \psi \rrbracket_c$. Il reste à montrer le cas de l'implication. Si $\Gamma \vdash_{\mathbf{IL}} \varphi \rightarrow \psi$, alors pour tout $\Delta \leq \Gamma$, on a $\Delta \vdash \varphi \rightarrow \psi$. Mais de plus, si $\Delta \vdash \varphi$, on en déduit que $\Delta \vdash \psi$. En utilisant la récurrence, on a donc montré que :

$$\Gamma \vdash_{\mathbf{IL}} \varphi \rightarrow \psi \Rightarrow \forall \Delta \leq \Gamma, (\Delta \in \llbracket \varphi \rrbracket_c \Rightarrow \Delta \in \llbracket \psi \rrbracket_c) \Rightarrow \Gamma \in \llbracket \varphi \rightarrow \psi \rrbracket_c$$

Réciproquement, si $\Gamma \not\vdash_{\mathbf{IL}} \varphi \rightarrow \psi$, cela implique que $\Gamma, \varphi \not\vdash_{\mathbf{IL}} \psi$. Alors, par application de la proposition 27, il existe un ensemble premier Δ tel que $\Gamma \cup \{\varphi\} \subseteq \Delta$ et $\Delta \not\vdash_{\mathbf{IL}} \psi$. Soit, exprimé autrement, $\exists \Delta \leq \Gamma : \Delta \in \llbracket \varphi \rrbracket_c$ et $\Delta \notin \llbracket \psi \rrbracket_c$, donc $\Gamma \notin \llbracket \varphi \rightarrow \psi \rrbracket_c$. \square

Corollaire 28.1

Pour toute formule $\varphi \in \mathcal{L}_\Psi$, on a :

$$\vdash_{\mathbf{IL}} \varphi \Leftrightarrow \models_{\mathcal{R}} \varphi$$

Preuve Le résultat de correction (prop. 26) montre que $\vdash_{\mathbf{IL}} \varphi$ implique $\forall \mathcal{R}, \nu, \llbracket \varphi \rrbracket_{\mathcal{R}, \nu} = \mathcal{R}$ soit par définition $\models_{\mathcal{R}} \varphi$. Réciproquement, en utilisant le modèle canonique \mathcal{R}_c de la proposition 28, on a :

$$\models_{\mathcal{R}} \varphi \Rightarrow \llbracket \varphi \rrbracket_c = \mathcal{R}_c \Rightarrow \vdash_{\mathbf{IL}} \varphi$$

On a donc bien l'équivalence voulue. □

B.2 Approximations de représentations

On retrouve ici la preuve d'un résultat donné dans [12].

B.2.1 Correction

Pour prouver le résultat de correction, on procède comme pour le cas avec les représentations seules, en ajoutant les règles correspondant aux opérateurs modaux.

Proposition 29

Pour tout séquent $\gamma_1, \dots, \gamma_n \vdash \varphi$ prouvable par $\mathbf{IS4} + \mathbf{KV}$, pour toute fonction d'interprétation $\llbracket \cdot \rrbracket$ vers une représentation \mathcal{R} munie d'un opérateur de clôture supérieure ρ , on a :

$$\bigcap_{i \in [1..n]} \llbracket \Gamma_i \rrbracket \subseteq \llbracket \varphi \rrbracket$$

Preuve On vérifie que les axiomes \mathbf{K} , \mathbf{T} , $\mathbf{4}$ et \mathbf{KV} vérifient cette inclusion. Pour l'axiome \mathbf{K} , on a :

$$\llbracket \Box(\varphi \rightarrow \psi) \rrbracket \cap \llbracket \Box\varphi \rrbracket \subseteq \llbracket \Box\psi \rrbracket$$

En effet, $d \in \llbracket \Box(\varphi \rightarrow \psi) \rrbracket$ implique que $\forall d' \leq \rho(d), d' \in \llbracket \varphi \rrbracket \Rightarrow d' \in \llbracket \psi \rrbracket$. Mais si $d \in \llbracket \Box\varphi \rrbracket$, on en déduit que $\rho(d) \in \llbracket \varphi \rrbracket$, d'où $\rho(d) \in \llbracket \psi \rrbracket$. Pour l'axiome \mathbf{T} , on a :

$$\llbracket \Box\varphi \rrbracket \subseteq \llbracket \varphi \rrbracket$$

Cela vient de l'extensivité de $\rho : d \in \llbracket \Box\varphi \rrbracket \Rightarrow \rho(d) \in \llbracket \varphi \rrbracket \Rightarrow d \in \llbracket \varphi \rrbracket$. L'axiome $\mathbf{4}$ vient lui de l'idempotence, qui implique que $\llbracket \Box\Box\varphi \rrbracket = \llbracket \Box\varphi \rrbracket$. Il reste l'axiome \mathbf{KV} qui se traite en remarquant que $\llbracket \Box(\varphi \vee \psi) \rrbracket = \llbracket \Box\varphi \rrbracket \cup \llbracket \Box\psi \rrbracket = \llbracket \Box\varphi \vee \Box\psi \rrbracket$. □

B.2.2 Complétude

On définit pour prouver la complétude le modèle canonique suivante : la représentation $\mathcal{R}_{\mathcal{AR}}$ est l'ensemble des ensembles premiers de $\mathcal{L}_{\Box, \Psi}$ pour **IS4** + **KV**, ordonné par l'inclusion inverse. L'opérateur de clôture $\rho_{\mathcal{AR}}$ est défini par :

$$\rho_{\mathcal{AR}}(\Gamma) = \{\varphi \mid \Box\varphi \in \Gamma\}$$

L'axiome **K** assure la clôture par déduction des $\rho_{\mathcal{AR}}(\Gamma)$, **KV** que $\varphi \vee \psi \in \rho_{\mathcal{AR}}(\Gamma)$ implique que φ ou bien ψ appartient à $\rho_{\mathcal{AR}}(\Gamma)$, **T** que les $\rho_{\mathcal{AR}}(\Gamma)$ sont consistents, ainsi que $\rho_{\mathcal{AR}}(\Gamma) \subseteq \Gamma$. **4** assure lui que $\rho_{\mathcal{AR}}$ est bien idempotent.

Il manque au modèle canonique la fonction d'interprétation atomique $\nu_{\mathcal{AR}}$ qui associe à une proposition atomique $\psi \in \Psi$ l'ensemble $\{\Gamma \mid \psi \in \Gamma\}$. En notant $\llbracket \cdot \rrbracket_{\mathcal{AR}}$ l'interprétation associée à ce modèle canonique, on peut maintenant interpréter la proposition principale de la preuve de complétude.

Proposition 30

Pour toute formule $\varphi \in \mathcal{L}_{\Box, \Psi}$ et pour tout ensemble premier $\Gamma \in \mathcal{R}_{\mathcal{AR}}$, on a l'équivalence :

$$\Gamma \vdash_{\mathbf{IS4+KV}} \varphi \Leftrightarrow \Gamma \in \llbracket \varphi \rrbracket_{\mathcal{AR}}$$

Preuve En s'appuyant par la preuve donnée pour les représentations et la logique intuitionniste, il suffit pour prouver ce résultat de s'intéresser au cas $\Box\varphi$. Par récurrence, on a :

$$\Gamma \in \llbracket \Box\varphi \rrbracket_{\mathcal{AR}} \Leftrightarrow \rho_{\mathcal{AR}}(\Gamma) \in \llbracket \varphi \rrbracket_{\mathcal{AR}} \Leftrightarrow \rho_{\mathcal{AR}}(\Gamma) \vdash \varphi$$

Il suffit donc de prouver qu'on a l'équivalence $\rho_{\mathcal{AR}}(\Gamma) \vdash \varphi \Leftrightarrow \Gamma \vdash \Box\varphi$. Mais si $\rho_{\mathcal{AR}}(\Gamma) \vdash \varphi$, l'axiome **K** nous permet de déduire $\{\Box\psi \mid \Box\psi \in \Gamma\} \vdash \Box\varphi$, et donc $\Gamma \vdash \Box\varphi$. Réciproquement, la primarité de Γ implique que si $\Gamma \vdash \Box\varphi$, alors $\Box\varphi \in \Gamma$, et donc $\varphi \in \rho_{\mathcal{AR}}(\Gamma)$. \square

Le théorème suit :

Théorème 12

Pour toute formule φ de $\mathcal{L}_{\Box, \Psi}$, on a :

$$\vdash_{\mathbf{IS4+KV}} \varphi \Leftrightarrow \models_{\mathcal{AR}} \varphi$$

Preuve En effet, si $\vdash_{\mathbf{IS4+KV}} \varphi$, alors on peut prouver le séquent $\emptyset \vdash \varphi$, et donc pour toute interprétation $\llbracket \cdot \rrbracket$ vers une représentation \mathcal{R} , on en déduit que $\llbracket \varphi \rrbracket = \mathcal{R}$ grâce à la proposition 29. Réciproquement, $\models_{\mathcal{AR}} \varphi$ implique en particulier que $\llbracket \varphi \rrbracket_{\mathcal{AR}} = \mathcal{R}_{\mathcal{AR}}$, et donc que $\vdash_{\mathbf{IS4+KV}} \varphi$ par complétude (Prop. 30). \square

B.3 Systèmes sans composition

B.3.1 Correction

Pour prouver le résultat de correction, comme pour les approximations de représentations, on procède en étudiant les règles correspondant aux opérateurs modaux.

Proposition 31

Pour tout séquent $\gamma_1, \dots, \gamma_n \vdash \varphi$ prouvable par OL^- , c'est-à-dire $\text{IKDL} + \text{KV}$, pour toute fonction d'interprétation $\llbracket \cdot \rrbracket$ vers un système de représentations $\mathcal{S} = \langle \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$, on a :

$$\forall j, \bigcap_{i \in [1..n]} \llbracket \Gamma_i \rrbracket_j \subseteq \llbracket \varphi \rrbracket_j$$

Preuve Chacune des fonctions d'interprétation $\llbracket \cdot \rrbracket_i$ se comportant comme dans le cas des représentations pour les connecteurs non modaux, il suffit de vérifier que les axiomes **K**, **D**, **L** et **KV** sont vérifiés eux aussi.

Pour l'axiome **K** : $K_i(\varphi \rightarrow \psi) \rightarrow K_i\varphi \rightarrow K_i\psi$, on prouve que :

$$\forall j, \llbracket K_i(\varphi \rightarrow \psi) \rrbracket_j \cap \llbracket K_i\varphi \rrbracket_j \subseteq \llbracket K_i\psi \rrbracket_j$$

Ainsi, soit $d \in \llbracket K_i(\varphi \rightarrow \psi) \rrbracket_j \cap \llbracket K_i\varphi \rrbracket_j$. Cela signifie que $f_{i|j}(d) \in \llbracket \varphi \rightarrow \psi \rrbracket_i \cap \llbracket \varphi \rrbracket_i$, ce qui donne le résultat voulu. Pour l'axiome **D**, dont on utilise la formulation $\neg K_i \perp$ (qui est équivalente à la formulation usuelle $K_i\varphi \rightarrow \neg K_i \neg\varphi$ en présence de l'axiome **K**), il suffit de remarquer que $\llbracket K_i \perp \rrbracket_j = \{d \mid f_{i|j}(d) \in \llbracket \perp \rrbracket_i\} = \emptyset$. L'égalité $\llbracket K_i\varphi \rrbracket_i = \{d \mid f_{i|i}(d) \in \llbracket \varphi \rrbracket_i\} = \llbracket \varphi \rrbracket_i$ entraîne $\llbracket \varphi \leftrightarrow K_i\varphi \rrbracket_i$, ce qui prouve la validité de l'axiome **L** : $K_i(\varphi \leftrightarrow K_i\varphi)$. Enfin, l'axiome **KV** : $K_i(\varphi \vee \psi) \rightarrow K_i\varphi \vee K_i\psi$ vient du fait que :

$$f_{i|j}(d) \in \llbracket \varphi \vee \psi \rrbracket_i \Leftrightarrow f_{i|j}(d) \in \llbracket \varphi \rrbracket_i \text{ ou } f_{i|j}(d) \in \llbracket \psi \rrbracket_i$$

□

B.3.2 Complétude

Le résultat se prouve comme précédemment en définissant un modèle canonique, avec quelques adaptations pour tenir compte de la modalité (voir [14, 25]).

Définition 9 (Modèle canonique)

On définit le modèle canonique comme $\mathcal{C}_{\mathcal{S}} = \langle \mathcal{I}, \{\mathcal{R}_i\}, \{f_{i|j}\}, \nu \rangle$ où les représentations \mathcal{R}_i sont définies comme :

$$\forall i \in \mathcal{I}, \mathcal{R}_i = \left\{ \Gamma_i \left| \begin{array}{l} \Gamma_i \text{ premier pour } \mathcal{L}_{\mathcal{I}, \Psi} \text{ selon } \vdash_{\text{OL}^-} \\ \forall \varphi, \varphi \leftrightarrow K_i\varphi \in \Gamma_i \end{array} \right. \right\}$$

ordonnées par l'inverse de l'inclusion ($\Gamma \leq_i \Delta \Leftrightarrow \Delta \subseteq \Gamma$) et les fonctions de transformation par $f_{i|j}(\Gamma_j) = \Gamma_{j/i} = \{\varphi \mid K_i \varphi \in \Gamma_j\}$. Enfin, les fonctions d'interprétation atomique sont données par $\nu_i(\psi) = \{\Gamma_i \mid \psi \in \Gamma_i\}$.

Le modèle canonique tel qu'il est défini est bien un système de représentations. En particulier, si $\Gamma_i \in \mathcal{R}_i$, on peut vérifier que $f_{j|i}(\Gamma_i) \in \mathcal{R}_j$. Pour cela, comme dans le cas des approximations d'une représentation par une autre, l'axiome **K** assure que l'ensemble $f_{j|i}(\Gamma_i)$ est clos pour la déduction, **D** qu'il est consistant, **KV** qu'il décide la disjonction. Enfin, par définition de la logique \mathbf{OL}^- , pour toute proposition φ , on a $K_j(\varphi \leftrightarrow K_j \varphi) \in \Gamma_i$, et donc $\varphi \leftrightarrow K_j \varphi \in f_{j|i}(\Gamma_i)$.

On peut donc définir une interprétation canonique $\llbracket \cdot \rrbracket_{\mathcal{C}_S}$ à partir du modèle canonique, et on montre :

Proposition 32

Pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$ et pour tout ensemble premier $\Gamma_i \in \mathcal{R}_i$ avec $i \in \mathcal{I}$, on a l'équivalence :

$$\Gamma_i \vdash_{\mathbf{OL}^-} \varphi \Leftrightarrow \Gamma_i \in \llbracket \varphi \rrbracket_{\mathcal{C}_S, i}$$

Preuve À nouveau, on prouve ce résultat par induction sur la taille du terme φ . Les connecteurs non modaux se traitent comme dans le cas des représentations (Prop. 28 page 169).

Il reste à étudier le cas $K_i \varphi$. Par définition, on a $\Gamma_j \in \llbracket K_i \varphi \rrbracket_{\mathcal{C}_S, j} \Leftrightarrow \Gamma_{j/i} \in \llbracket \varphi \rrbracket_{\mathcal{C}_S, i}$. Par induction, c'est équivalent à $\Gamma_{j/i} \vdash \varphi$. Il suffit donc de montrer $\Gamma_{j/i} \vdash \varphi \Leftrightarrow \Gamma_j \vdash K_i \varphi$ pour avoir le résultat voulu. Mais le sens \Rightarrow découle de l'axiome **K**, tandis que le sens \Leftarrow vient du fait que :

$$\Gamma_j \vdash K_i \varphi \Rightarrow K_i \varphi \in \Gamma_j \Rightarrow \varphi \in \Gamma_{j/i}$$

□

Avant de passer au théorème principal, nous proposons deux lemmes qui servent à montrer que si une proposition φ est validée par la représentation \mathcal{R}_i du modèle canonique, alors on peut prouver $K_i \varphi$ dans \mathbf{OL}^- .

