THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : Université Toulouse III Paul Sabatier (UPS)

Discipline ou spécialité : Informatique

Présentée et soutenue par :

Pablo SEBAN

Le 20 Juin 2011

Titre :

Who May Say What ? Thougths about Objectivity, Group Ability and Permission in Dynamic Epistemic Logic

---

Directeurs de thèse :

Philippe BALBIANI
Directeur de Recherche CNRS
Université Paul Sabatier

Hans VAN DITMARSCH
Chercheur associé
Université de Séville

Rapporteurs :

Frédéric CUPPENS
Professeur
Télécom Bretagne

Viorica SOFRONIE-STOKKERMAN
Chercheuse
Institut Max-Planck

Rineke VERBRUGGE
Professeure
université de Groningen

Examinateurs :

Claudette CAYROL
Professeure - Université Paul Sabatier

Emiliano LORINI
Chercheur CNRS - Université Paul Sabatier

École doctorale :

Mathématiques Informatique Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de recherche en Informatique de Toulouse (IRIT)

Who May Say What?

Thoughts on Objectivity, Group Ability, and Permission in Dynamic Epistemic Logic

Pablo Seban

To my father,
the greatest card player
I have ever met.

Remerciements

Je remercie d'abord mes directeurs, Philippe Balbiani et Hans van Ditmarsch. Philippe pour le sérieux avec lequel il travaille, dans ses recherches et dans ses collaborations, pour son soutien quotidien, ses éclaircissements permanents, sa très grande disponibilité. Hans pour avoir voulu être mon co-directeur et avoir joué ce rôle parfaitement, pour m'avoir accueilli à deux reprises, pour ses concerts de Violoncelle et cette magnifique balade à vélo. Merci à tous les deux pour m'avoir fait confiance dès le début – et pour l'avoir maintenue durant ces années bien que je vous aie donné plus d'une occasion d'être déçus.

Merci à Rineke Verbrugge, Viorica Sofronie-Stokkermans et Frédéric Cuppens d'avoir accepté le travail pesant de relecture du présent mémoire, à Claudette Cayrol et Emiliano Lorini pour avoir accepté d'être membres de mon jury. Merci également aux relecteurs anonymes des différents articles que j'ai pu soumettre, leurs éclairages étaient toujours utiles et profitables.

La liste est longue ensuite de celles et ceux avec qui j'ai pu collaborer de près ou de loin durant ces trois années et que je tiens à remercier. Toute l'équipe LILaC d'abord, qui forme un cadre de travail idéal si l'on cherche à chercher. Je remercie en particulier mes co-bureau et amis François et Nadine pour... tout, et Bilal, Fahima, Marwa, Mounira, Srdjan, Manu, Anaïs, Pierre et les autres "jeunes chercheurs" du 3ème étage de l'IRIT. Merci pour les discussions, le thé du bureau d'à côté,pour la musique, et j'en passe.

Les corrections d'anglais de Marwa, de mon frère Pierrot et d'Agathe m'ont permis de finaliser ce mémoire, ainsi que les dessins de ma soeur Brune. Merci à vous pour ça.

Merci à ceux qui m'ont accueilli dans leur équipe: Jan Broersen et son équipe de l'Université de Ütrecht, Barteld Kooï de l'Université de Groningen, Angel Nepomuceno et toute l'équipe de l'Université de Séville où j'ai passé des moments inoubliables.

Merci à vous qui vous reconnaitrez pour les discussions stimulantes dans les conférences, workshop ou écoles d'été, merci à celles et ceux qui se sont passioné-e-s avec moi pour Eleusis, et merci à toi Fede.

Ces trois années, de nombreuses personnes se sont battues pour que l'université reste un lieu d'émancipation plutôt que de rentabilité et de compétition. Les débats que ces mouvements ont soulevés ont été importants pour le monde de l'enseignement et de la recherche... et pour moi. Merci encore à toutes celles et tous ceux qui y ont participé.

Merci à Mme Albert, M. Alessandri et feu M. Exbrayat du lycée Joffre de Montpellier pour m'avoir fait aimer la beauté des maths (que j'avais intuitée), à Alex pour m'avoir donné envie de chercher et à Jacques Sauloy pour m'avoir permis d'apprendre à enseigner.

Et puis il y a tous les autres, famille, amours, camarades ou amis, qui ont partagé ces trois années avec moi, sans qui je n'aurais pu ni commencer ni surtout terminer ce travail: les châtelains officiels ou officieux qui m'ont apporté ce cadre de solidarité et de respect dans lequel je vis, Richard, Renata, Julien, Martina, Robin, Alicia, Aurélie (à qui je dois tant), les Couvreu-Planche & Co., la "God save", les membres de "la guerre tue" et de "Générations spontanées", tous mes camarades et mes amis de Toulouse ou d'ailleurs, pour l'énergie toujours renouvelée.

Un dernier merci et j'arrête: les membres de ma petite famille qui m'ont toujours soutenu, parfois malgré moi. Merci à vous *anche per questo*.

Acknowledgments

I would like to start by thanking my supervisors Philippe Balbiani and Hans van Ditmarsch. Philippe for how seriously his work is, both in his researches and his collaborations, for his daily support, his permanent clearing up, his grand availability. Hans for having wanted to be my co-supervisor and perfectly played the role, for welcoming me twice, for when he played the cello and the beautiful bicycle ride. Thanks to both of you for trusting me from the start – and keeping on doing it, three years long, although I gave you more than one occasion to be disappointed.

Thanks to Rineke Verbrugge, Viorica Sofronie-Stokkermans and Frédéric Cuppens for accepting the heavy task of rereading the present paper, to Claudette Cayrol and Emiliano Lorini for accepting to be part of my jury. Also thanks to the anonymous reviewers of the articles I have submitted, the light they spread was always useful and beneficial.