Tout d'abord, nous allons dans le premier lemme définir un élément particulier des représentations canoniques \mathcal{R}_i . Pour cela, étant donné une proposition $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$, notons $\text{sf}(\varphi)$ l'ensemble des sous-formules de φ et $\text{ssf}(\varphi)$ l'ensemble des parties finies de $\mathcal{L}_{\mathcal{I},\Psi}$ contenant $\text{sf}(\varphi)$. On définit alors l'ensemble $\Gamma_{i,\top}$ comme :

$$\Gamma_{i,\top} = \left\{ \varphi \mid \exists \Phi \in \text{ssf}(\varphi) : \vdash \bigwedge_{\varphi' \in \Phi} (\varphi' \leftrightarrow K_i \varphi') \rightarrow \varphi \right\}$$

Lemme 1

Pour tout indice $i \in \mathcal{I}$ et toute proposition $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$, on a :

$$\varphi \in \Gamma_{i,\perp} \Leftrightarrow \vdash_{\text{OL-}} K_i \varphi$$

Preuve Pour le sens \Leftarrow , si $\vdash K_i \varphi$, alors comme $\varphi \in \text{sf}(K_i \varphi)$, on a dans la conjonction d'ensemble Φ de la définition le terme $\varphi \leftrightarrow K_i \varphi$ qui permet de conclure.

Réciproquement, si $\varphi \in \Gamma_{i,\perp}$, soit $\Phi \in \text{ssf}(\varphi)$ tel que $\vdash \bigwedge_{\varphi' \in \Phi} (\varphi' \leftrightarrow K_i \varphi') \rightarrow \varphi$. En appliquant la règle de nécessité et l'axiome **K**, on obtient :

$$\vdash \bigwedge_{\varphi' \in \Phi} K_i (\varphi' \leftrightarrow K_i \varphi') \rightarrow K_i \varphi$$

On conclut en remarquant que la conjonction ne contient plus que des instances de l'axiome **L**.
□

Lemme 2

Pour tout indice $i \in \mathcal{I}$, l'ensemble $\Gamma_{i,\top}$ appartient à \mathcal{R}_i .

Preuve On se sert du lemme précédent pour montrer ce résultat. En particulier, les axiomes **K** et **D** assurent la clôture par déduction et la consistance de $\Gamma_{i,\perp}$. De même, **KV** assure la décision de la disjonction, puisque :

$$\begin{aligned} \varphi_1 \vee \varphi_2 \in \Gamma_{i,\top} &\Rightarrow \vdash K_i (\varphi_1 \vee \varphi_2) \\ &\Rightarrow \vdash K_i \varphi_1 \vee K_i \varphi_2 \\ &\Rightarrow \vdash K_i \varphi_1 \text{ ou } \vdash K_i \varphi_2 \\ &\Rightarrow \varphi_1 \in \Gamma_{i,\top} \text{ ou } \varphi_2 \in \Gamma_{i,\top} \end{aligned}$$

Enfin, on a bien : $\forall \varphi, \varphi \leftrightarrow K_i \varphi \in \Gamma_{i,\top}$ du fait de l'axiome **L**. □

Lemme 3

On a pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I},\Psi}$ et tout indice $i \in \mathcal{I}$:

$$\llbracket \varphi \rrbracket_i = \mathcal{R}_i \Rightarrow \vdash_{\text{OL-}} K_i \varphi$$

Preuve En utilisant le lemme précédent, on a :

$$\llbracket \varphi \rrbracket_i = \mathcal{R}_i \Rightarrow \varphi \in \Gamma_{i,\perp} \Rightarrow \vdash_{\text{OL-}} K_i \varphi$$

□

On peut maintenant prouver le résultat principal :

Théorème 13

Pour toute formule φ de $\mathcal{L}_{\mathcal{I},\Psi}$, on a :

$$\vdash_{\mathbf{OL}^-} \varphi \quad \Leftrightarrow \quad \models_{\mathcal{S}} \varphi$$

Preuve En effet, si $\vdash_{\mathbf{OL}^-} \varphi$, alors on peut prouver le séquent $\emptyset \vdash \varphi$, et donc pour toute interprétation $\llbracket \cdot \rrbracket$ vers un système de représentations $\mathcal{S} = \langle \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$, on en déduit que $\forall i, \llbracket \varphi \rrbracket_i = \mathcal{R}_i$ grâce à la proposition 31. Réciproquement, $\models_{\mathcal{S}} \varphi$ implique en particulier que $\forall i, \llbracket \varphi \rrbracket_{c_{\mathcal{S},i}} = \mathcal{R}_i$, et donc que $\forall i, \vdash_{\mathbf{OL}^-} K_i \varphi$ par complétude (Prop. 32). On conclut en utilisant la règle **Univ** pour obtenir $\vdash_{\mathbf{OL}^-} \varphi$. \square

B.4 Systèmes compositionnels**B.4.1 Correction**

Pour la correction, il suffit de s'intéresser à l'axiome **T₂** qui est le seul à avoir été ajouté à \mathbf{OL}^- pour obtenir **OL**.

Proposition 33

Pour tout séquent $\gamma_1, \dots, \gamma_n \vdash \varphi$ prouvable par **OL** pour toute fonction d'interprétation $\llbracket \cdot \rrbracket$ vers un système compositionnel $\mathcal{S} = \langle \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$, on a :

$$\forall j, \bigcap_{i \in [1..n]} \llbracket \Gamma_i \rrbracket_j \subseteq \llbracket \varphi \rrbracket_j$$

Preuve On montre que l'axiome **T₂** : $K_i K_j \varphi \rightarrow K_j \varphi$ est bien vérifié par les systèmes compositionnels :

$$\begin{aligned} d \in \llbracket K_i K_j \varphi \rrbracket_k &\Rightarrow f_{i|k}(d) \in \llbracket K_j \varphi \rrbracket_i \\ &\Rightarrow f_{j|i} \circ f_{i|k}(d) \in \llbracket \varphi \rrbracket_j \\ &\Rightarrow f_{j|k}(d) \in \llbracket \varphi \rrbracket_j \\ &\Rightarrow d \in \llbracket K_j \varphi \rrbracket_k \end{aligned}$$

On a donc le résultat espéré. \square

B.4.2 Complétude

Le résultat de complétude pour les hiérarchies compositionnelles vis-à-vis de **OL** se prouve de la même façon que pour les hiérarchies sans composition vis-à-vis de **OL**⁻. Seulement on utilise pour le modèle canonique \mathcal{C}_{CS} les ensembles premiers selon **OL** :

Définition 10 (Modèle canonique)

On définit le modèle canonique comme $\mathcal{C}_S = \langle \mathcal{I}, \{\mathcal{R}_i\}, \{f_{i|j}\}, \nu \rangle$ où les représentations \mathcal{R}_i sont définies comme :

$$\forall i \in \mathcal{I}, \mathcal{R}_i = \left\{ \Gamma_i \mid \begin{array}{l} \Gamma_i \text{ premier pour } \mathcal{L}_{\mathcal{I}, \Psi} \text{ selon } \vdash_{\mathbf{OL}} \\ \forall \varphi, \varphi \leftrightarrow K_i \varphi \in \Gamma_i \end{array} \right\}$$

ordonnées par l'inverse de l'inclusion ($\Gamma \leq_i \Delta \Leftrightarrow \Delta \subseteq \Gamma$) et les fonctions de transformation par $f_{i|j}(\Gamma_j) = \Gamma_j/i = \{\varphi \mid K_i \varphi \in \Gamma_j\}$. Enfin, les fonctions d'interprétation atomique sont données par $\nu_i(\psi) = \{\Gamma_i \mid \psi \in \Gamma_i\}$.

La similitude avec la définition de \mathcal{C}_S nous assure que \mathcal{C}_{CS} est bien un système de représentations. Il reste à prouver que la propriété de composition est vérifiée. Il suffit d'écrire :

$$\begin{aligned} \varphi \in f_{i|j} \circ f_{j|k}(\Gamma_k) &\Rightarrow K_i \varphi \in f_{j|k}(\Gamma_k) \\ &\Rightarrow K_j K_i \varphi \in \Gamma_k \\ &\Rightarrow K_i \varphi \in \Gamma_k \\ &\Rightarrow \varphi \in f_{i|k}(\Gamma_k) \end{aligned}$$

On peut donc définir une interprétation canonique $\llbracket \cdot \rrbracket_{\mathcal{C}_{CS}}$ à partir du modèle canonique, et on montre :

Proposition 34

Pour toute formule $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$ et pour tout ensemble premier $\Gamma_i \in \mathcal{R}_i$, on a l'équivalence :

$$\Gamma_i \vdash_{\mathbf{OL}} \varphi \Leftrightarrow \Gamma_i \in \llbracket \varphi \rrbracket_{\mathcal{C}_{CS}, i}$$

Preuve La preuve est identique à celle de la proposition équivalente pour les systèmes non compositionnels (Prop. 32). □

On peut mettre ensemble ces deux résultats pour prouver :

Théorème 14

Pour toute formule φ de $\mathcal{L}_{\mathcal{I}, \Psi}$, on a :

$$\vdash_{\mathbf{OL}} \varphi \Leftrightarrow \models_{\mathcal{C}_S} \varphi$$

Preuve Comme précédemment, ce résultat est la mise en commun de la proposition de correction (Prop. 33) et de complétude (Prop. 34). \square

B.5 Hiérarchies consistantes

Nous allons maintenant nous pencher sur le résultat de complétude des hiérarchies consistantes vis-à-vis de la logique $\text{IS4} + \text{KV}$. Pour cela, nous allons tout d'abord nous intéresser aux hiérarchies consistantes ayant une représentation minimale pour l'ordre \triangleleft . Notons comme précédemment \perp l'indice correspondant. Cette contrainte permet de simplifier la situation. En effet, on note que d'après les définitions de la fonction Loc et l'interprétation $\{\{\cdot\}\}$, on a :

$$\begin{aligned} \forall \varphi, \perp \in \text{Loc}(\varphi) \\ \forall \varphi, (\forall i, \{\{\varphi\}\}_i = \mathcal{R}_i) \Leftrightarrow \{\{\varphi\}\}_\perp = \mathcal{R}_\perp \end{aligned}$$

Ainsi, on n'a besoin de calculer les $\{\{\varphi\}\}_i$ que dans le cas où $i = \perp$. Mais alors, l'interprétation de $\{\{K_i \varphi\}\}_\perp$ devient égale à $\{d \mid f_{\perp|i} \circ f_{i|\perp}(d) \in \{\{\varphi\}\}_\perp\}$. Les fonctions de la forme $f_{\perp|i} \circ f_{i|\perp}$ étant par définition de l'ordre \triangleleft des opérateurs de clôture inférieure, on se retrouve dans un cas analogue à celui apparu pour l'étude de l'approximation d'une représentation par une autre. Le cas où existent plusieurs opérateurs d'approximation est étudié dans [12], où nous avons montré que la logique modélisée est alors $\text{IS4} + \text{KV}$. Dans le cas présent cependant, il faut de plus exprimer qu'à la représentation la plus précise correspond un opérateur de modalité K_\perp . Cela se traduit par l'ajout d'un axiome $\text{GD}'_\perp : \varphi \rightarrow K_\perp \varphi$. L'ajout de cet axiome est suffisant pour obtenir la logique modélisée par les hiérarchies consistantes vérifiant $\forall i, \perp \triangleleft i$, puisque dans le modèle canonique, ρ_\perp est l'identité.

Nous avons donc montré le résultat suivant :

$$\forall \varphi, \models_{\mathcal{CH}, \perp} \varphi \Leftrightarrow \vdash_{\text{IS4} + \text{KV} + \text{GD}'_\perp} \varphi$$

Ce résultat s'applique en particulier au cas où l'ensemble \mathcal{I} est fini, puisqu'il y a alors forcément une représentation plus précise que toutes les autres.

Considérons maintenant le cas général où l'ensemble \mathcal{I} est infini, et soit $\varphi \in \mathcal{L}_{\mathcal{I}, \Psi}$. On note comme précédemment $\text{ind}(\varphi)$ l'ensemble des indices apparaissant dans le terme φ , et comme \mathcal{I} est infini et $\text{ind}(\varphi)$ fini, on peut définir un indice $j \in \mathcal{I} \setminus \text{ind}(\varphi)$. Notons \mathcal{J} l'ensemble $\text{ind}(\varphi) \cup \{j\}$.

Si φ est modélisée par les hiérarchies de représentations (ce que l'on note $\models_{\mathcal{CH}} \varphi$ ou $\models_{\mathcal{CH}, \mathcal{I}} \varphi$ en précisant l'ensemble d'indices), alors en particulier $\models_{\mathcal{CH}, \mathcal{J}, j} \varphi$ (où l'on se restreint aux

hiérarchies sur \mathcal{J} ayant j minimal pour \triangleleft). En effet, une telle hiérarchie de représentations peut être étendue pour être définie sur \mathcal{I} , en ajoutant pour $i \in \mathcal{I} \setminus \mathcal{J}$ la représentation vide à un seul élément. On en déduit alors que :

$$\vdash_{\mathbf{IS4+KV+GD}'_j} \varphi$$

On conclut en remarquant que toute preuve de $\mathbf{IS4} + \mathbf{KV} + \mathbf{GD}'_j$ peut être transformée en une preuve de $\mathbf{IS4} + \mathbf{KV}$ en effaçant toutes les occurrences de K_j . En particulier, tous les axiomes concernant les modalités de $\mathbf{IS4} + \mathbf{KV}$ restent prouvables dans cette logique après “nettoyage”, et l’axiome \mathbf{GD}'_j devient la tautologie $\varphi \rightarrow \varphi$. On a donc bien le résultat voulu :

$$\models_{c\mathcal{H}} \varphi \Rightarrow \vdash_{\mathbf{IS4+KV}} \varphi$$

Annexe C

Preuve d'élimination de coupures

Dans cette annexe, nous allons développer la preuve d'élimination de coupures pour le calcul de séquents présenté dans la section 7.1 p.75 et donc nous rappelons les règles dans la figure 10 p.180. La preuve utilisée est une adaptation de la preuve de Gentzen [29] modifiée par Tait [61] et Girard [30], telle que présentée dans [27].

C.1 La réduction de localisation

Avant de nous pencher sur le théorème de réduction de coupures, nous allons tout d'abord nous pencher sur la proposition 8 p.82 dont on rappelle l'énoncé :

Étant donné une preuve Π du séquent $\Gamma \vdash [\varphi]_{\Lambda}$, il est possible de la transformer en une preuve Π' de $\Gamma \vdash [\varphi]_{\Lambda'}$ où $\Lambda' \leq_{\star} \Lambda$. De plus, la structure de Π' ne diffère par rapport à Π que par l'ajout de règles Loc.