Then stands a long list of those I have worked with, more or less directly, these three years, and whom I want to thank. The whole LILaC team, in the first place, the ideal researching environment if one wants to research. I particularly thank my office companions, and friends, François and Nadine for... everything, and Bilal, Fahima, Marwa, Mounira, Srdjan, Manu, Anaïs, Pierre and all the other IRIT's third floor's "young researchers". Thanks for discussions, next-door tea, music, and so on.

English corrections by Marwa, my brother Pierrot and Agathe allowed me to finalise this paper, and so did my sister Brune's drawing. Thank you for that.

Thanks to those who welcomed me in their team : Jan Broersen and his research team in Ütrecht University, Barteld Kooï from Groningen University, Angel Nepomuceno and the whole team in the University of Seville where I spend moments I wont forget.

Thanks to you who entered in challenging discussions with me, during conferences, workshops or summer schools, thanks to those who got passionate for Eleusis like I did, and thanks to you Fede.

The last three years, many people fought for the university to remain a place of emancipation rather than profit and competition. The debates raised by these movements were important to the world of teaching and research– and for me. Thanks again to everyone who participated in them.

Thanks to Mme Albert, M. Alessandri and the late M. Exbrayat, from Joffre high-school in Montpellier, for making me love the beauty of mathematics (that I had intuited), to Alex for giving me the will to research and to Jacques Sauloy for allowing me to learn how to teach.

And then, all the others, relatives, loves, comrades and friends who shared the last three years with me, and without whom I couldn't have started and especially not finished this work : the inhabitants of the castle, official or not, who brought me the environment of solidarity and respect I live in, Richard, Renata, Julien, Martina, Robin, Alicia, Aurélie (to whom I owe so much), the Couvreu-Planche & Co., the "God save", all the members of "la guerre tue" and "Générations spontanées", all my comrades and my friends from Toulouse or elsewhere, for their always renewed energy.

A last thank you before the end : to the members of my kind family who always supported me, at times even against my will. Thanks to you, *anche per questo*.

Contents

0 Introduction en français **1**

0.1 Introduction 1

0.1.1 Qu'est-ce que la *logique*? 1

0.1.2 Aperçu général 4

0.2 Logiques modales pour la représentation de la connaissance 5

0.2.1 Représentation de la connaissance 5

0.2.2 Notions techinques classiques en logique modale 10

1 Introduction **23**

1.1 What is *Logic*? 23

1.2 Outline 26

2 Modal Logic for the Representation of Knowledge **27**

2.1 Representation of Knowledge 27

2.1.1 The Language of Epistemic Logic 27

2.1.2 The *Possible Worlds* Semantics 29

2.1.3 Characterisation of Knowledge 31

2.2 Classical Technical Notions in Modal Logic 32

2.2.1 Properties of the Language 32

2.2.2 Properties of Models 34

2.2.3 Axiomatization 37

2.2.4 Decidability and Classes of Complexity 40

3 State of Art **43**

3.1 Dynamic Epistemic Logics 43

3.1.1 Public Announcement Logic 43

3.1.2 What Is Knowable? - Arbitrary Public Announcement Logic 47

3.1.3 Not All Announcements Are Public and Made by an Omniscient Agent - Action Model Logic 49

3.1.4 Objective Beliefs 52

3.1.5 Partial Conclusion 54

3.2 Deontic Logics 54

3.2.1 Standard Deontic Logic 55

3.2.2 Dynamic Logics of Permission 57

3.3 Permission and Epistemic Actions 62

4 Logics of Objective Beliefs **65**
4.1 Introduction . . . 65
4.2 Some Properties of the Logic of Objective Beliefs . . . 66
4.2.1 Decidability and Complexity: . . . 68
4.2.2 Adding Update . . . 72
4.3 Update of Objective Beliefs . . . 75
4.3.1 Syntax and Semantics . . . 75
4.3.2 Axiomatization and Completeness . . . 77
4.3.3 Decidability and Complexity . . . 81
4.4 Arbitrary Update of Objective Beliefs . . . 82
4.4.1 Syntax and Semantics . . . 82
4.4.2 Axiomatization and Completeness . . . 84
4.5 Case Study . . . 89

5 Group Announcement Logic **93**
5.1 Group Announcement Logic . . . 94
5.1.1 Language . . . 94
5.1.2 Semantics . . . 94
5.1.3 Logical Properties . . . 95
5.1.4 Axiomatization . . . 102
5.2 Expressivity . . . 108
5.3 Model Checking . . . 113
5.3.1 Preliminary Results . . . 113
5.3.2 Model Checking Algorithm . . . 114
5.4 Announcements and Ability . . . 118
5.4.1 Singleton Coalitions . . . 118
5.4.2 More Than One Agent . . . 122
5.5 Security Protocols . . . 124

6 Permission and Public Announcements **129**
6.1 The Logic of Permission and Obligation to Speak . . . 131
6.1.1 Syntax . . . 131
6.1.2 Semantics . . . 132
6.1.3 Example: Art School . . . 134
6.1.4 Valid Principles and Other Semantic Results . . . 135
6.2 Axiomatization . . . 141
6.2.1 Soundness . . . 142
6.2.2 Completeness . . . 142
6.3 Decidability . . . 146

6.4 Extended Example: La Belote 150
6.5 Comparison to the Literature 153
6.5.1 Classic Deontic Principles and Paradoxes 153
6.5.2 Deontic Action Logics 155
6.6 Tableaux 157
6.6.1 Definitions 157
6.6.2 Properties 161
6.6.3 Implementation 166

7 Private Permissions 171
7.1 Introduction 171
7.2 Logic of Permitted and Obligatory Private Announcements 172
7.2.1 Introducing Agency and Individual Permission 172
7.2.2 Syntax of $\mathcal{L}_{popral}$ 174
7.2.3 Semantics for $\mathcal{L}_{pral}$ 176
7.2.4 Semantics for $\mathcal{L}_{popral}$ 177
7.2.5 Comparison With the *Non-Agent* Version 179
7.2.6 Properties 182
7.2.7 Soundness and Completeness 185
7.3 Case Study: AIDS 189

8 Conclusion 193
8.1 Dealing with Privacy Policies 194
8.2 Dynamic Epistemic Logic with Permission 197

Bibliography 201

Chapter 0

Introduction en français

0.1 Introduction

De nombreuses situations font intervenir la notion de communication ainsi que des restrictions sur cette communication. C'est le cas lorsque l'on pense à des informations militaires, des communications médicales, des normes morales, des jeux, *etc.* Dans certaines des ces situations, il se peut qu'existent des structures pour penser et organiser le *droit de communiquer*. Dans l'armée, par exemple, une telle structure est assez simple, et facile à comprendre: plus on est haut-placé dans la hiérarchie militaire, plus on a le droit de savoir et moins on a l'autorisation de dire. En effet, un général a accès à de nombreuses informations secrètes sans avoir le droit de les divulguer à ses soldats, alors qu'un soldat peut donner toutes les informations qu'il possède (il se peut même qu'il *doive* les donner) sans avoir accès à de nombreuses autres. Le champ médical est un exemple où des restrictions plus subtiles empêchent un patron d'avoir accès à des données médicales d'un de ses travailleurs, alors qu'un docteur devrait pouvoir y avoir accès. Souvent, ces structures sont présentées sous la forme d'un ensemble de règles informelles, ensemble qui peut être incomplet et même contradictoire, laissant la justice décider ce qu'il convient de faire en cas de conflits.

Mais il n'existe pas de cadre général pour analyser ce genre de situations. L'objectif de ce mémoire est d'apporter quelques éléments, dans le champ de la logique, pour une meilleure compréhension de la notion de 'droit de savoir', éléments qui pourraient nous aider à comprendre et répondre aux problèmes pour lesquels cette notion rentre en jeu. On concentre notre réflexion sur la partie informative de la communication (et non sur sa forme), ce qui amène notre sujet central à la notion de 'droit de donner une information'.

0.1.1 Qu'est-ce que la *logique*?

La logique est l'étude formelle de l'argumentation humaine. En un sens, elle peut être considérée comme l'étude du raisonnement humain (si l'on considère que les arguments traduisent le raisonnement interne inclus dans une communication entre personnes). Son but est d'obtenir des résultats formels (et sans ambiguïtés). Pourtant, le langage naturel (dans lequel sont formés les arguments) est particulièrement ambigü, chaque mot ayant différents sens possibles et chaque concept ayant différentes interprétations dans un même langage. Pour former une théorie logique, il est donc nécessaire de modéliser une partie seulement du raisonnement, en suivant des conditions prédéterminées. C'est ce qui se passe dans le fameux syllogisme suivant, attribué à Aristote[1]: *"Tous les hommes sont mortels. Socrate est un homme. Donc Socrate est mortel."* En effet, il suppose que les notions de mortalité, d'homme, est d''être' sont sans ambiguïtés. Ça pourrait paraître acceptable dans ce cas précis, mais cette autre phrase qui a la même structure, et qui est assez connue également, montrera que ce n'est pas si évident en général: "*Les oiseaux volent. Tweety le manchot est un oiseau. Donc Tweety vole*"[2]. 'Il est inexact que que tous les oiseaux volent' pourrait me rétorquer un lecteur avisé, et nous pourrions être d'accord. Mais le point important est qu'il existe une ambiguïté dans le langage naturel concernant ces concepts: lorsque l'on dit que les oiseaux volent, entend-on 'toujours'? 'Généralement'? 'Dans toute situation normale'? Si un aigle se casse une de ses ailes, est-ce que ça rend inexact le fait que 'les aigles volent'?

Par conséquent, pour créer une théorie logique, nous devons définir un langage exempt de toute ambiguïté, et une interprétation déterministe de ses formules. 'Interpréter une formule' signifie ici 'dire si une formule est vraie ou fausse dans un contexte donné'. Notez que rien ne nous oblige à considérer la valeur de vérité comme une fonction biniare: vrai ou faux. Qui plus est, dans notre conception de la réalité certains concepts ne sont pas binaires: je mesure 1m76, suis-je grand? Certaines diraient que oui, d'autres que non, mais notre compréhension commune nous mènerait plutôt à dire que je suis *assez* grand, mais pas *très* grand. Certaines théories logiques (voir par exemple [Dubois and Prade, 1988]) permettent de considérer ce genre de concepts dont la valeur de vérité est à la fois qualitative et quantitative. Dans cette thèse, tous les concepts (abstraits) que l'on considère ne peuvent être que vrais ou faux (et certainement pas les deux à la fois).

Nous pouvons alors représenter le monde par une liste de tout ce qui est vrai (le reste étant faux). Une telle liste serait impossible à obtenir si l'on veut considérer toutes les propriétés du monde (et combien y en a-t-il?), mais dans des situations données il est possible de se limiter à un nombre fini de propriétés intéressantes et ne considérer que celles-là.

Figure 1: Un modèle booléen des enfants lunatiques
Les propositions suivantes sont vraies dans ce modèle : G_B, G_C

Voyons un exemple: voici quatre enfants, Alex, Brune, Cha et Dan. Nous ne nous

[1]Nous n'avons aucune référence pour affirmer que cette phrase est effectivement d'Aristote, et la notion de syllogisme dans les écrits d'Aristote est probablement plus proche de: '*Si tous les hommes sont mortels et si tous les grecs sont des hommes, alors tous les grecs sont mortels*'.

[2]Dans l'exemple classique, en français, Tweety est souvent un pingouin. Mais il se trouve que les pingouins volent (!) ce qui enlève l'effet escompté de cet exemple

intéressons qu'à leur humeur, considérée comme binaire: ils sont joyeux ou tristes. Par contre elle peut ne pas être statique: de fait, ces enfants sont lunatiques, leur humeur change tout le temps!