Tout d'abord, la définition de l'ordre \leq sur les mots de \mathcal{I}^{\star} est équivalente à l'existence d'une fonction croissante entre les lettres des mots. Plus précisément, deux mots Λ et Λ' vérifient $\Lambda \leq \Lambda'$ si et seulement si il existe une fonction croissante $\text{pos} : [1, |\Lambda|] \rightarrow [1, |\Lambda'|]$ (où $|\Lambda|$ désigne la longueur du mot Λ) telle que :

$$\forall i \in [1, |\Lambda|], \Lambda(i) = \Lambda'(\text{pos}(i))$$

On parle alors de fonction d'inclusion de Λ dans Λ' . De plus, on remarque qu'il est possible de définir un ordre partiel sur les fonctions vérifiant la propriété précédente en comparant les différents résultats (soit $p_1 \leq p_2 \Leftrightarrow \forall i \in [1, |\Lambda|], p_1(i) \leq p_2(i)$), et qu'il existe une fonction

$$\begin{array}{c}
\frac{\Lambda' \leq_* \Lambda}{[\varphi]_\Lambda \vdash [\varphi]_{\Lambda'}} \text{Axiom} \quad \frac{\Lambda' \leq \Lambda}{[\perp]_\Lambda \vdash [\varphi]_{\Lambda'}} \perp \\
\\
\frac{\Gamma \vdash [\psi]_\Lambda}{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda} \text{Weak} \quad \frac{\Gamma, [\varphi]_\Lambda, [\varphi]_\Lambda \vdash \psi}{\Gamma, [\varphi]_\Lambda \vdash \psi} \text{Contract} \quad \frac{\forall i, \Gamma_{oi} \vdash [\varphi]_{\Lambda \circ i}}{\Gamma \vdash [\varphi]_\Lambda} \text{Univ} \\
\\
\frac{\Gamma \vdash [\varphi]_\Lambda \quad \Delta, [\varphi]_{\Lambda'} \vdash \psi \quad \Lambda' \leq_* \Lambda}{\Gamma, \Delta \vdash \psi} \text{Cut} \quad \frac{\Gamma, [\varphi]_{\Lambda'} \vdash \psi \quad \Lambda' \leq_* \Lambda}{\Gamma, [\varphi]_\Lambda \vdash \psi} \text{Loc} \\
\\
\frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta \quad \Gamma, [\psi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \vee \psi]_\Lambda \vdash \vartheta} \vee L \quad \frac{\Gamma \vdash [\varphi]_\Lambda}{\Gamma \vdash [\varphi \vee \psi]_\Lambda} \vee R_1 \quad \frac{\Gamma \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \vee \psi]_\Lambda} \vee R_2 \\
\\
\frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \wedge \psi]_\Lambda \vdash \vartheta} \wedge L_1 \quad \frac{\Gamma, [\varphi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \wedge \psi]_\Lambda \vdash \vartheta} \wedge L_2 \quad \frac{\Gamma \vdash [\varphi]_\Lambda \quad \Gamma \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \wedge \psi]_\Lambda} \wedge R \\
\\
\frac{\Gamma \vdash [\varphi]_\Lambda \quad \Gamma, [\psi]_\Lambda \vdash \vartheta}{\Gamma, [\varphi \rightarrow \psi]_\Lambda \vdash \vartheta} \rightarrow L \quad \frac{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda}{\Gamma \vdash [\varphi \rightarrow \psi]_\Lambda} \rightarrow R \\
\\
\frac{\Gamma, [\varphi]_{i \circ \Lambda} \vdash [\psi]_{\Lambda'} \quad \Lambda' \leq \Lambda}{\Gamma, [K_i \varphi]_\Lambda \vdash [\psi]_{\Lambda'}} KL \quad \frac{\Gamma \vdash [\varphi]_{i \circ \Lambda}}{\Gamma \vdash [K_i \varphi]_\Lambda} KR
\end{array}$$

Table 10 – Formulation en calcul des séquents de **OL**

minimale pour cet ordre partiel. En effet, définissons $\min(\Lambda, \Lambda')$ de la façon suivante :

$$\min(\Lambda, \Lambda')(1) = \min \{j \mid \Lambda(1) = \Lambda'(j)\}$$

$$\forall i \in \{2, \dots, |\Lambda|\}, \min(\Lambda, \Lambda')(i) = \min \{j \mid j > \min(\Lambda, \Lambda')(i-1) \text{ et } \Lambda(i) = \Lambda'(j)\}$$

Il est facile de vérifier que $\min(\Lambda, \Lambda')$ est bien une fonction d'inclusion de Λ dans Λ' , et sa définition nous assure qu'elle est bien la plus petite.

De façon duale, on peut définir $\max(\Lambda, \Lambda')$ qui est la plus grande fonction d'inclusion de Λ dans Λ' en posant :

$$\max(\Lambda, \Lambda')(|\Lambda|) = \max \{j \mid \Lambda(|\Lambda|) = \Lambda'(j)\}$$

$$\forall i \in \{1, \dots, |\Lambda| - 1\}, \max(\Lambda, \Lambda')(i) = \max \{j \mid j < \min(\Lambda, \Lambda')(i+1) \text{ et } \Lambda(i) = \Lambda'(j)\}$$

La fonction $\max(\Lambda, \Lambda')$ est par sa définition strictement croissante. On en déduit en particulier que : $\forall i \in \{1, \dots, |\Lambda|\}$, $\max(\Lambda, \Lambda')(i) > i$. De plus, si l'on a trois mots $\Lambda_1 \leq \Lambda_2 \leq \Lambda_3$, on a :

$$\forall i, \max(\Lambda_2, \Lambda_3) \circ \max(\Lambda_1, \Lambda_2)(i) \leq \max(\Lambda_1, \Lambda_3)(i)$$

Cette dernière inégalité vient de la maximalité de $\max(\Lambda_1, \Lambda_3)$ et du fait que $\max(\Lambda_2, \Lambda_3) \circ \max(\Lambda_1, \Lambda_2)$ est une fonction d'insertion de Λ_1 dans Λ_3 . Nous allons maintenant présenter quelques propositions de manipulation de localisations.

Proposition 35

Étant donné quatre mots $\Lambda_1, \Lambda_2, \Lambda, \Lambda' \in \mathcal{I}^*$ tels que $\Lambda \leq \Lambda'$, alors on a :

$$\Lambda_1 \circ \Lambda \circ \Lambda_2 \leq \Lambda_1 \circ \Lambda' \circ \Lambda_2$$

De plus, si $\Lambda \leq_* \Lambda'$ alors on a en fait :

$$\Lambda_1 \circ \Lambda \circ \Lambda_2 \leq_* \Lambda_1 \circ \Lambda' \circ \Lambda_2$$

Preuve La preuve de la comparaison selon \leq provient directement de la définition de cet ordre partiel. Pour le cas \leq_* , on vérifie que les deux mots $\Lambda_1 \circ \Lambda \circ \Lambda_2$ et $\Lambda_1 \circ \Lambda' \circ \Lambda_2$ commencent par la même lettre. \square

Introduisons maintenant l'opération suivante : étant donné deux mots $\Lambda \leq \Lambda'$, on définit le mot Λ'/Λ comme étant :

$$\Lambda'(1 \dots \max(\Lambda, \Lambda')(1) - 1) \circ \Lambda$$

Cette transformation a pour fonction de “nettoyer” un mot Λ' selon l'insertion la plus à droite de Λ . Pour illustrer cela, on a par exemple :

$$\text{abcdadbcbca}/\text{aca} = \text{abcdaca}$$

On peut énoncer deux propriétés simples à propos de cette construction.

Proposition 36

Si l'on a deux mots $\Lambda \leq \Lambda'$, alors on a :

$$\Lambda \leq \Lambda'/\Lambda \leq \Lambda'$$

De plus, si ces mots vérifient $\Lambda \leq_* \Lambda'$, alors on peut écrire :

$$\Lambda \leq_* \Lambda'/\Lambda \leq_* \Lambda'$$

Preuve Les inégalités avec \leq proviennent de la définition de \bullet/\bullet . Pour le cas \leq_* , on vérifie que les trois mots commencent à par la même lettre. \square

Proposition 37

Étant donné trois mots $\Lambda \leq \Lambda_1 \leq \Lambda_2$, on a :

$$\Lambda_1/\Lambda \leq \Lambda_2/\Lambda$$

De plus, si l'on a $\Lambda_1 \leq_* \Lambda_2$, on a :

$$\Lambda_1/\Lambda \leq_* \Lambda_2/\Lambda$$

Preuve Puisque $\max(\Lambda_1, \Lambda_2) \circ \max(\Lambda, \Lambda_1)(1) \leq \max(\Lambda, \Lambda_2)(1)$ on en déduit que :

$$\Lambda_1(1 \dots \max(\Lambda, \Lambda_1)(1) - 1) \leq \Lambda_2(1 \dots \max(\Lambda, \Lambda_2)(1) - 1)$$

Cela entraîne $\Lambda_1/\Lambda \leq \Lambda_2/\Lambda$. Maintenant, si de plus $\Lambda_1 \leq_* \Lambda_2$, alors on a bien $\Lambda_1/\Lambda \leq_* \Lambda_2/\Lambda$ puisque deux mots de la forme Λ et Λ/Λ' (pour $\Lambda' \leq \Lambda$) commencent forcément par la même lettre. \square

Définissons enfin une dernière transformation : étant donné trois mots $\Lambda_1 \leq \Lambda_2 \leq \Lambda_3$, on définit $\Lambda_3/\Lambda_2:\Lambda_1$ comme étant le mot :

$$\Lambda_3(1 \dots \max(\Lambda_2, \Lambda_3)(1) - 1) \circ \Lambda_1$$

En combinant les propositions précédentes, on a :

Proposition 38

Étant donné des mots $\Lambda' \leq_* \Lambda$ et $\Lambda \leq \Lambda_1 \leq \Lambda_2$, on a :

$$\Lambda_1/\Lambda':\Lambda \leq \Lambda_2/\Lambda':\Lambda$$

Si de plus $\Lambda_1 \leq_* \Lambda_2$, on a même :

$$\Lambda_1/\Lambda':\Lambda \leq_* \Lambda_2/\Lambda':\Lambda$$

On a de plus dans tous les cas :

$$\Lambda_1/\Lambda':\Lambda \leq_* \Lambda_1$$

Enfin, nous proposons un résultat sur l'imbrication des constructions $\bullet/\bullet : \bullet :$

Proposition 39

Étant donné quatre mots tels que $\Lambda' \leq_* \Lambda \leq \Lambda_1\Lambda_2$, on a :

$$\Lambda_2/\Lambda_1 : (\Lambda_1/\Lambda:\Lambda') \leq_* \Lambda_2/\Lambda:\Lambda'$$

Preuve Tout d'abord, on remarque que l'on peut écrire :

$$\begin{aligned}\Lambda_2/\Lambda_1 : (\Lambda_1/\Lambda : \Lambda') &= \Lambda_2(1 \dots \max(\Lambda_1, \Lambda_2)(1) - 1) \circ \Lambda_1(1 \dots \max(\Lambda, \Lambda_1)(1) - 1) \circ \Lambda' \\ \Lambda_2/\Lambda : \Lambda' &= \Lambda_2(1 \dots \max(\Lambda, \Lambda_2)(1) - 1) \circ \Lambda'\end{aligned}$$

Mais comme les fonctions \max sont strictement croissantes, on en déduit que :

$$\begin{aligned}\max(\Lambda_2, \Lambda)(1) &\geq \max(\Lambda_1, \Lambda_2)(\max(\Lambda, \Lambda_1)(1)) \\ &\geq \max(\Lambda_1, \Lambda_2)(1) + \max(\Lambda, \Lambda_1)(1) - 1\end{aligned}$$

Cela implique que $\Lambda_2(1 \dots \max(\Lambda_1, \Lambda_2)(1) - 1) \circ \Lambda_1(1 \dots \max(\Lambda, \Lambda_1)(1) - 1)$ est plus petit que $\Lambda_2(1 \dots \max(\Lambda, \Lambda_2)(1) - 1)$. Les deux mots de départ sont donc comparables pour \leq . Pour avoir la comparaison selon \leq_* , on vérifie qu'ils commencent par la même lettre. Mais cela découle du fait que l'on a $\Lambda' \leq_* \Lambda$ qui implique que $\Lambda_1/\Lambda : \Lambda'$ et Λ_1 commencent eux par la même lettre. \square

Nous avons donc défini quelques transformations et prouvé qu'elles se comportent de façon satisfaisante vis-à-vis des ordres \leq et \leq_* . Nous allons nous en servir pour manipuler les localisations de propositions dans des preuves.

Tout d'abord, étudions la façon dont évoluent les localisations des propositions à droite du symbole \vdash dans une preuve d'un séquent $\Gamma \vdash [\varphi]_\Lambda$ ne comportant ni coupure, ni instance de la règle $\rightarrow L$. Elles sont forcément de la forme $\Lambda_1 \circ \Lambda \circ \Lambda_2$. En effet, seules les règles KR et $Univ$ modifient ces localisations, et elles ne font qu'ajouter un indice à gauche ou à droite à la localisation si l'on regarde du bas vers le haut. Il s'agit d'une version plus précise de la proposition 5 p.81.

Concernant les propositions à gauche d'un séquent, leur localisation est supérieur à celle du terme à droite du même séquent (voir la proposition 4 p.80). Nous pouvons donc définir une transformation en "nettoyant" les localisations à gauche selon celle de droite.

Dans la proposition suivante, nous allons en fait combiner les deux propriétés pour transformer une preuve en diminuant la localisation du terme à droite dans le séquent résultat.

Proposition 40

Étant donné une preuve Π du séquent $[\gamma_1]_{\Lambda_1}, \dots, [\gamma_n]_{\Lambda_n} \vdash [\varphi]_\Lambda$ (que l'on peut abrégé en $\Gamma \vdash [\varphi]_\Lambda$) et un mot $\Lambda' \leq_ \Lambda$, on peut transformer Π en une preuve de $[\gamma_1]_{\Lambda_1/\Lambda:\Lambda'}, \dots, [\gamma_n]_{\Lambda_n/\Lambda:\Lambda'} \vdash [\varphi]_{\Lambda'}$ (ou de façon abrégée $\Gamma_{/\Lambda:\Lambda'} \vdash [\varphi]_{\Lambda'}$) avec pour seules modifications l'ajout d'applications de la règle Loc en plus du changement des localisations.*

Preuve On procède par récurrence sur la hauteur de la preuve. Considérons les cas de base, correspondant aux règles *Axiom* et \perp . Dans le premier cas, on part du séquent $[\varphi]_{\Lambda_1} \vdash [\varphi]_\Lambda$

avec $\Lambda \leq_* \Lambda_1$. La proposition précédent nous assure que $\Lambda/\Lambda:\Lambda' \leq_* \Lambda_1/\Lambda:\Lambda'$. Comme de plus $\Lambda/\Lambda:\Lambda'$, on peut transformer la preuve en :

$$\frac{}{[\varphi]_{\Lambda_1/\Lambda:\Lambda'} \vdash [\varphi]_{\Lambda'}} \text{Axiom}$$

Pour la règle \perp , on a bien $\Lambda_1 \leq \Lambda$ qui implique $\Lambda_1/\Lambda:\Lambda' \leq \Lambda'$.

Regardons maintenant le cas des autres règles. Les règles structurelles *Weak*, *Contract*, les règles d'introduction des connecteurs \wedge et \vee ainsi que la règle $\rightarrow R$ ne posent pas de problème. En effet, il n'y a pas de manipulation des localisations. Par exemple, pour la règle *Weak*, on part de :

$$\frac{\Pi_1 \quad \Gamma \vdash [\psi]_{\Lambda}}{\Gamma, [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda}} \text{Weak}$$

L'hypothèse de récurrence permet d'obtenir une preuve Π'_1 de $\Gamma/\Lambda:\Lambda' \vdash [\psi]_{\Lambda'}$, que l'on utilise pour obtenir une preuve de $\Gamma/\Lambda:\Lambda', [\varphi]_{\Lambda:\Lambda'} \vdash [\psi]_{\Lambda'}$ en appliquant la règle *Weak*.