Notre langage est basé sur les propositions suivantes: *Alex_est_joyeux* (G_A), *Brune_est_joyeuse* (G_B), *Cha_est_joyeuse* (G_C) et *Dan_est_joyeuse* (G_D)[3]. Mises ensemble, elles forment l'ensemble des *propositions atomiques* du langage, noté $PROP$. Donc, $PROP = \{G_A, G_B, G_C, G_D\}$.

Il est alors possible de représenter le monde réel par une liste des valeurs de vérité (vrai ou faux) des propositions (prises dans l'ensemble $PROP$). La figure 1 donne un exemple d'une telle représentation, appelée *modèle propositionnel booléen.*

Il y a différents mondes possibles, ici exactement seize. Ils sont représentés dans la figure 2.

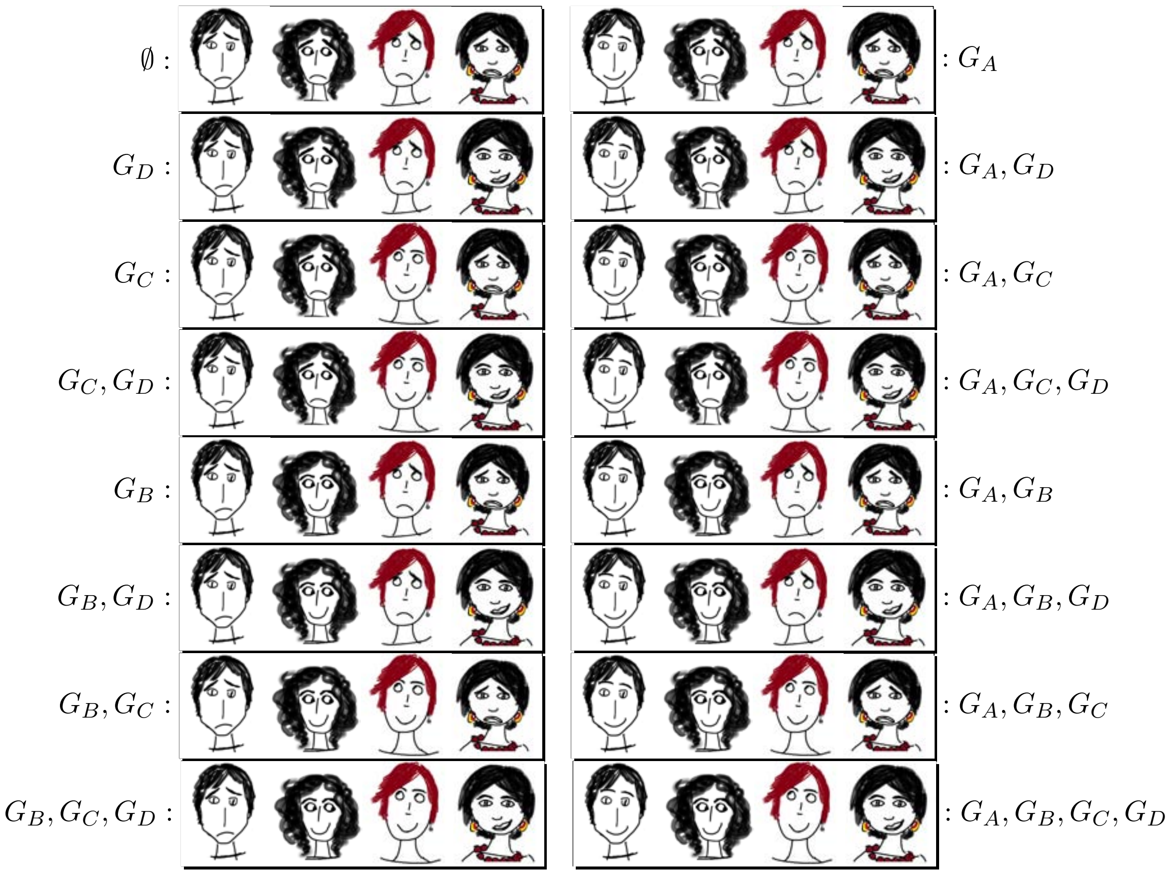


Figure 2: Enfants lunatiques: tous les mondes possibles

Ces mondes possibles sont la base de la représentation du monde réel avec des modalités (qui peuvent être de temps, de croyance, de connaissance, de résultats d'actions *etc.*). De telles représentations sont introduites dans le chapitre suivant, à travers la modalité de connaissance.

---

[3]Le 'G' vient de l'anglais 'Good'

Il y a un double lien entre la logique et l'informatique. D'un côté, les théories informatiques donnent à la logique des résultats techniques importants, comme des algorithmes déterministes qui peuvent prouver qu'une formule est vraie dans un contexte donné, ou dans tout contexte possible. Le temps nécessaire à l'obtention d'une telle réponse, en fonction de la taille de la formule initiale, peut aussi être obtenu. Nous présenterons dans ce travail des résultats théoriques de ce type. Une introduction à ceux-ci est proposée dans la section 0.2.2.

D'un autre côté, les théories logiques peuvent donner aux informaticiens des méthodes utiles pour résoudre des problèmes concrets. Un exemple à la mode est le SUDOKU: un algorithme classique peut être très long à écrire, alors qu'une procédure formalisant dans un langage logique les propriétés qu'il faut satisfaire est très simple à développer.