Pour la règle *Loc*, on peut appliquer la récurrence, sachant que que l'on a :

$$\Lambda'_1 \leq_* \Lambda_1 \Rightarrow \Lambda'_1/\Lambda:\Lambda' \leq_* \Lambda_1/\Lambda:\Lambda'$$

Dans le cas de *Univ*, il suffit de voir que $(\Lambda_1/\Lambda:\Lambda') \circ i = (\Lambda_1 \circ i)/(\Lambda \circ i):(\Lambda' \circ i)$. On peut donc appliquer la récurrence aux prémisses avec $\Lambda \circ i$ et $\Lambda' \circ i$. Pour la règle *KL*, on remarque de façon similaire que $(i \circ \Lambda_1)/\Lambda:\Lambda' = i \circ (\Lambda_1/\Lambda:\Lambda')$, alors que la règle *KR* se traite en remarquant que :

$$\Lambda_1/(i \circ \Lambda):(i \circ \Lambda') = \Lambda_1/\Lambda:\Lambda'$$

Il reste donc à traiter le cas des règles *Cut* et $\rightarrow L$, qui sont similaire du point de vue de la localisation. Regardons la règle de coupure. On part d'une preuve Π de la forme :

$$\frac{\Pi_1 \quad \Pi_2 \quad \Gamma \vdash [\varphi]_{\Lambda_1} \quad \Delta, [\varphi]_{\Lambda_1} \vdash [\psi]_{\Lambda}}{\Gamma, \Delta \vdash [\psi]_{\Lambda}} \text{Cut}$$

La récurrence appliquée à Π_2 permet d'obtenir une preuve Π'_2 du séquent $\Delta/\Lambda:\Lambda', [\varphi]_{\Lambda_1/\Lambda:\Lambda'} \vdash [\psi]_{\Lambda'}$. Maintenant, appliquant la récurrence à Π_1 avec Λ_1 et $\Lambda_1/\Lambda:\Lambda'$ à la place de Λ et Λ' . On obtient une preuve Π'_1 du séquent :

$$\Gamma/\Lambda_1:(\Lambda_1/\Lambda:\Lambda') \vdash [\varphi]_{\Lambda_1/\Lambda:\Lambda'}$$

Maintenant, le point important est de remarquer que si $\Lambda_1 \leq \Lambda_2$ et $\Lambda' \leq_* \Lambda$, alors on a d'après la proposition 39 p.182 :

$$\Lambda_2/\Lambda_1 : (\Lambda_1/\Lambda : \Lambda') \leq_* \Lambda_2/\Lambda : \Lambda'$$

Aussi, en appliquant la règle *Loc* de façon adéquate, on obtient une preuve de $\Gamma/\Lambda : \Lambda' \vdash [\varphi]_{\Lambda_1/\Lambda : \Lambda'}$. Pour résumer, la preuve que l'on obtient est :

$$\frac{\frac{\Pi'_1}{\Gamma/\Lambda_1 : (\Lambda_1/\Lambda : \Lambda') \vdash [\varphi]_{\Lambda_1/\Lambda : \Lambda'}}{\Gamma/\Lambda : \Lambda' \vdash [\varphi]_{\Lambda_1/\Lambda : \Lambda'}} \text{Loc} \quad \frac{\Pi'_2}{\Delta/\Lambda : \Lambda', [\varphi]_{\Lambda_1/\Lambda : \Lambda'} \vdash [\psi]_{\Lambda'}} \text{Cut}}{\Gamma/\Lambda : \Lambda', \Delta/\Lambda : \Lambda' \vdash [\psi]_{\Lambda'}}$$

Cette manipulation peut s'adapter facilement au cas de la règle $\rightarrow L$, ce qui conclut notre étude par cas. \square

On peut tirer deux corollaires de cette proposition.

Corollaire 40.1

Étant donné une preuve Π du séquent $\Gamma \vdash [\varphi]_{\Lambda}$ et un mot Λ' tel que $\Lambda' \leq_* \Lambda$, on peut obtenir une preuve Π' de $\Gamma \vdash [\varphi]_{\Lambda'}$ où la structure de la preuve Π' diffère de celle de Π uniquement par l'ajout d'applications de la règle *Loc*.

Preuve En appliquant la proposition 40, on obtient une preuve Π'' de $\Gamma/\Lambda : \Lambda' \vdash [\varphi]_{\Lambda'}$. Puisque l'on a pour tout $\Lambda_1 \geq \Lambda$ la propriété $\Lambda_1/\Lambda : \Lambda' \leq_* \Lambda_1$, on obtient la preuve espérée en appliquant la règle *Loc* aux propositions à gauche du symbole \vdash :

$$\frac{\frac{\Pi''}{\Gamma/\Lambda : \Lambda' \vdash [\varphi]_{\Lambda'}}{\Gamma \vdash [\varphi]_{\Lambda'}} \text{Loc}}{\Gamma \vdash [\varphi]_{\Lambda'}}$$

\square

Corollaire 40.2

Étant donné une preuve Π du séquent $\Gamma, [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda}$ et un mot Λ' tel que $\Lambda' \leq_* \Lambda$, on peut obtenir une preuve Π' de $\Gamma, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda'}$ où la structure de la preuve Π' diffère de celle de Π uniquement par l'ajout d'applications de la règle *Loc*.

Preuve La preuve est la même que pour le corollaire précédent, sachant que l'on peut choisir précisément à quelles propositions on applique la règle *Loc*. \square

C.2 La réduction de coupures

C.2.1 Quelques définitions pratiques

Avec de commencer la preuve proprement dite, nous allons définir quelques notions qui seront utiles pour spécifier l'ordre utilisé pour la récurrence.

Definition 7 (Degré d'une formule)

Le degré $|\varphi|$ d'une formule φ est le nombre de connecteurs logiques présents dans φ .

Pour les deux définitions suivantes, étant donné une preuve Π , les prémisses de la racine de Π sont notées $\Pi_1 \dots \Pi_n$ lorsqu'elles existent. On rappelle que l'on suppose que l'on a un nombre fini d'indices, aussi la règle *Univ* a bien un nombre fini de prémisses.

Definition 8 (Hauteur de coupure d'une preuve)

La hauteur de coupure $ch(\Pi)$ d'une preuve Π est le nombre maximum d'instances de la règle de coupure situées l'une au dessus de l'autre. On la définit inductivement par :

- Si Π est une instance de la règle Axiom ou de \perp , on a $ch(\Pi) = 0$.
- Si la racine de Π est la règle de coupure, on pose $ch(\Pi) = 1 + \max(ch(\Pi_1), ch(\Pi_2))$.
- Enfin, dans les autres cas, on pose $ch(\Pi) = \max(\{ch(\Pi_i)\})$.

Definition 9 (Rang de coupure d'une preuve)

Le rang de coupure $cr(\Pi)$ d'une preuve Π est définie inductivement de la façon suivante :

- Si Π est simplement une instance de la règle Axiom ou de \perp , alors $cr(\Pi) = 0$.
- Si la racine de Π est la règle de coupure Cut avec φ comme formule de coupure, on a $cr(\Pi) = \max(\{|\varphi| + 1, cr(\Pi_1), cr(\Pi_2)\})$.
- Sinon, on a $cr(\Pi) = \max(\{cr(\Pi_i)\})$ pour les règles autres que Cut et ayant au moins une prémisses.

Definition 10 (Profondeur logique d'une preuve)

La profondeur logique $ld(\Pi)$ d'une preuve Π est définie inductivement de la façon suivante :

- Si Π est simplement une instance de la règle Axiom ou de \perp , alors $ld(\Pi) = 0$.
- Si la racine de Π est un règle structurelle à une seule prémisses (soit Weak, Contract ou Loc), $ld(\Pi) = ld(\Pi_1)$.
- Si la racine de Π est une règle logique à une seule prémisses, $ld(\Pi) = ld(\Pi_1) + 1$.
- Si la racine de Π est une règle avec deux prémisses ou plus, $ld(\Pi) = 1 + \max(\{ld(\Pi_i)\})$.

Definition 11 (Taille d'une preuve)

La taille $ps(\Pi)$ d'une preuve Π est le nombre de règles qui la composent. Ainsi, on a $ps(\Pi) = 1 + \sum ps(\Pi_i)$.

Enfin, pour simplifier la présentation des règles de réécriture, nous considérons dans la suite une version plus puissante de la règle de coupure :

Definition 12 (Coupure étendue)

La règle de coupure étendue est définie par :

$$\frac{\Gamma \vdash [\varphi]_{\Lambda} \quad \Delta, n [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda'}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{ExtCut}$$

où la notation $n [\varphi]_{\Lambda}$ désigne une famille de $n \geq 1$ termes de la forme $[\varphi]_{\Lambda_i}$ telle que $\forall i, \Lambda_i \leq_{\star} \Lambda$.

En d'autres termes, il faut au moins que l'une des occurrences de φ ait la localisation Λ , et que pour les autres, la règle de localisation *Loc* puisse s'appliquer.

On vérifie sans peine que les règles *Cut* et *ExtCut* sont équivalentes en présence de *Contract* et de *Loc*. Aussi par la suite, nous n'utiliserons que cette dernière que nous noterons *Cut*.

C.2.2 La réduction proprement dite

La preuve d'élimination de coupures repose principalement sur le lemme suivant qui donne explicitement les différentes transformations à appliquer à une preuve suivant sa structure.

Proposition 41 (Réduction des coupures)

Étant donné une preuve Π du séquent $\Gamma \vdash [\varphi]_{\Lambda}$, il est possible de transformer la preuve Π en une preuve sans coupures du même séquent.

Preuve On procède par induction en utilisant l'ordre lexicographique défini par $(ch(\Pi_1) \leq ch(\Pi_2), cr(\Pi_1) \leq cr(\Pi_2), ld(\Pi_1) \leq ld(\Pi_2), ps(\Pi_1) \leq ps(\Pi_2))$. Il s'agit d'un ordre bien fondé puisque défini comme un ordre lexicographique basé sur des ordres bien fondés.

Les seuls cas à développer sont ceux où la preuve Π est de la forme :

$$\frac{\Pi_1 \quad \Pi_2}{\Gamma \vdash [\varphi]_{\Lambda} \quad \Delta, n [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda'}} \text{Cut}$$

Il existe plusieurs cas à considérer (pas forcément mutuellement exclusifs) suivant la structure des sous-arbres Π_1 et Π_2 .

1. La racine de Π_1 et de Π_2 sont les conclusions d'inférences logiques ayant pour résultat la formule de coupure φ . On dit alors qu'elle est active.

Chaque transformation est donnée en deux versions, l'une pour la règle de coupure standard, l'autre pour la règle de coupure étendue.

Cas de la conjonction On part de la preuve Π :

$$\frac{\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Gamma \vdash [\psi]_\Lambda}}{\Gamma \vdash [\varphi \wedge \psi]_\Lambda} \wedge R \quad \frac{\frac{\pi_3}{\Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Delta, [\varphi \wedge \psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \wedge L}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

que l'on transforme en une nouvelle preuve Π' :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_3}{\Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

Avec cette transformation, on a bien $ch(\Pi') \leq ch(\Pi)$, $cr(\Pi') \leq ch(\Pi)$ et $ld(\Pi') < ld(\Pi)$. On peut donc appliquer la récurrence à la preuve Π' .

Pour la règle de coupure étendue, nous partons maintenant d'une preuve Π de la forme :

$$\frac{\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Gamma \vdash [\psi]_\Lambda}}{\Gamma \vdash [\varphi \wedge \psi]_\Lambda} \wedge R \quad \frac{\frac{\pi_3}{\Delta, [\varphi]_{\Lambda'}, n [\varphi \wedge \psi]_\Lambda \vdash [\vartheta]_{\Lambda''}}{\Delta, (n+1) [\varphi \wedge \psi]_\Lambda \vdash [\vartheta]_{\Lambda''}} \wedge L}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

Soit Π'_2 la preuve obtenue en appliquant l'hypothèse de récurrence à la preuve suivante qui est bien strictement plus petite que Π :

$$\frac{\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Gamma \vdash [\psi]_\Lambda}}{\Gamma \vdash [\varphi \wedge \psi]_\Lambda} \wedge R \quad \frac{\pi_3}{\Delta, [\varphi]_{\Lambda'}, n [\varphi \wedge \psi]_\Lambda \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}$$

De même, on peut appliquer l'hypothèse de récurrence à π_1 pour obtenir π'_1 . Définissons alors Π' comme étant :

$$\frac{\frac{\pi'_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\Pi'_2}{\Gamma, \Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Contract}$$

On a $ch(\Pi') = 1 \leq ch(\Pi)$, alors que $cr(\Pi') = |\varphi| + 1 < |\varphi \wedge \psi| + 1 \leq cr(\Pi)$. On peut donc appliquer la récurrence à Π' .

Cas de la disjonction Dans la version de base, on transforme la preuve Π :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \vee R \quad \frac{\frac{\pi_2}{\Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \quad \frac{\pi_3}{\Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \vee L}{\Delta, [\varphi \vee \psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

en une nouvelle preuve Π' :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Delta, [\varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

De même, la preuve Π' est strictement plus petite que Π , on peut lui appliquer la récurrence. La version étendue se traite comme pour le cas de la conjonction.

Cas de l'implication Cette fois-ci, la transformation à effectuer est la suivante : la preuve Π suivante

$$\frac{\frac{\pi_1}{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda} \rightarrow R \quad \frac{\frac{\pi_2}{\Delta \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi_3}{\Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \rightarrow L}{\Delta, [\varphi \rightarrow \psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

devient la preuve Π' en utilisant la preuve π'_1 obtenue en appliquant le corollaire 40.2 p.185 à π_1 :

$$\frac{\frac{\frac{\pi_2}{\Delta \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi'_1}{\Gamma, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda'}}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{Cut} \quad \frac{\pi_3}{\Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\frac{\Gamma, \Delta, \Delta \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Contract}}$$

On applique alors la récurrence au sous-arbre de racine $\Gamma, \Delta \vdash [\psi]_{\Lambda'}$, à π_3 , en remplaçant dans Π' les résultats obtenus pour obtenir Π'' . On a alors $ch(\Pi'') = 1 \leq ch(\Pi)$ et $cr(\Pi'') = |\psi| + 1 < |\varphi \rightarrow \psi| + 1 \leq cr(\Pi)$. On peut donc conclure en appliquant maintenant la récurrence à Π'' .

Examinons maintenant en détail le cas de la coupure étendue. La preuve Π de départ est de la forme :

$$\frac{\frac{\pi_1}{\Gamma, [\varphi]_\Lambda \vdash [\psi]_\Lambda} \quad \frac{\frac{\pi_2}{\Delta, n[\varphi \rightarrow \psi]_\Lambda \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi_3}{\Delta, n[\varphi \rightarrow \psi]_\Lambda, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}}{\Delta, (n+1)[\varphi \rightarrow \psi]_\Lambda \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

Appliquons l'hypothèse de récurrence pour transformer les deux preuves :

$$\frac{\frac{\pi_1}{\frac{\Gamma, [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda} \rightarrow R}{\Gamma \vdash [\varphi \rightarrow \psi]_{\Lambda}}} \quad \frac{\pi_2}{\Delta, n[\varphi \rightarrow \psi]_{\Lambda} \vdash [\varphi]_{\Lambda'}} \text{Cut}}{\Gamma, \Delta \vdash [\varphi]_{\Lambda'}} \text{Cut}$$

$$\frac{\frac{\pi'_1}{\frac{\Gamma, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda'} \rightarrow R}{\Gamma \vdash [\varphi \rightarrow \psi]_{\Lambda'}}} \quad \frac{\pi_3}{\Delta, n[\varphi \rightarrow \psi]_{\Lambda}, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\psi]_{\Lambda} \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

en Π'_1 et Π'_2 respectivement. On définit alors la preuve Π' comme étant :

$$\frac{\frac{\Pi'_1}{\Gamma, \Delta \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi'_1}{\Gamma, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda'}} \text{Cut}}{\Gamma, \Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{Cut} \quad \frac{\Pi'_2}{\Gamma, \Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\frac{\Gamma, \Gamma, \Gamma, \Delta, \Delta, \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Contract}} \text{Cut}$$

Comme précédemment, on applique la récurrence au sous-arbre de racine $\Gamma, \Gamma, \Delta \vdash [\psi]_{\Lambda'}$ (sa hauteur de coupure est de 1 et son rang de coupe égal à $|\varphi| + 1$ est strictement inférieur à $cr(\Pi)$) puis à la preuve obtenue à partir de Π' après cette dernière transformation.