0.1.2 Aperçu général

Je crois que la recherche scientifique devrait faire un effort permanent pour être accessible au plus grand nombre. Il est clair que tout travail scientifique n'est pas forcément compréhensible par tout le monde, par contre chaque chercheur peut faire tout son possible pour donner des éléments qui rendent au moins une partie de son travail compréhensible à des personnes extérieures à son champ de recherche. Il me semble que c'est particulièrement vrai pour une thèse de doctorat qui synthétise plusieurs années de travail, avec une taille finale non-imposée et qui pourrait être lue par des lecteurs novices (famille, amis...). Le chapitre 2 est donc consacré à la présentation des notions basiques de la logique modale, dans le contexte de l'étude de la connaissance. Certaines de ces notions sont cependant beaucoup plus générales et peuvent être utilisées pour tout type de logique modale. Ce chapitre est traduit intégralement en français ci après. J'espère que ça incitera les lecteurs non initiés à s'intéresser aux éléments basiques de la logique modale.

Des travaux plus développés en logiques épistémique, dynamique et déontique sont présentés dans le chapitre 3. On y situe également notre travail dans le cadre de la recherche actuelle, et on y présente des données nécessaires à la présentation ultérieure de nos travaux. On discute également quelques principes qu'il nous faut suivre pour une bonne compréhension des notions liées au 'droit de savoir'.

Dans un travail de plusieurs mois sur un sujet donné, de nombreuses questions parallèles apparaissent et demandent à être résolues. Les chapitres 4 et 5 présentent les travaux qui ont suivi ce processus. En effet, le chapitre 4 traite du concept de croyance objective, une notion intermédiaire entre la connaissance et la croyance, et présente des résultats techniques qui complètement ceux de [Hommersom et al., 2004]. Quant au chapitre 5, il présente un travail collectif ([Ågotnes et al., 2010]) sur la capacité d'un groupe d'agents à communiquer une information.

Les chapitres 6 et 7 présentent le résultat le plus important de cet essai: une formalisation du 'droit de dire'. Le premier présente cette notion dans le contexte d'une communication publique, *id est* dans des situations où tout échange d'information est public, et pour laquelle les restrictions sur cette communication sont indépendantes de la nature de l'agent qui communique. Dans ce formalisme, il est impossible de déterminer qui est en train de parler, la seule chose qui compte est ce qui est dit. La présentation est basée sur un exemple: la belote.

Le second généralise la première proposition, en donnant une formalisation qui inclut des permissions individuelles et qui considère des communications privées aussi bien que publiques.

Le dernier chapitre conclut ces travaux et ouvre la voie à des perspectives futures. En effet, ce travail est un premier pas dans une voie inachevée qu'il s'agit de poursuivre, en généralisant cette formalisation ou en analysant différentes situations qui utilisent ces concepts.

0.2 Logiques modales pour la représentation de la connaissance

Qu'est-ce que cela veut dire que quelqu'un *sait* quelque chose? Est-il seulement possible que quelque chose soit *su*? Ces questions ne sont pas nouvelles, elles ont été étudiées au moins depuis les philosophes grecs (voir [Plato, BC]) et forment le champ de l'Epistémologie, l'étude de la connaissance. Plusieurs siècles après Platon, [Hintikka, 1962] a proposé une analyse logique formelle de la connaissance dans un contexte multi-agent. Son formalisme, comme nous allons le voir, utilise la sémantique des mondes possibles. Depuis lors, des logiques épistémiques ont été utilisées dans de nombreux champs d'étude, comme l'intelligence artificielle, l'économie, la linguïstique ou l'informatique théorique, en se concentrant sur les aspects multi-agents (donc sur l'interaction entre agents, qui peuvent être des êtres humains ou des systèmes informatiques) bien plus que sur la compréhension philosophique de la connaissance.

Il est possible dans ce formalisme de raisonner sur ce que l'on sait, sur ce qu'un agent sait de la connaissance d'un autre, sur ce qui constitue l'ensemble des connaissances partagées par les agents. Mais comment le formalisme de Hintikka représente-t-il cette connaissance?

0.2.1 Représentation de la connaissance

Langage de la logique épistémique

Tout d'abord, il nous faut définir proprement notre langage de la *logique épistémique*, noté $\mathcal{L}_{el}$, en partant d'un ensemble dénombrable d'agents AG et d'un ensemble dénombrable d'atomes propositionnels $PROP$. Dans l'exemple présenté dans le premier chapitre, on considère $AG = \{a, b, c, d\}$ pour *Alex*, *Brune*, *Cha* et *Dan*, et $PROP = \{G_A, G_B, G_C, G_D\}$. Voici alors quelques formules exprimables dans notre langage:

$K_b(G_A)$: "Brune sait qu'Alex est joyeux"

$(G_C) \longrightarrow K_c(G_C)$: 'Si Cha est joyeuse, alors elle le sait"

$K_b(G_D \vee \neg G_D)$: "Brune sait que Dan est joyeuse ou triste"

$\neg G_C \wedge K_a(G_C)$: "Cha est triste est Alex sait qu'elle est joyeuse".

Plus formellement, voici comment les formules du langage sont construites:

Definition 0.1 (Le Langage $\mathcal{L}_{el}$) *L'ensemble $\mathcal{L}_{el}(AG, PROP)$ de formules épistémiques est obtenu à partir de AG et $PROP$ en itérant indéfiniment les opérations suivantes:*

- *pour tout $p \in PROP$, p est une formule,*
- *$\bot$ ("faux") est une formule,*
- *si φ est une formule alors $\neg\varphi$ ("non φ") est une formule,*
- *si φ est une formule et si ψ est une formule alors $(\varphi \vee \psi)$ ("φ ou ψ") est une formule,*
- *si φ est une formule alors pour tout agent $a \in AG$, $K_i\varphi$ ("i sait que φ") est une formule.*

Dans le cas où les ensembles d'atomes ($PROP$) et d'agents (AG) sont clairs nous les omettons. Cette définition peut être écrite de façon plus concise de la manière suivante[4]:

Definition 0.2 (Le langage $\mathcal{L}_{el}$) *Le langage $\mathcal{L}_{el}$ basé sur un ensemble dénombrable d'agents AG et sur un ensemble dénombrable d'atomes propositionnels $PROP$ est défini de la façon suivante:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid (\varphi_1 \vee \varphi_2) \mid K_i\varphi$$

où $i \in AG$ et $p \in PROP$.

On ajoute les abréviations suivantes:

- $\top$ ("vrai") est une abréviation de $\neg\bot$
- $(\varphi \wedge \psi)$ ("φ et ψ") est une abréviation de $\neg(\neg\varphi \vee \neg\psi)$
- $(\varphi \longrightarrow \psi)$ ("φ implique ψ") est une abréviation de $(\neg\varphi \vee \psi)$
- $(\varphi \longleftrightarrow \psi)$ ("φ est équivalent à ψ") est une abréviation de $((\varphi \longrightarrow \psi) \wedge (\psi \longrightarrow \varphi))$
- $\hat{K}_i\varphi$ ("i envisage φ") est une abréviation de $\neg K_i\neg\varphi$. On dit que $\hat{K}_i$ est le dual de K_i.

Comme indiqué plus haut, $\neg G_C \wedge K_a(G_C)$ (lire "Cha est triste et Alex sait qu'elle est joyeuse") est une formule du langage. Ceci explicite le fait que toute les formules appartenant au langage ne sont pas forcément intuitivement vraies. Mais personne n'a dit que toutes les formules exprimables étaient vraies. De fait, on n'a pas pour l'instant défini comment évaluer la valeur de vérité d'une formule épistémique. Faisons-le maintenant.

[4]Cette description formelle d'un langage est inspirée par la grammaire classique en informatique baptisée Backus-Naur Form (BNF). On l'utilise dans cet essai pour décrire des langages formels.

Sémantique *des mondes possibles*

D'abord, nous considérons qu'il existe une interprétation objective du monde *réel*, indépendamment de qui le regarde. Cette interprétation est une liste des valeurs de vérité de tous les faits objectifs dans l'état courant. Si nous appelons *propositions* ces faits objectifs, on comprend aisément que cette représentation du monde n'est autre qu'un modèle propositionnel Booléen, tel qu'introduit dans le chapitre 0.1. Dans notre exemple, il s'agit d'une liste des humeurs de tous les enfants.

Le manque de connaissance peut alors être vu comme un doute sur lequel des états possibles est l'état courant. Hintikka représente le *monde épistémique* (c'est à dire le monde et la connaissance de chaque agent) par un graphe où les noeuds sont des représentations de mondes possibles (donc des modèles propositionnels) et une arrête, idexée par un agent i, représente le fait que l'agent i *ne sait pas* si l'état courant est l'un ou l'autre des mondes reliés par cette arrête. Réciproquement, on dit que i *sait* une assertion φ si φ est vraie dans tous les états reliés à l'état courant par une arrête indexée par i. Voici une représentation d'une situation dans laquelle Brune ne connaît pas l'humeur de Cha:

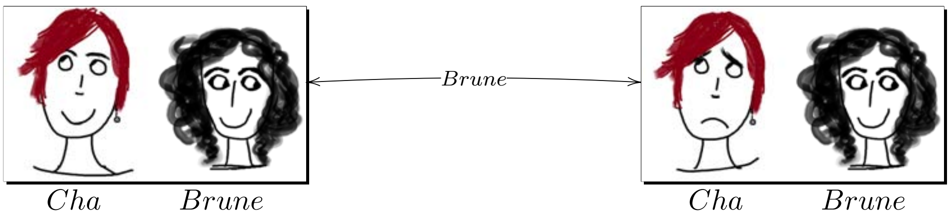


Figure 3: Modèle épistémique

La figure 4 donne une représentation plus complète de ce genre de situations épistémiques: Alex connaît sa propre humeur mais ne connaît pas celle de Dan, et il sait que Dan connaît sa propre humeur mais pas la sienne à lui. Et Dan est consciente de ça, *etc.*

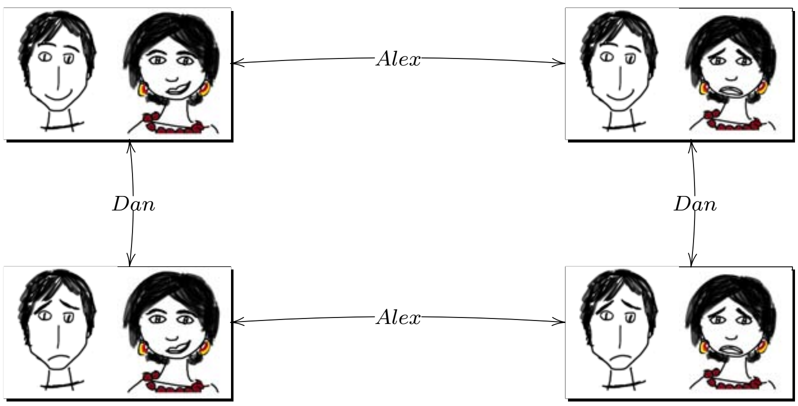


Figure 4: Un autre modèle épistémique

On omet dans cette figure les arrêtes réflexives (celles qui pointent un monde vers lui-même) qui représentent le fait que les enfants envisagent le monde réel comme une possibilité, ce qui est considéré toujours vrai en logique épistémique.