Cas de l'opérateur de modalité La preuve de la forme :

$$\frac{\frac{\pi_1}{\frac{\Gamma \vdash [\varphi]_{i\circ\Lambda} \text{ KR}}{\Gamma \vdash [K_i \varphi]_{\Lambda}}} \quad \frac{\pi_2}{\frac{\Delta, [\varphi]_{i\circ\Lambda'} \vdash [\vartheta]_{\Lambda''} \text{ KL}}{\Delta, [K_i \varphi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

se transforme de façon directe en une preuve strictement plus petite :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_{i\circ\Lambda}} \quad \frac{\pi_2}{\Delta, [\varphi]_{i\circ\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

On peut appliquer la récurrence à cette dernière.

Dans la version étendue, il faut considérer la preuve Π' obtenue en transformant :

$$\frac{\frac{\pi_1}{\frac{\Gamma \vdash [\varphi]_{i\circ\Lambda} \text{ KR}}{\Gamma \vdash [K_i \varphi]_{\Lambda}}} \quad \frac{\pi_2}{\Delta, [\varphi]_{i\circ\Lambda'}, n[K_i \varphi]_{\Lambda} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\varphi]_{i\circ\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

pour construire la preuve Π :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_{i\circ\Lambda} \quad \Gamma, \Delta, [\varphi]_{i\circ\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta \vdash [\vartheta]_{\Lambda''}} \text{Contract}$$

Les inégalités voulues s'obtiennent de façon classique.

2. La racine de Π_1 ou de Π_2 est l'application d'une règle mettant en jeu une formule φ qui n'est pas la formule de coupe (elle est alors qualifiée de *passive*).

Cas de la conjonction On a une preuve de la forme (où Λ'' est inférieur à Λ et Λ' , lesquels ne sont pas forcément comparables en eux, contrairement aux conventions précédentes) :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_{\Lambda} \quad \Delta, n[\varphi]_{\Lambda}, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Delta, n[\varphi]_{\Lambda}, [\psi_1 \wedge \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \wedge L}{\Gamma, \Delta, [\psi_1 \wedge \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

Cette preuve peut être transformée en une nouvelle preuve où l'on inverse les règles *Cut* et $\wedge L$:

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_{\Lambda} \quad \Delta, n[\varphi]_{\Lambda}, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\psi_1 \wedge \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \wedge L$$

On conclut en appliquant l'hypothèse au sous-arbre dont la racine est $\Gamma, \Delta, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}$, puisque par rapport à la preuve de départ, les hauteur et rang de coupure sont conservés alors que la profondeur logique a diminué de 1 exactement.

Cas de la disjonction On traite une formule de la forme :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_{\Lambda} \quad \Delta, n[\varphi]_{\Lambda}, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''} \quad \Delta, n[\varphi]_{\Lambda}, [\psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Delta, n[\varphi]_{\Lambda}, [\psi_1 \vee \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \vee L}{\Gamma, \Delta, [\psi_1 \vee \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

que l'on peut transformer en :

Comme précédemment, on conclut en appliquant la récurrence aux sous-arbres de racine $\Gamma, \Delta, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}$ et $\Gamma, \Delta, [\psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}$.

Cas de l'implication et de l'opérateur modal On procède comme au-dessus.

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta, n [\varphi]_\Lambda, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta, [\psi_1]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut} \quad \frac{\pi_2 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta, n [\varphi]_\Lambda, [\psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta, [\psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\psi_1 \vee \psi_2]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \vee L$$

Cas de la règle de coupure On part d'une preuve Π de la forme :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta_1, n_1 [\varphi]_\Lambda \vdash [\psi]_{\Lambda'} \quad \Delta_2, n_2 [\varphi]_\Lambda, m [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Delta_1, \Delta_2, n_1 + n_2 [\varphi]_\Lambda \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta_1, \Delta_2 \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

Soient Π'_1 et Π'_2 les preuves obtenues en appliquant la proposition aux preuves strictement plus petites que Π :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta_1, n_1 [\varphi]_\Lambda \vdash [\psi]_{\Lambda'}}{\Gamma, \Delta_1 \vdash [\psi]_{\Lambda'}} \text{Cut} \quad \frac{\pi_2 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta_2, n_2 [\varphi]_\Lambda, m [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta_2, m [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta_1, \Delta_2 \vdash [\vartheta]_{\Lambda''}} \text{Cut}$$

et définissons Π' comme étant la preuve :

$$\frac{\frac{\Pi'_1 \quad \Gamma, \Delta_1 \vdash [\psi]_{\Lambda'} \quad \Pi'_2 \quad \Gamma, \Delta_2, m [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Gamma, \Delta_1, \Delta_2 \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta_1, \Delta_2 \vdash [\vartheta]_{\Lambda''}} \text{Contract}$$

On a alors $ch(\Pi') = 1 < 2 \leq ch(\Pi)$. On peut donc appliquer la récurrence à Π' .

Cas de la contraction, de l'affaiblissement et de la localisation Les trois cas se traitent de façon absolument similaire, en permutant la règle en question avec la règle de coupure. Ainsi, dans le cas de la règle *Contract*, on a une preuve de la forme :

$$\frac{\frac{\pi_1 \quad \Gamma \vdash [\varphi]_\Lambda \quad \Delta, n [\varphi]_\Lambda, [\psi]_{\Lambda'}, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Delta, n [\varphi]_\Lambda, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Contract}}{\Gamma, \Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Contract}$$

que l'on transforme en :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Delta, n [\varphi]_\Lambda, [\psi]_{\Lambda'}, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}}{\Gamma, \Delta, [\psi]_{\Lambda'}, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Cut}}{\Gamma, \Delta, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}} \text{Contract}$$

On applique alors la proposition au sous-arbre dont la racine est $\Gamma, \Delta, [\psi]_{\Lambda'}, [\psi]_{\Lambda'} \vdash [\vartheta]_{\Lambda''}$. Celui-ci a les mêmes hauteur et rang de coupure, ainsi que profondeur logique que la preuve de départ, mais sa taille de preuve est strictement inférieure.

Règle d'universalité Cette règle ne s'applique pas ici, puisque concernant les séquents en entier.

3. La racine de Π_1 ou de Π_2 est le résultat de l'application de la règle *Axiom* ou de \perp . En ce qui concerne la règle *Axiom* utilisée dans le sous-arbre gauche, on peut transformer la preuve Π :

$$\frac{\frac{\text{Axiom}}{[\varphi]_\Lambda \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi_2}{\Delta, n [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''}}{\Delta, [\varphi]_\Lambda \vdash [\psi]_{\Lambda''}} \text{Cut}}$$

en une nouvelle preuve Π' :

$$\frac{\frac{\frac{\pi_2}{\Delta, n [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''}}{\Delta, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''}} \text{Contract}}{\Delta, [\varphi]_\Lambda \vdash [\psi]_{\Lambda''}} \text{Loc}}$$

On vérifie que $ch(\Pi') = ch(\Pi) - 1$. Si la règle *Axiom* est utilisée dans le sous-arbre droite, on transforme le sous-arbre gauche en une preuve de $\Gamma \vdash [\varphi]_{\Lambda'}$ (on utilise ici les notations usuelles) en appliquant le corollaire 40.1 p.185. On applique la récurrence à la preuve obtenue, puisque sa hauteur de coupure est diminuée de 1 par rapport à la preuve de départ.

La règle \perp dans le sous-arbre gauche se traite de façon similaire, en transformant la preuve :

$$\frac{\frac{\perp}{[\perp]_\Lambda \vdash [\varphi]_{\Lambda'}} \quad \frac{\pi_2}{\Delta, n [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''}}{\Delta, [\perp]_\Lambda \vdash [\psi]_{\Lambda''}} \text{Cut}}$$

en une nouvelle preuve :

$$\frac{[\perp]_\Lambda \vdash [\psi]_{\Lambda''}}{\Delta, [\perp]_\Lambda \vdash [\psi]_{\Lambda''}} \text{Weak}$$

Si la règle \perp est utilisée dans le sous-arbre droite, c'est-à-dire si \perp est la formule de coupure, on transforme la preuve :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\perp]_\Lambda} \quad \frac{\perp}{[\perp]_\Lambda \vdash \varphi_{\Lambda'}}}{\Gamma \vdash [\varphi]_{\Lambda'}} \text{Cut}$$

en la preuve de $\Gamma \vdash [\varphi]_{\Lambda'}$ obtenue en appliquant la proposition 6 p.81 à la preuve π_1 . Puisqu'alors la transformation ne change pas la structure de la preuve, à par la transformation éventuelle d'instances de la règle *Axiom* en instances de \perp , on vérifie que la preuve obtenue vérifie les contraintes sur le rang de coupe et la profondeur logique.

4. Le dernier cas à considérer est celui l'on utilise une règle structurelle s'appliquant à la formule de coupe.

Cas de la contraction La règle de contraction se traite de la façon suivante. Partant de la preuve Π :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\frac{\pi_2}{\Delta, (n-1) [\varphi]_\Lambda, [\varphi]_{\Lambda'}, [\varphi]_{\Lambda'} \vdash [\psi]_{\Lambda''}}{\Delta, n [\varphi]_\Lambda \vdash [\psi]_{\Lambda''}} \text{Contract}}{\Gamma, \Delta \vdash [\psi]_{\Lambda''}} \text{Contract}$$

on la transforme en une nouvelle preuve Π' :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\pi_2}{\Delta, (n+1) [\varphi]_\Lambda \vdash [\psi]_{\Lambda''}}{\Gamma, \Delta \vdash [\psi]_{\Lambda''}}$$

et on applique la proposition à cette dernière, ayant $ph(\Pi') < ph(\Pi)$ et des égalités pour *ch*, *cr* et *ld*.

Cas de l'affaiblissement Pour la règle *Weak*, deux cas sont à considérer. Dans le premier cas, on supprime la seule occurrence de la formule de coupe :

$$\frac{\frac{\pi_1}{\Gamma \vdash [\varphi]_\Lambda} \quad \frac{\frac{\pi_2}{\Delta \vdash [\psi]_{\Lambda'}}{\Delta, [\varphi]_\Lambda \vdash [\psi]_{\Lambda'}} \text{Weak}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{Cut}$$

On peut la transformer en :

$$\frac{\frac{\pi_2}{\Delta \vdash [\psi]_{\Lambda'}}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{Weak}$$

Dans le cas où il reste des formules de coupe, on procède comme précédemment, en supprimant l'application de la règle d'affaiblissement, et en appliquant la propriété au résultat obtenu.

Cas de la localisation Pour cette règle, la définition de la règle de coupure étendue que l'on utilise permet de supprimer purement et simplement une occurrence de la règle de localisation dans le cadre d'une coupure. La taille de la preuve diminue alors strictement de 1, on peut appliquer la récurrence à la preuve obtenue.

Cas de l'universalité La preuve de départ à la forme suivante :

$$\frac{\frac{\pi_{1,i}}{\forall i, \Gamma_{oi} \vdash [\varphi]_{\Lambda_{oi}}} \text{Univ} \quad \frac{\pi_2}{\Delta, [\varphi]_{\Lambda} \vdash [\psi]_{\Lambda'}} \text{Cut}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}} \text{Cut}$$

Définissons les preuves $\pi_{2,i}$ obtenues en appliquant la proposition 7 p.82 à π_2 pour ajouter l'indice i à droite, et soient Π'_i les preuves obtenues en appliquant la récurrence à :

$$\frac{\frac{\pi_{1,i}}{\Gamma_{oi} \vdash [\varphi]_{\Lambda_{oi}}} \quad \frac{\pi_{2,i}}{\Delta_{oi}, [\varphi]_{\Lambda_{oi}} \vdash [\psi]_{\Lambda'_{oi}}} \text{Cut}}{\Gamma_{oi}, \Delta_{oi} \vdash [\psi]_{\Lambda'_{oi}}} \text{Cut}$$

On définit alors la preuve Π comme étant :

$$\frac{\frac{\Pi'_i}{\forall i, \Gamma_{oi}, \Delta_{oi} \vdash [\psi]_{\Lambda'_{oi}}} \text{Univ}}{\Gamma, \Delta \vdash [\psi]_{\Lambda'}}$$

On a alors bien $c(\Pi) = \max(\{c(\Pi'_i)\}) \leq |\varphi|$, et $d(\Pi) \leq 1 + 2d(\pi_2) + 2 \max(\{d(\pi_{1,i})\}) \leq 2(d(\Pi_1) + d(\Pi_2))$.

□

Annexe D

Manipulation des hiérarchies

Nous allons maintenant exclusivement nous intéresser à ces hiérarchies de représentation, puisque toutes les conditions que l'on impose sur ces structures vont permettre de définir les constructions que l'on espère (en particulier le produit cartésien et le raffinage) de façon simple.

D.1 Opérations de manipulation des indices

Les premières opérations que nous définissons ici sont des briques de bases permettant des manipulations élémentaires de hiérarchies.

D.1.1 Hiérarchie atomique

La première construction à considérer est la transformation d'une représentation en hiérarchie. Son rôle est purement formel, et permet de construire les hiérarchies "atomiques" que l'on combinera par la suite pour aboutir à des hiérarchies plus complexes.

Ainsi, étant donné une représentation \mathcal{R} avec une description maximale et une conjonction partielle (ce que l'on impose aux représentations présentes dans une hiérarchie), on définit :

$$H(\mathcal{R}) = \langle \{\iota\}, \{\langle \iota, \iota \rangle\}, \{\mathcal{R}_\iota\}, \{f_{\iota|\iota}\} \rangle$$

où l'unique indice est noté ι , la représentation \mathcal{R}_ι est \mathcal{R} elle-même, et la fonction $f_{\iota|\iota}$ est l'identité de \mathcal{R} .

D.1.2 Ajout de la représentation vide

On peut définir une représentation vide \mathcal{R}_\top , ayant une unique description \top_\top (qui donc est la description vide, celle ne fournissant aucune information) en posant :

$$\mathcal{R}_\top = \langle R_\top, \leq_\top, \top_\top \rangle$$

avec $R_{\top} = \{\top\}$, l'ordre partiel \leq_{\top} se déduisant directement de cette définition, puisque R_{\top} ne contient qu'un élément.

Il est possible d'ajouter cette représentation à une hiérarchie de représentations. En effet, soit une hiérarchie définie à l'aide de surjections de Galois

$$\mathcal{H} = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i\}, \{\langle \alpha_{j|i}, \gamma_{i|j} \rangle\} \rangle$$

En supposant que le symbole \top n'appartient pas à \mathcal{I} , on peut définir une nouvelle hiérarchie :

$$\mathcal{H}_{\top} = \langle \mathcal{I}', \triangleleft', \{\mathcal{R}'_i\}_{i \in \mathcal{I}'}, \{\langle \alpha'_{j|i}, \gamma'_{i|j} \rangle\} \rangle$$

où le nouvel ensemble d'indices est $\mathcal{I}' = \mathcal{I} \cup \{\top\}$, l'ordre partiel \triangleleft' est tel que :

$$\begin{aligned} \forall i \in \mathcal{I}', i \triangleleft' \top \\ \forall i, j \in \mathcal{I}, i \triangleleft' j \Leftrightarrow i \triangleleft j \\ \top \triangleleft' i \Rightarrow i = \top \end{aligned}$$

De plus, \mathcal{R}'_{\top} est la représentation vide \mathcal{R}_{\top} , et $\forall i \in \mathcal{I}, \mathcal{R}'_i = \mathcal{R}_i$. Enfin, pour les surjections de Galois, on a :

$$\begin{aligned} \forall i, j \in \mathcal{I}, i \triangleleft j \Rightarrow \langle \alpha'_{j|i}, \gamma'_{i|j} \rangle &= \langle \alpha_{j|i}, \gamma_{i|j} \rangle \\ \forall i \in \mathcal{I}', \begin{cases} \alpha'_{\top|i} = x \mapsto \top_{\top} \\ \gamma'_{i|\top} = x \mapsto \top_i \end{cases} \end{aligned}$$

Cette construction, bien qu'un peu fastidieuse à définir, sera très pratique, utilisée conjointement avec le produit cartésien.

D.1.3 Suppression d'indices

La dernière construction de base sur les hiérarchies est la suppression d'indices. En effet, il peut arriver plusieurs situations où il peut être intéressé de retirer certains indices et les représentations qui correspondent d'une hiérarchie de représentations.

De façon générale, si \mathcal{H} est une hiérarchie de représentation, et si un sous-ensemble $\mathcal{J} \subseteq \mathcal{I}$ vérifie la condition de fermeture 9.3 p. 107, on peut définir la restriction de $\mathcal{H}|_{\mathcal{J}}$ en posant :

$$\mathcal{H}|_{\mathcal{J}} = \langle \mathcal{J}, \triangleleft|_{\mathcal{J} \times \mathcal{J}}, \{\mathcal{R}_i\}_{i \in \mathcal{J}}, \{f_{i|j}\}_{i, j \in \mathcal{J}} \rangle$$

Deux cas particuliers peuvent apparaître régulièrement pour la suppression d'indices :

Suppression vers le haut Étant donné un filtre \mathcal{F} de \mathcal{I} (c'est-à-dire un sous-ensemble tel que $i \in \mathcal{F}$ et $j \triangleleft i \Rightarrow j \in \mathcal{F}$), le complément $\mathcal{I} \setminus \mathcal{F}$ est toujours un ensemble fermé d'indices. On peut donc définir $\mathcal{H}|_{\mathcal{I} \setminus \mathcal{F}}$. Un cas particulier est, en présence d'une hiérarchie avec un indice maximal \top , de considérer $\mathcal{H}|_{\mathcal{I} \setminus \{\top\}}$. En effet, dans la pratique, on peut avoir besoin d'ajouter une représentation vide pour les besoins d'une construction, et de la supprimer après.

Suppression vers le bas On peut adapter la suppression vers le haut pour retirer des indices vers le bas, mais pour que la condition de fermeture soit toujours vérifiée, il faut faire quelques adaptations. Étant donné un indice $i \in \mathcal{I}$, notons $i \downarrow = \{j \mid j \triangleleft i\}$ pour obtenir la définition suivante :

Définition 11 (Idéal premier)

Un idéal premier d'un e.p.o. $\langle P, \leq \rangle$ est un idéal I de P vérifiant la propriété suivante :

$$\forall x, y \in P, x \downarrow \cap y \downarrow \subseteq I \Rightarrow x \in I \text{ ou } y \in I$$

Cette notion d'idéal premier va nous permettre de caractériser les idéaux que l'on peut retirer à un ensemble d'indice pour qu'il soit toujours fermé au sens de la condition 9.3. En effet, si \mathcal{I} est fermé, et $\mathcal{J} \subseteq \mathcal{I}$ est un idéal premier, alors $\mathcal{I} \setminus \mathcal{J}$ est encore fermé.

Cette "direction" de suppression d'indices pourra être utilisée pour retirer les représentations contenant une certaine information. Ainsi, si l'ensemble d'indices \mathcal{I} est le produit $\mathcal{I}_1 \times \mathcal{I}_2$ où \mathcal{I}_2 a un indice maximum \top_2 (ce qui sera une configuration classique en pratique), alors on peut définir $\mathcal{J} = \{\langle i_1, i_2 \rangle \mid i_1 \in \mathcal{I}_1, i_2 \in \mathcal{I}_2 \setminus \{\top_2\}\}$. Il s'agit bien d'un idéal premier de \mathcal{I} , et le complément est $\mathcal{I} \setminus \mathcal{J} = \{\langle i_1, \top_2 \rangle \mid i_1 \in \mathcal{I}_1\}$, de sorte que l'on a effacé toutes les informations correspondant aux représentations associées aux indices de \mathcal{I}_2 . On peut ainsi masquer les informations relatives à une partie du système.

D.2 Produit de systèmes

On rappelle que dans la section 4.2 p.43, on a défini le produit de deux représentation en prenant comme ensemble de descriptions le produit des deux ensembles de départ, et comme ordre partiel l'ordre composante par composante correspondant aux ordres de départ. On peut l'étendre aux représentation avec description vide et conjonction partielle en définissant :

$$\langle R_1, \leq_1, \top_1 \rangle \otimes \langle R_2, \leq_2, \top_2 \rangle = \langle \mathcal{R}_1 \times \mathcal{R}_2, \leq_{1 \otimes 2}, \langle \top_1, \top_2 \rangle \rangle$$

où $\langle d_1, d_2 \rangle \leq_{1 \otimes 2} \langle e_1, e_2 \rangle \Leftrightarrow d_1 \leq_1 e_1$ et $d_2 \leq_2 e_2$.

À partir de ce produit défini pour les représentations, on peut définir le produit de deux hiérarchies de représentations. Intuitivement, on construit le produit en considérant toutes les paires de représentations des hiérarchies de départ, et les fonctions entre elles agissent composante par composante de la façon la plus simple possible. Ainsi, étant donné deux hiérarchies $\mathcal{H}_1 = \langle \mathcal{I}, \triangleleft_1, \{\mathcal{R}_i\}, \{f_{i|j}\} \rangle$ et $\mathcal{H}_2 = \langle \mathcal{J}, \triangleleft_2, \{\mathcal{S}_i\}, \{g_{i|j}\} \rangle$, on définit :

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \left\langle \mathcal{I} \times \mathcal{J}, \triangleleft_{1 \otimes 2}, \{\mathcal{T}_{i,j}\}_{i \in \mathcal{I}, j \in \mathcal{J}}, \{h_{\langle i_1, j_1 \rangle | \langle i_2, j_2 \rangle}\} \right\rangle$$

où les représentations sont définies par $\mathcal{T}_{i,j} = \mathcal{R}_i \otimes \mathcal{S}_j$ et les fonctions de transformations par :

$$h_{\langle i_1, j_1 \rangle | \langle i_2, j_2 \rangle}(\langle e, f \rangle) = \langle f_{i_1 | i_2}(e), g_{j_1 | j_2}(f) \rangle$$

Dans cette définition simple, le produit ne fait que mettre dans une même structure deux hiérarchies de représentations, mais sans ajouter d'autre information. Ainsi, si chaque hiérarchie de départ représente diverses façons d'observer une partie distincte d'un système, le produit de ces hiérarchies n'exprime aucune corrélation entre les représentations des deux parties.

Maintenant que nous avons défini comment étendre le produit à des hiérarchies de représentations, nous allons maintenant aborder la construction de raffinement, qui permettra entre autres d'exprimer des relations entre les différentes composantes d'un produit.

D.3 Raffinage de systèmes

Nous allons ici reprendre une version renforcée parmi les définitions de raffinement données dans la section 4.2 p.43. Ainsi, étant donné une représentation \mathcal{R} , le raffinement que nous considérerons maintenant est défini par un opérateur partiel de clôture inférieure ϕ sur \mathcal{R} . Cet opérateur fournit trois types d'informations sur le raffinement :

- Les descriptions que l'on doit maintenant considérer comme incohérentes. Il s'agit du complément du domaine de définition de ϕ , soit $\mathcal{R} \setminus \text{dom}(\phi)$.
- Les descriptions qui sont suffisamment riches pour le raffinement. Il s'agit de l'ensemble des points fixes $\{d \mid d = \phi(d)\}$.
- La façon dont les autres descriptions doivent être transformées pour inclure les informations relatives à l'application du raffinement. En effet, si l'on a une description $d \in \text{dom}(\phi)$, celle-ci est transformée en $\phi(d)$.

Précisons un peu plus l'application d'un tel opérateur dans le cas où la représentation possède un élément maximum, et une conjonction partielle, comme c'est le cas pour les représentations présentes dans une hiérarchie de représentations.

Definition 13

Étant donné une représentation $\mathcal{R} = \langle R, \leq, \top \rangle$ et un opérateur de clôture inférieure ϕ défini sur \mathcal{R} , son raffinement est la représentation $\phi(\mathcal{R}) = \langle R', \leq', \top' \rangle$ où l'on a :

- $R' = \{\phi(d) \mid d \in R\}$
- $\leq' = \leq|_{R' \times R'}$
- $\top' = \phi(\top)$ puisque $\forall d, d \leq \top \Rightarrow \phi(d) \leq \phi(\top)$.

Pour pouvoir étendre l'application du raffinement à une hiérarchie de représentations, il faut prendre en compte plusieurs éléments. Tout d'abord, il faut considérer qu'il existe un opérateur

de clôture inférieure ϕ_i pour chacune des représentations \mathcal{R}_i , avec des conditions reliant ces opérateurs. De plus, on doit aussi définir la façon de transformer les fonctions de transformations pour tenir compte des changements de représentations. L'approche que nous allons suivre dans la suite va être d'appliquer un opérateur de clôture en un point de la hiérarchie (c'est-à-dire de l'appliquer à une représentation bien précise), et d'étendre son action à l'ensemble de la hiérarchie.

Pour cela, nous allons définir deux façons de modifier les opérateurs de clôture, en allant d'une part "vers le haut" et d'autre part "vers le bas". Nous allons par la suite considérer deux indices i, j de la hiérarchie de représentation tels que $i \triangleleft j$.

Vers le haut

Nous allons nous intéresser premièrement au premier sens, qui est le plus simple. En effet, si l'on a un opérateur de clôture inférieur ϕ_i défini sur \mathcal{R}_i , on peut définir $\phi_j = \alpha_{j|i} \circ \phi_i \circ \gamma_{i|j}$. On vérifie facilement que la fonction ainsi définie est intensive : $\phi_j(d) \leq \alpha_{j|i} \circ \gamma_{i|j}(d) = d$, elle est croissante puisque définie comme la composition de fonctions croissantes, et idempotente. Pour cette dernière propriété, on a en effet :

$$\begin{aligned} \phi_j(d) &= \alpha \circ \phi_i \circ \gamma(d) \\ &= \alpha \circ \phi_i \circ \phi_i \circ \gamma(d) \\ &\leq \alpha \circ \phi_i \circ \gamma \circ \alpha \circ \phi_i \circ \gamma(d) \\ &= \phi_j \circ \phi_j(d) \end{aligned}$$

Vers le bas

Si maintenant on a un opérateur de clôture ϕ_j défini sur \mathcal{R}_j , comment définir un autre opérateur ϕ_i sur \mathcal{R}_i compatible avec le premier, c'est-à-dire tel que :

$$\phi_i \circ \gamma_{i|j}(d) = \gamma_{i|j} \circ \phi_j(d)$$

Tout d'abord, la croissance nous impose :

$$\phi_i(d) \leq \phi_i(\gamma_{i|j} \circ \alpha_{j|i}(d)) = \gamma_{i|j} \circ \phi_j \circ \alpha_{j|i}(d)$$

En tenant compte de l'intensivité, on obtient l'inégalité suivante :

$$\phi_i(d) \leq d \wedge \gamma_{i|j} \circ \phi_j \circ \alpha_{j|i}(d)$$

L'opérateur que l'on cherche est donc le plus petit opérateur de clôture vérifiant cette inégalité. Définissons la suite suivante :

$$d_0 = d \quad d_{k+1} = d_k \wedge \gamma \circ \phi_j \circ \alpha(d_k)$$

Du fait de l'idempotence, on trouve que $\forall k, \phi_i(d) \leq d_k$. En effet, par récurrence, on peut écrire que si $\phi_i(d) \leq d_k$, alors :

$$\begin{aligned}\phi_i(d) &= \phi_i \circ \phi_i(d) \\ &\leq \phi_i(d) \wedge \gamma \circ \phi_j \circ \alpha(\phi_i(d)) \\ &\leq d_k \wedge \gamma \circ \phi_j \circ \alpha(d_k) \\ &= d_{k+1}\end{aligned}$$

Cependant, si la suite $\{d_k\}$ a une limite d_∞ , on a $\phi_i(d) \leq d_\infty$, mais puisque $d_\infty \leq \gamma \circ \phi_j \circ \alpha(d_\infty)$ par définition de la limite, on trouve que c'est en fait la bonne valeur de $\phi_i(d)$. On peut donc écrire :

$$\phi_i(d) = \text{lfp} (x \mapsto x \wedge \gamma_{i|j} \circ \phi_j \circ \alpha_{j|i}(x)) (d)$$

Grâce à la condition 9.4 définie plus haut, on peut s'assurer que le domaine de définition de ϕ_i vérifie bien la condition :

$$d_1 \leq d_2 \Rightarrow d_1 \in \text{dom}(\phi_i) \Rightarrow d_2 \in \text{dom}(\phi_i)$$

On retrouve donc vérifiée la condition imposée sur les opérateurs de clôture pour être utilisables dans le cadre de représentations.

Nous allons maintenant nous intéresser à une propriété de composition des transformations précédentes. Pour cela, soient trois indices $i < j < k$, et supposons que l'on a un opérateur ϕ_k défini sur \mathcal{R}_k . Il existe deux manières de définir un opérateur de clôture sur \mathcal{R}_j à partir de ϕ_k . La première façon est de descendre directement de \mathcal{R}_k à \mathcal{R}_j . On définit donc :

$$\phi_j^1 = d \mapsto \text{lfp} (x \mapsto x \wedge \gamma_{j|k} \circ \phi_k \circ \alpha_{k|j}(x)) (d)$$

Mais on peut définir un second opérateur en descendant jusqu'à \mathcal{R}_i puis en remontant vers \mathcal{R}_j . On définit donc :

$$\phi_j^2 = d \mapsto \alpha_{j|i} \circ \text{lfp} (x \mapsto x \wedge \gamma_{i|k} \circ \phi_k \circ \alpha_{k|i}(x)) \circ \gamma_{i|j}(d)$$

Nous allons montrer que ces deux fonctions sont égales. Tout d'abord, soit $d \in \text{dom}(\phi_j^1)$. On peut donc définir une suite par $d_0 = d$ et $d_{n+1} = d_n \wedge \gamma_{j|k} \circ \phi_k \circ \alpha_{k|j}(d_n)$, et cette suite a une limite d_∞ .

Définissons donc une suite dans \mathcal{R}_i en posant $e_n = \gamma_{i|j}(d_n)$. Comme $\forall n, \gamma_{i|j}(d_\infty) \leq e_n$, on en déduit que la suite $\{e_n\}$ converge. Soit e_∞ sa limite. Or, il est facile de vérifier que cette suite vérifie de plus :

$$e_{n+1} = e_n \wedge \gamma_{i|k} \circ \phi_k \circ \alpha_{k|i}(e_n)$$

On en déduit donc que $\gamma_{i|j}(d)$ (qui est égal à e_0) appartient à $\text{dom}(\phi_i)$, et donc d appartient à $\text{dom}(\phi_j^2)$. Réciproquement, si $d \in \text{dom}(\phi_j^2)$, on en déduit que $\gamma_{i|j}(d) \in \text{dom}(\phi_i)$. Mais alors,

en reprenant les suites $\{d_n\}$ et $\{e_n\}$ définies précédemment, on en déduit que la suite $\{e_n\}$ converge, et donc c'est aussi le cas pour $\{d_n\}$.