Avant de définir ces notions proprement, voici une définition plus précise de ce qu'est un modèle:

Definition 0.3 (Modèle de Kripke) *Etant donné un ensemble dénombrable d'agents AG et un ensemble dénombrable d'atomes propositionnels $PROP$, un modèle de Kripke est un tuple $\mathcal{M} = (S, \mathcal{R}, V)$ où:*

- *S est un ensemble dont les éléments sont appelés "mondes" ou "états",*
- *$V : PROP \longrightarrow 2^S$ est une fonction de valuation qui attribue à chaque proposition p l'ensemble $V(p)$ des mondes dans lesquels p est considérée vraie, et*
- *$\mathcal{R} = \{R_i\}_{i \in AG}$ avec pour tout $i \in AG$, $R_i \subseteq S \times S$ est une fonction binaire sur S.*

On appelle modèle pointé *un modèle de Kripke $\mathcal{M}, s$ accompagné d'un de ses états.*

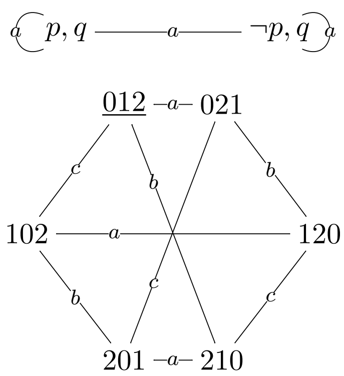


Figure 5: Quelques exemples de modèles de Kripke

La figure 5 donne deux représentations plus classiques de situations épistémiques à l'aide de modèles de Kripke. Le premier modèle représente le doute d'un agent a concernant p alors que q est su par l'agent. Le second, explicité (en anglais) à la page 45, est une représentation de l'état épistémique suite à la distribution de trois cartes 0, 1 et 2 parmi trois joueurs a,b et c. Les arrêtes réflexives sont omises une fois de plus dans ce second dessin.

Ces modèles nous permettent d'interpréter des phrases qui traitent de la vérité d'un fait objectif, de la connaissance qu'ont les agents à propos de ces faits, et de la connaissance des agents concernant ce genre de phrases.

Definition 0.4 (Relation de satisfaisabilité pour $\mathcal{L}_{el}$) *Soit $\mathcal{M}$ un modèle. On définit la relation de satisfaisabilité $\models : S \times \mathcal{L}_{el} \longrightarrow \{0, 1\}$ par récurrence sur la structure de φ[5] de la façon suivante:*
(On note $\mathcal{M}, s \models \varphi$, qui se lit "$\varphi$ est vraie dans l'état s du modèle $\mathcal{M}$", si $\models (s, \varphi) = 1$ et $\mathcal{M}, s \not\models \varphi$, qui se lit "$\varphi$ est faux dans l'état s du modèle $\mathcal{M}$", si $\models (s, \varphi) = 0$)

[5] Cette notion est présentée en détail dans la remarque 0.7

- *pour tout* $s \in S$, $\mathcal{M}, s \models p$ *ssi* $s \in V(p)$
- *pour tout* $s \in S$, $\mathcal{M}, s \not\models \bot$
- *pour tout* $s \in S$, $\mathcal{M}, s \models \neg\psi$ *ssi* $\mathcal{M}, s \not\models \psi$
- *pour tout* $s \in S$, $\mathcal{M}, s \models \psi_1 \vee \psi_2$ *ssi* $(\mathcal{M}, s \models \psi_1$ *ou* $\mathcal{M}, s \models \psi_2)$
- *pour tout* $s \in S$, $\mathcal{M}, s \models K_i\psi$ *ssi pour tout* $t \in S$ *tel que* sR_it, $\mathcal{M}, t \models \psi$

On dit que φ *est valide dans le modèle* $\mathcal{M}$*, et on note* $\mathcal{M} \models \varphi$*, si pour tout* $s \in S$, $\mathcal{M}, s \models \varphi$*. On dit que* φ *est valide si pour tout modèle* $\mathcal{M}$, $\mathcal{M} \models \varphi$*, c'est à dire si* φ *est valide dans tout état de tout modèle. Enfin, on note* $[\![\varphi]\!]_{\mathcal{M}}$ *l'ensemble des mondes* s *du modèle* $\mathcal{M}$ *tels que* $\mathcal{M}, s \models \varphi$*.*

Caractérisation de la connaissance

Nous avons affirmé, dans la définition 0.1, que l'on pouvait interpréter $K_i\varphi$ par "l'agent i sait φ". Comme nous l'avons expliqué en introduction de ce chapitre, cette affirmation, pour être raisonnable, doit être suivie d'arguments qui rendent la sémantique appropriée à l'interprétation de la connaissance. Examinons les validités données par la sémantique, et celles que l'on devrait s'assurer d'avoir pour représenter une conception, même idéalisée, de la connaissance.

Avant tout, la sémantique de Kripke, présentée dans la définition 0.4, impose que nos agents, qui sont capables de savoir, ont une capacité de déduction sans limite. Pourquoi? Supposez que, dans un état donné d'un modèle donné, un agent i sache que ψ est vrai et que ψ implique φ. Alors les formules φ et $\psi \longrightarrow \varphi$ sont satisfaites dans tout état envisagé par i, et donc φ y est satisfaite aussi (car cet état est un modèle propositionnel). En d'autres termes, les formules suivantes sont valides pour tous ψ, φ dans le langage et tout agent i

$$(K_i\psi \wedge K_i(\psi \longrightarrow \varphi)) \longrightarrow K_i\varphi. \tag{K}$$

Ça pourrait sembler simuler raisonnablement la capacité de déduction d'un agent rationnel. Mais elle implique, par exemple, que tout agent "connaît" toutes les tautologies, c'est à dire toutes les phrases qui sont toujours vraies. Or, bien que vous soyez probablement rationnel-le-, pouvez vous affirmer "savoir" que la formule suivante est tautologique?

$$(((p \vee t) \to s) \wedge (q \leftrightarrow u)) \vee ((p \vee t \wedge \neg s) \wedge ((q \wedge u) \vee (\neg q \wedge \neg u))) \longrightarrow (q \vee (v \wedge u \wedge (v \longrightarrow \bot)))$$