On en déduit donc que $d \in \text{dom}(\phi_j^1) \Leftrightarrow d \in \text{dom}(\phi_j^2)$. Les deux domaines sont égaux. Enfin, en utilisant les notations précédentes, on a :

$$\phi_j^1(d) = d_\infty \quad \phi_j^2(d) = \alpha_{j|i}(e_\infty)$$

Mais on a vu que $\gamma_{i|j}(d_\infty) \leq e_\infty$ ce qui implique que $d_\infty \leq \alpha_{j|i}(e_\infty)$. Réciproquement, on a : $\forall n, \alpha_{j|i}(e_\infty) \leq d_n$ et donc $\alpha_{j|i}(e_\infty) \leq d_\infty$ ce qui donne l'égalité entre les deux fonctions.

Construction générale

Nous pouvons maintenant nous pencher sur la méthode pour définir à partir d'un opérateur de clôture inférieure ϕ_i défini sur une représentation \mathcal{R}_i une famille d'opérateurs de clôture $\{\phi'_j : \mathcal{R}_j \rightarrow \mathcal{R}_j\}$ telle que :

$$j \triangleleft k \Rightarrow \phi'_k = \gamma_{k|j} \circ \phi'_j \circ \alpha_{j|k}$$

Pour cela, étant donné $j \in \mathcal{I}$, on pose :

$$\phi'_j = d \mapsto \alpha_{j|k} \circ \text{lfp} (x \mapsto x \wedge \gamma_{k|i} \circ \phi_i \circ \alpha_{i|k}(x)) \circ \gamma_{k|j}(d)$$

où k est un indice appartenant à $i \downarrow \cap j \downarrow$. On peut tout d'abord vérifier que le ϕ'_j ainsi défini ne dépend pas de la valeur de k choisie. Supposons en effet que l'on a deux indices $k_1, k_2 \in i \downarrow \cap j \downarrow$. Soit de plus $l \in k_1 \downarrow \cap k_2 \downarrow$. D'après le résultat précédent, on a :

$$\begin{aligned} & \alpha_{j|k_1} \circ \text{lfp} (x \mapsto x \wedge \gamma_{k_1|i} \circ \phi_i \circ \alpha_{i|k_1}(x)) \circ \gamma_{k_1|j} \\ &= \alpha_{j|k_1} \circ \alpha_{k_1|l} \circ \text{lfp} (x \mapsto x \wedge \gamma_{l|i} \circ \phi_i \circ \alpha_{i|l}(x)) \circ \gamma_{l|k_1} \circ \gamma_{k_1|j} \\ &= \alpha_{j|l} \circ \text{lfp} (x \mapsto x \wedge \gamma_{l|i} \circ \phi_i \circ \alpha_{i|l}(x)) \circ \gamma_{l|j} \\ &= \alpha_{j|k_2} \circ \alpha_{k_2|l} \circ \text{lfp} (x \mapsto x \wedge \gamma_{l|i} \circ \phi_i \circ \alpha_{i|l}(x)) \circ \gamma_{l|k_2} \circ \gamma_{k_2|j} \\ &= \alpha_{j|k_2} \circ \text{lfp} (x \mapsto x \wedge \gamma_{k_2|i} \circ \phi_i \circ \alpha_{i|k_2}(x)) \circ \gamma_{k_2|j} \end{aligned}$$

Ainsi, la famille $\{\phi'_j\}$ est bien définie. On vérifie sans peine que $\phi'_i = \phi_i$, ainsi que $j \triangleleft k \Rightarrow \phi'_k = \gamma_{k|j} \circ \phi'_j \circ \alpha_{j|k}$, grâce à la propriété de composition des α et γ .

On peut donc définir pour chaque représentation sa transformée en fonction de ϕ_i en posant :

$$\forall j, \mathcal{R}'_j = \phi'_j(\mathcal{R}_j)$$

Pour obtenir une hiérarchie de représentations, il reste à transformer les fonctions de transformations, et de vérifier que les conditions sur l'ordre \triangleleft sont bien vérifiées. Si l'on part d'une

description d dans une représentation donnée, et qu'on la transforme à l'aide de fonctions α et γ , pour être sûr que le résultat est stable par l'opérateur de clôture de la représentation d'arrivée, il suffit de l'appliquer. On peut donc définir :

$$\forall i \triangleleft j, \quad \alpha'_{j|i} = \phi'_j \circ \alpha_{j|i} \quad \gamma'_{i|j} = \phi'_i \circ \gamma_{i|j}$$

On peut légèrement simplifier cette notation en remarquant que $\phi'_j = \alpha_{j|i} \circ \phi'_i \circ \gamma_{i|j}$, ce qui fait que si $d = \phi'_i(d)$ alors $\alpha_{j|i}(d) = \phi'_j(\alpha_{j|i}(d))$. Le dernier point à montrer est que $i \triangleleft j \Rightarrow \alpha'_{j|i} \circ \gamma'_{i|j} = \text{id}_j$. Mais on a :

$$\alpha'_{j|i} \circ \gamma'_{i|j} = \alpha_{j|i} \circ \phi'_i \circ \gamma_{i|j} = \phi'_j$$

Or, par définition de \mathcal{R}'_j , on a $d \in \mathcal{R}'_j \Rightarrow d = \phi'_j(d)$.

Toutes ces considérations nous ont montré comment modifier une hiérarchie de représentations à partir d'un opérateur de clôture défini pour une représentation donnée. Nous avons montré comment en déduire un ensemble d'opérateurs permettant de modifier toutes les représentations ainsi que les fonctions de transformations, sans pour autant changer l'ordre sur les indices \triangleleft . Nous résumons maintenant cette construction.

Définition 12

Étant donné une hiérarchie de représentations $\mathcal{H} = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i\}, \{\langle \alpha_{j|i}, \gamma_{i|j} \rangle\} \rangle$, un indice $p \in \mathcal{I}$ et un opérateur de clôture inférieure ϕ défini sur \mathcal{R}_p , on définit le raffinement de \mathcal{H} par ϕ comme la hiérarchie :

$$\mathcal{H}' = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}'_i\}, \{\langle \alpha'_{j|i}, \gamma'_{i|j} \rangle\} \rangle$$

où pour tout $i \in \mathcal{I}$, on a $\mathcal{R}'_i = \phi'_i(\mathcal{R}_i)$, $\alpha'_{i|j} = \alpha_{j|i}|_{\mathcal{R}'_i}$ et $\gamma'_{i|j} = \phi'_i \circ \gamma_{i|j}|_{\mathcal{R}'_j}$. On utilise dans ces définitions les opérateurs ψ'_i définis par :

$$\phi'_i = d \mapsto \alpha_{i|j} \circ \text{lfp} (x \mapsto x \wedge \gamma_{j|p} \circ \phi \circ \alpha_{p|j}(x)) \circ \gamma_{j|i}(d)$$

Nous avons ainsi défini une transformation de hiérarchies de représentations basée sur l'application d'un opérateur de clôture inférieure en un point précis de cette hiérarchie, soit sur une représentation déterminée. Cela traduit le fait que dans l'étude d'un système, si l'on ajoute de l'information structurelle, on le fait en général en se restreignant à une partie bien précise du système, et non au système entier. En utilisant les constructions de base que l'on a définies dans ce chapitre, nous allons présenter plusieurs exemples de manipulations de hiérarchies de représentations.

D.4 Autres constructions

D.4.1 Produit de deux représentations

Nous donnons ici un exemple simple de l'utilisation du produit de hiérarchies conjointement aux opérations simples définies juste avant. L'exemple est le suivant : on part de deux représentations \mathcal{R}_1 et \mathcal{R}_2 et l'on montre comment aboutir à une hiérarchie de représentation contenant trois représentations (celles de départ \mathcal{R}_1 et \mathcal{R}_2 et leur produit $\mathcal{R}_1 \otimes \mathcal{R}_2$) ainsi que les fonctions usuellement associées aux produits cartésiens.

On effectue cette construction en plusieurs étapes :

1. Tout d'abord, on transforme les deux représentations de départ en hiérarchies atomiques.

$$H(\mathcal{R}_1) \quad H(\mathcal{R}_2)$$

2. On ajoute à chacune de ces hiérarchies atomiques une représentation vide.

$$H(\mathcal{R}_1)_\top \quad H(\mathcal{R}_2)_\top$$

3. On effectue le produit cartésien de ces deux représentations.

$$H(\mathcal{R}_1)_\top \otimes H(\mathcal{R}_2)_\top$$

4. On supprime la nouvelle représentation vide, d'indice $\langle \top, \top \rangle$.

$$(H(\mathcal{R}_1)_\top \otimes H(\mathcal{R}_2)_\top) \setminus \{\langle \top, \top \rangle\}$$

Si l'on note $\{\mathcal{R}'_i\}$ l'ensemble des représentations de la hiérarchie finale, on a $\mathcal{R}'_{\langle \ell, \ell \rangle}$ qui est égale $\mathcal{R}_1 \otimes \mathcal{R}_2$, alors que $\mathcal{R}'_{\langle \ell, \top \rangle}$ (resp. $\mathcal{R}'_{\langle \top, \ell \rangle}$) est isomorphe à \mathcal{R}_1 (resp. \mathcal{R}_2). De plus, on a pour les fonctions :

$$\alpha_{\langle \ell, \top \rangle | \langle \ell, \ell \rangle} = \langle x, y \rangle \mapsto x \quad \gamma_{\langle \ell, \ell \rangle | \langle \ell, \top \rangle} = x \mapsto \langle x, \top \rangle$$

Les différentes étapes que l'on vient de décrire sont représentées dans la figure D.1. On voit dans cette construction que l'on a bien les représentations, ainsi que les projections et insertions qui sont présentes.

D.4.2 Application de fonctions

L'ajout d'une contrainte structurelle sous forme d'une fonction peut être formalisée en utilisant la construction de raffinement. Un tel ajout de contrainte peut s'exprimer de la façon suivante : supposons que l'on a deux indices $i, j \in \mathcal{I}$, et que la contrainte s'exprime par une fonction croissante $f : \mathcal{R}_i \rightarrow \mathcal{R}_j$ dont l'interprétation est la suivante : une connaissance depuis le point

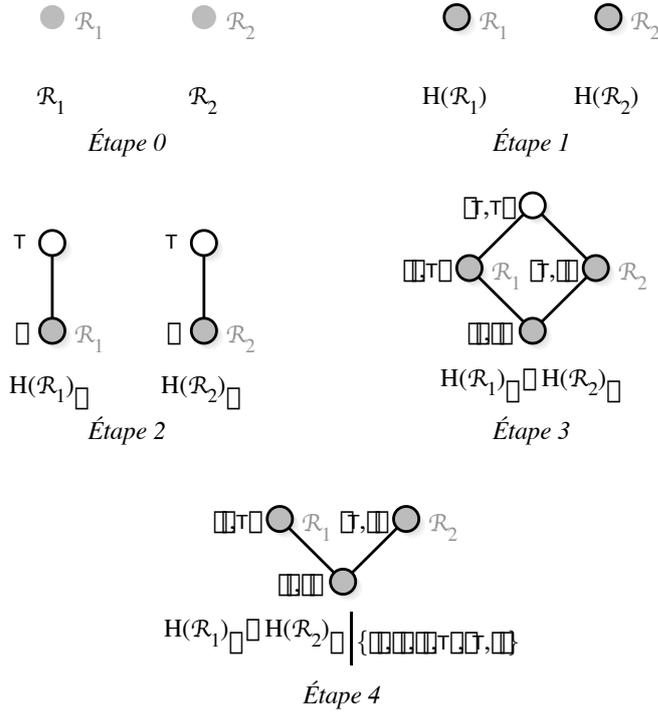


FIG. D.1 – Produit de deux représentations

d’observation \mathcal{O}_i exprimée par une description d_i entraîne des informations sur la représentation \mathcal{R}_j exprimée par $f(d_i)$.

Grâce à la propriété de fermeture de l’ensemble des indices d’une hiérarchie, il est possible d’avoir un indice k tel que $k \triangleleft i$ et $k \triangleleft j$. On peut alors transformer la fonction précédente en une fonction ayant la même représentation (ici \mathcal{R}_k) comme domaine et co-domaine en posant $f' = \gamma_{k|j} \circ f \circ \alpha_{i|k}$. Cette fonction f' représente exactement le même apport d’information que la fonction f , puisque $f = \alpha_{j|k} \circ f' \circ \gamma_{k|i}$. On peut donc se limiter aux fonctions d’une représentation vers elle-même.

Ainsi, considérons une représentation \mathcal{R}_i et une fonction $f : \mathcal{R}_i \rightarrow \mathcal{R}_i$ qui traduit l’apport d’information structurelle. En partant d’une description $d \in \mathcal{R}_i$, il faut y ajouter des informations correspondant à la description $f(d) \in \mathcal{R}_i$, pour obtenir $d \wedge f(d)$. Il est donc intéressant de considérer la fonction $d \mapsto d \wedge f(d)$ pour effectuer le raffinement de la représentation. Mais ce n’est pas en général un opérateur de clôture, car si elle est bien croissante et intensive, l’idempotence n’est pas toujours vérifiée.

On peut identifier quelques cas où l’on peut facilement déduire un opérateur de clôture de cette fonction :

- On peut tout d’abord s’intéresser au cas où la fonction $d \mapsto d \wedge f(d)$ est un opérateur de

clôture. Pour cela, il faut que la propriété d'idempotence soit vérifiée. Elle s'écrit :

$$\forall d, d \wedge f(d) = d \wedge f(d) \wedge f(d \wedge f(d))$$

On en déduit la condition nécessaire et suffisante :

$$f(d \wedge f(d)) \leq d \wedge f(d)$$

- Un cas particulier de la situation précédente est exprimé par la condition suivante :

$$\forall d, f(d \wedge f(d)) = f(d)$$

Cette situation est celle que l'on rencontre par exemple dans le cas d'un produit cartésien de représentation : si l'on a deux représentations \mathcal{R}_i et \mathcal{R}_j , et leur produit $\mathcal{R}_k = \mathcal{R}_i \otimes \mathcal{R}_j$, pour une fonction croissante $f : \mathcal{R}_i \rightarrow \mathcal{R}_j$, en posant $f' = \gamma_{k|j} \circ f \circ \alpha_{i|k} = \langle x, y \rangle \mapsto \langle x, f(x) \rangle$, on a :

$$f'(\langle x, y \rangle \wedge f'(\langle x, y \rangle)) = f'(\langle x, y \wedge f(x) \rangle) = \langle x, y \wedge f(x) \rangle = f'(\langle x, y \rangle)$$

- Si la représentation \mathcal{R}_k est complète pour l'ordre, on peut définir une suite x_n en posant $x_0 = x$ et $x_{n+1} = x_n \wedge f(x_n)$. Deux cas se présentent alors : soit la suite est définie pour tout n , auquel cas on définit $\phi_f^\infty(x) = \liminf \{x_n\}_{n \in \mathbb{N}}$, soit la suite n'est pas définie pour tout n , auquel cas ϕ_f^∞ n'est pas définie pour x . On vérifie aisément que ϕ_f^∞ vérifie les propriétés suivantes : c'est un opérateur de clôture inférieure dont le domaine de définition est un filtre de \mathcal{R}_k , qui vérifie $\forall d \in \text{dom}(\phi_f^\infty), \phi_f^\infty(d) \leq f(d)$ et c'est le plus grand opérateur de clôture ayant cette propriété.

D.4.3 Partition d'une représentation

La partition d'une représentation permet son découpage en deux parties, permettant d'imiter le comportement de booléens au sein d'une représentation. On définit tout d'abord la représentation booléenne \mathcal{R}_B comme étant composée de trois éléments \top, t, f avec \top qui est la description vide et les autres descriptions t et f ne sont pas comparables.

L'idée de la partition est alors d'associer un idéal à la description t et un autre à f . Pour cela, on se donne deux idéaux P et N de \mathcal{R} , ainsi que deux descriptions d_P et d_N vérifiant :

$$\forall d \in P, d \leq d_P \quad \forall d \in N, d \leq d_N$$

Ainsi, P (resp. N) correspond à t (resp. f). On se donne de plus deux descriptions (d_P et d_N) pour représenter les deux valeurs booléennes. Il est à noter que l'on ne veut pas que \mathcal{R}_B deviennent une approximation de \mathcal{R} , ce qui implique que l'on n'impose ni $d_P \in P$, ni $d_N \in N$.

On définit donc la partition de \mathcal{R} suivant $\langle P, d_P \rangle$ et $\langle N, d_N \rangle$ comme $\phi_C(\mathcal{R} \otimes \mathcal{R}_B)$ où la fonction C est défini par les gardes suivantes :

$$\begin{aligned} \forall d \in P, \langle d, \top \rangle &\mapsto \langle \top, t \rangle \\ \forall d \in N, \langle d, \top \rangle &\mapsto \langle \top, f \rangle \\ \langle \top, t \rangle &\mapsto \langle d_P, \top \rangle \\ \langle \top, f \rangle &\mapsto \langle d_N, \top \rangle \end{aligned}$$

Il est facile de vérifier que C vérifie : $\forall d, C(d \wedge C(d)) = C(d)$. Un schéma résumant cette construction est donné en figure D.2.

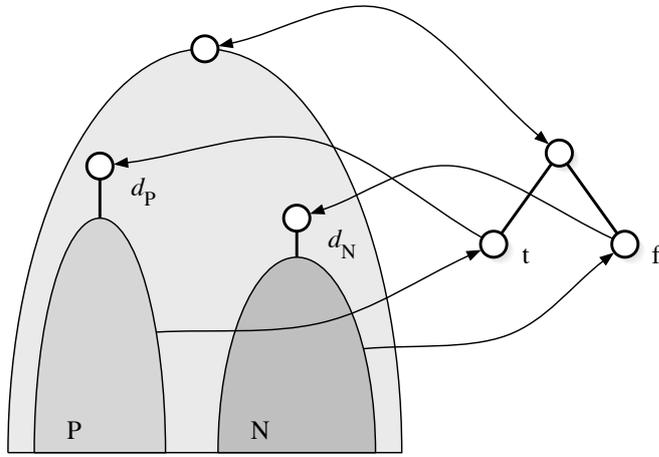


FIG. D.2 – Partition d'une représentation

D.4.4 Ensembles d'idéaux

Étant donné une représentation \mathcal{R} , on peut définir un autre ensemble partiellement ordonné en considérant le poset de ses idéaux $\mathcal{R}^\downarrow = \langle \wp^\downarrow(\mathcal{R}), \subseteq \rangle$. Cette construction permet d'étendre la structure d'une hiérarchie de représentations en considérant des ensembles de descriptions. Cette extension peut aussi être appliquée aux fonctions croissantes, puisque si l'on a une telle fonction $f : \mathcal{R}_1 \rightarrow \mathcal{R}_2$, on peut définir :

$$\begin{aligned} f^\downarrow : \wp(\mathcal{R}_1) &\rightarrow \wp(\mathcal{R}_2) \\ I &\mapsto \{d \mid \exists d' \in I : d \leq f(d')\} \\ &= \bigcup_{d \in I} \{d' \mid d' \leq f(d)\} \end{aligned}$$

On peut adapter la transformation d'une représentation en l'ensemble de ses idéaux à une hiérarchie de représentations, en définissant (avec les notations habituelles) :

$$\mathcal{H}^\downarrow = \langle \mathcal{I}, \triangleleft, \{\mathcal{R}_i^\downarrow\}, \{f_{i|j}^\downarrow\} \rangle$$

Proposition 42

Étant donné une hiérarchie de représentations, la structure \mathcal{H}^\downarrow en est une elle aussi.

Preuve Il suffit pour prouver cela de vérifier les propriétés suivantes :

$$\begin{aligned} \forall i \triangleleft j, f_{j|i}^\downarrow \circ f_{i|j}^\downarrow &= \text{id}_j^\downarrow \\ \forall i, j, k, f_{i|k}^\downarrow &\subseteq f_{i|j}^\downarrow \circ f_{j|k}^\downarrow \end{aligned}$$

En effet, on a :

$$\begin{aligned} i \triangleleft j \Rightarrow f_{j|i}^\downarrow \circ f_{i|j}^\downarrow(I) &= f_{j|i}^\downarrow \left(\bigcup_{d \in I} \{d' \mid d' \leq f_{i|j}(d)\} \right) \\ &= \bigcup_{d \in I} f_{j|i}^\downarrow(\{d' \mid d' \leq f_{i|j}(d)\}) \\ &= \bigcup_{d \in I} \{d' \mid d' \leq f_{j|i} \circ f_{i|j}(d)\} \\ &= \bigcup_{d \in I} \{d' \mid d' \leq d\} \\ &= I \end{aligned}$$

$$\begin{aligned} f_{i|j}^\downarrow \circ f_{j|k}^\downarrow(I) &= \bigcup_{d \in I} \{d' \mid d' \leq f_{i|j} \circ f_{j|k}(d)\} \\ &\supseteq \bigcup_{d \in I} \{d' \mid d' \leq f_{i|k}(d)\} \\ &= f_{i|k}^\downarrow(I) \end{aligned}$$

□

On peut encore enrichir cette construction en reliant la hiérarchie \mathcal{H}^\downarrow à celle d'origine (\mathcal{H}) dans le cas où il est possible d'établir une surjection de Galois entre toutes les représentations \mathcal{R}_i et leur ensemble d'idéaux \mathcal{R}_i^\downarrow . C'est par exemple possible lorsque l'on peut définir une disjonction complète \vee_i sur chacune des \mathcal{R}_i . On définit alors $\gamma_i : \mathcal{R}_i \rightarrow \mathcal{R}_i^\downarrow$ en posant $\gamma_i(d) = \{d' \mid d' \leq d\}$ et inversement, $\alpha_i(I) = \vee_i \{d \in I\}$. Ces fonctions vérifient $\alpha_i \circ \gamma_i(d) = \vee_i \{d' \mid d' \leq d\} = d$ et :

$$\alpha_i(I) \leq d \Leftrightarrow \forall d' \in I, d' \leq d \Leftrightarrow I \subseteq \gamma_i(d)$$

Dans ce cas, on peut définir une hiérarchie de représentation :

$$\mathcal{H}' = \langle \mathcal{I}', \triangleleft', \{\mathcal{R}'_i\}, \{\langle \alpha'_{j|i}, \gamma'_{i|j} \rangle\} \rangle$$

où l'on a :

$$\begin{aligned}
\mathcal{I}' &= \{\langle i, j \rangle \mid i \in \mathcal{I}, j \in \{1, 2\}\} \\
\langle i_1, j_1 \rangle \triangleleft' \langle i_2, j_2 \rangle &\Leftrightarrow i_1 \leq i_2 \text{ et } j_1 \leq j_2 \\
\mathcal{R}'_{i,1} &= \mathcal{R}_i^\downarrow \quad \mathcal{R}'_{i,2} = \mathcal{R}_i \\
\alpha'_{\langle j,1 \rangle | \langle i,1 \rangle} &= \alpha_{j|i} \quad \gamma'_{\langle i,1 \rangle | \langle j,1 \rangle} = \gamma_{i|j} \\
\alpha'_{\langle j,2 \rangle | \langle i,2 \rangle} &= \alpha_{j|i}^\downarrow \quad \gamma'_{\langle i,2 \rangle | \langle j,2 \rangle} = \gamma_{i|j}^\downarrow \\
\alpha'_{\langle j,2 \rangle | \langle i,1 \rangle} &= \alpha_j' \circ \alpha_{j|i} \quad \gamma'_{\langle i,1 \rangle | \langle j,2 \rangle} = \gamma_i' \circ \gamma_{i|j}
\end{aligned}$$

Il est aisé de vérifier que \mathcal{H}' est bien une hiérarchie de représentations.

D.4.5 Union de deux éléments

On propose ici une autre méthode pour enrichir une représentation de telle sorte que l'on ajoute une description correspondant à la réunion de deux autres. Cette méthode utilise l'opération de raffinement en conjonction avec une fonction se comportant comme la disjonction booléenne. On définit pour cela la fonction u sur le produit $\mathcal{R}_B \otimes \mathcal{R}_B \otimes \mathcal{R}_B$ par les gardes suivantes (on note \cdot à la place de \top pour améliorer la lisibilité) :

$$u \left\{ \begin{array}{l} \langle f, f, \cdot \rangle \mapsto \langle \cdot, \cdot, f \rangle \\ \langle t, \cdot, \cdot \rangle \mapsto \langle \cdot, \cdot, t \rangle \\ \langle \cdot, t, \cdot \rangle \mapsto \langle \cdot, \cdot, t \rangle \\ \langle \cdot, \cdot, f \rangle \mapsto \langle f, f, \cdot \rangle \\ \langle f, \cdot, t \rangle \mapsto \langle \cdot, t, \cdot \rangle \\ \langle \cdot, f, t \rangle \mapsto \langle t, \cdot, \cdot \rangle \end{array} \right.$$

La figure D.3 montre comment, en appliquant $\phi_u : d \mapsto d \wedge u(d)$, on obtient une représentation équivalente à $\mathcal{R}_B \otimes \mathcal{R}_B$ à laquelle on a ajouté une description correspondant à l'union de $\langle t, \top \rangle$ et $\langle \top, t \rangle$ (il s'agit de la représentation $\langle \cdot, \cdot, t \rangle$ marquée en noir).

On peut appliquer cette transformation dans le cas général en se servant des partitions. Ainsi, si on veut construire la réunion de deux descriptions d_1 et d_2 , on effectue deux fois la partition, l'une selon $P_1 = \{d \mid d \leq d_1\}$ et $N_1 = \{d \mid \forall d' \leq d, d' \not\leq d_1\}$ et l'autre selon P_2 et N_2 définis similairement. On peut alors appliquer l'union booléenne définie ci-dessus aux deux booléens représentant d_1 et d_2 .

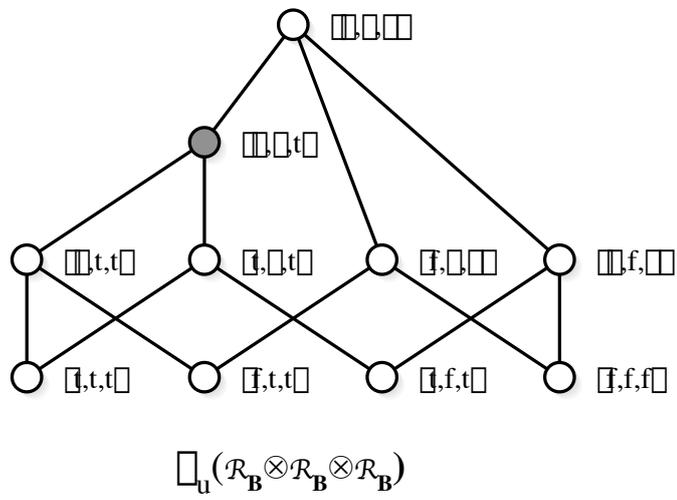
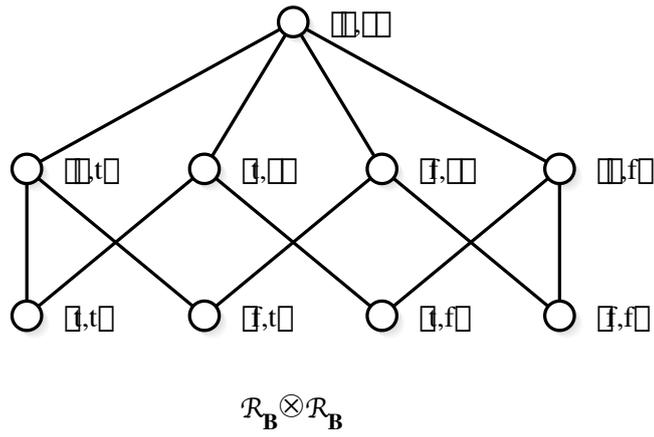


FIG. D.3 – Construction de l’union de deux éléments

Résumé

On s'intéresse à la connaissance que l'on peut avoir d'un système en se basant uniquement sur des observations que l'on peut en faire et où certaines informations peuvent rester cachées. On peut structurer ces observations en comparant leur contenu et la quantité d'informations qu'elles fournissent, pour obtenir ce que nous nommons des représentations. On peut de plus étudier les relations existant entre les différentes façons d'observer un même système, pour obtenir certaines fonctions reliant les représentations entre elles.

Avec ce formalisme, on se livre à une étude logique du comportement de l'information pour cette approche. Le premier résultat est que l'on se base sur la logique intuitionniste, puisque les propositions que l'on considère expriment des connaissances sûres, et que l'ajout d'information n'en modifie pas la véracité.

On étend cette logique en utilisant des opérateurs "modaux" pour symboliser les différentes façons d'observer le système et exprimer le fait qu'une information est accessible ou non depuis le point de vue correspondant. Suivant les contraintes que l'on impose, on obtient plusieurs comportements de ces opérateurs dont découlent plusieurs logiques proches de la logique nommée **IS4**.

Le postulat de base utilisé (on étudie un système en l'observant) est très général. Or, notre étude montre que cela impose une logique relativement faible, puisque ni le tiers-exclus, ni l'axiome modal **5** ne sont vérifiés, et ne peuvent l'être même en ajoutant des hypothèses. Cela signifie que seuls les éléments que l'on manipule, soit les résultats d'observations, sont importants. On est donc obligé de raisonner de façon constructive à partir de ceux-ci et la non-observation d'un fait ne permet pas d'en déduire sa négation.

Ainsi, les seuls éléments dont il faut tenir compte dans l'observation et l'étude de la nature sont les observations que l'on en fait et toute connaissance s'obtient de façon strictement déductive à partir de celles-ci.

Mots-clés: Représentation de connaissances, Logique modale, Logique intuitionniste, Théorie de l'observation

Abstract

We focus on the study of knowledge obtained by performing partial observations of a system. The result of those observations can be structured by comparing their informational

content, and this is how we define what we call representations. It is also possible to study the possible relationships existing between different observations methods, which leads to the existence of functions between representations.

With this formalism, we make a logical study of the way information behaves in such a context. The first point which arises from this is that one has to use intuitionistic logic, since propositions are about facts and not beliefs, and the addition of information does not change their veracity.

This logic is extended by adding “modal” operators, which correspond to different ways of observing a system and express the fact that a piece of information is accessible or not from a given point of view. Depending on the constraints which are applied to our structures, one gets several behaviors for those operators and as many logics, all similar to the modal logic IS4.

The basic postulate, that is a system is studied by observing it, is extremely general. However, our study shows that this leads to a rather weak logic, since neither the excluded middle principle nor the modal axiom 5 can be verified, even if extra hypotheses are considered. This means that the only important elements are the results of observations, and any reasoning is done constructively from them. As a consequence, the non-observation of a fact can not be used as a proof of its negation.

In conclusion, the only elements which have to be taken into account when observing and studying nature are the observations which can be done of it, and any knowledge is obtained deductively from it.

Keywords: Knowledge Representation, Modal Logic, Intuitionistic Logic, Observation Theory