Il est aussi largement accepté que si un agent sait quelque chose, alors cette chose est vraie. Autrement dit, les formules suivantes sont valides:

$$K_i\varphi \longrightarrow \varphi. \tag{T}$$

Il s'agit là d'une différence importante entre "connaissance" et "croyance", bien que nous ne puissions pas résumer la connaissance à de la croyance vraie (voir [Burnyeat and Barnes, 1980]). Les logiques de la croyance nient souvent ce principe de vérité de la croyance, et en assument un plus faible: la cohérence. En effet, on considère alors que si un agent croit quelque chose, il ne croit pas en même temps sa négation. Cela pourrait être traduit dans l'un de ces principes équivalents entre eux (et qui restent vrais dans le cas de la connaissance):

$$K_i\varphi \longrightarrow \neg K_i\neg\varphi \quad ; \quad \neg K_i\bot \tag{D}$$

On attribue également à la connaissance une introspection positive et négative. Autrement dit, on considère que si un agent sait quelque chose alors il sait qu'il le sait et, ce qui est plus fort encore, que s'il ignore quelque chose alors il sait qu'il l'ignore. C'est là une supposition très forte: savez-vous réellement quelle est l'ensemble de vos connaissances? Et pouvez vous énoncer la liste de tout ce que vous ignorez? Si l'on accepte ces propriétés, alors on accepte la validités des formules suivantes, pour tout φ dans le langage:

$$K_i\varphi \longrightarrow K_iK_i\varphi \tag{4}$$

$$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi. \tag{5}$$

Une dernière notion qu'il nous faut introduire est celle de *connaissance commune*. Alex et Cha sont des habitués du Poker (comme vous le verrez, nos enfants lunatiques aiment jouer aux cartes). Alex connaît les règles du jeu. Il sait aussi que Cha connaît les règles – sans quoi il serait tenté de tricher. Mais il sait également que Cha sait qu'il connaît les règles – et il peut donc supposer qu'elle ne tentera pas de tricher. On pourrait continuer à faire des phrases de ce type... En fait, les règles du jeu sont connaissance commune.

Plus formellement, la connaissance commune de φ est l'abréviation syntaxique d'une conjonction infinie de formules. Soit G un ensemble d'agents, alors la connaissance commune par les agents de G de la formule φ est:

$$CK_G\varphi := \bigwedge_{n\in\mathbb{N}} \bigwedge_{i_1,\ldots,i_n\in G} K_{i_1}\ldots K_{i_n}\varphi$$

Comme nous le verrons, cette notion est très importante lorsque l'on considère un apprentissage collectif: si Brune et Alex apprennent quelque chose ensemble, et s'ils peuvent voir que cet apprentissage est mutuel, alors l'information apprise devient connaissance commune. Voir [van Ditmarsch et al., 2009] pour plus de détails.

0.2.2 Notions techinques classiques en logique modale

Il se peut que ce chapitre soit plus difficile à comprendre pour un lecteur novice, et qu'il soit parfaitement redondant pour un expert. Mais il semble important de définir et d'expliquer correctement les notions d'informatique qui sont utiles en logique. Ces notions ne se limitent pas à l'étude de la connaissance, au contraire la majorité d'entre elles sont communes à tous les champs de la logique modale. Nous les présenterons toutefois en utilisant le langage et la sémantique de la logique épistémique.

Propriétés du langage

Commençons par des notions basiques concernant la syntaxe de langages logiques.

Definition 0.5 (Taille d'une formule) *Etant donnée une formule φ d'un langage $\mathcal{L}$ on appelle taille de φ, noté $|\varphi|$, le nombre de symboles qui constituent φ.*

Definition 0.6 (Sous-formule) *Pour toute formule $\varphi \in \mathcal{L}_{el}$ on définit $Sub(\varphi)$ l'ensemble des sous-formules de φ en fonction de la forme de la formule φ:*

- $Sub(p) = \{p\}$
- $Sub(\bot) = \{\bot\}$
- $Sub(\neg\psi) = \{\neg\psi\} \cup Sub(\psi)$
- $Sub(\varphi_1 \vee \varphi_2) = \{\varphi_1 \vee \varphi_2\} \cup Sub(\psi_1) \cup Sub(\varphi_2)$
- $Sub(K_a\psi) = \{K_a\psi\} \cup Sub(\psi)$.

Si $\psi \in Sub(\varphi)$ on dit que ψ est une sous-formule de φ.

On peut prouver que $Sub(\varphi)$ est bien définie par récurrence sur la taille de φ.

Remark 0.7 (Récurrecne "sur la structure de φ") *Dorénanvant, "prouver* (resp. définir) *une propriété $P(\varphi)$ par récurrence sur la structure de φ" signifie "prouver* (resp. définir) *$P(\psi)$ pour tout $\psi \in PROP \cup \{\bot\}$ et prouver* (resp. définir) *$P(\varphi)$ en admettant l'hypothèse de récurrence suivante: $P(\psi)$ est vraie* (resp. définie) *pour toute sous-formule ψ de φ". Il s'agit d'une récurrence à travers l'ordre partiel 'être sous formule de'.*

Le langage que l'on étudie ici ne peut exprimer qu'un nombre limité de notions. En ajoutant un opérateur modal (donc un nouveau symbole) à un langage donné, sans changer la sémantique des symboles précédemment introduits, le langage obtenu peut clairement exprimer *au moins* ce que pouvait exprimer l'ancien, et peut-être plus. Précisons ce concept d'expressivité d'un langage.

Definition 0.8 (Expressivité d'un langage) *Etant donnés deux langages $\mathcal{L}_1$ et $\mathcal{L}_2$ et une classe de modèles $\mathcal{C}$, $\mathcal{L}_1$ est* au moins aussi expressif *que $\mathcal{L}_2$ par rapport à $\mathcal{C}$ ssi pour toute formule φ de $\mathcal{L}_1$ il existe une formule ψ de $\mathcal{L}_2$ qui lui soit équivalente. Autrement dit, pour tout modèle $\mathcal{M}$ de $\mathcal{C}$, $[\![\varphi]\!]_{\mathcal{M}} = [\![\psi]\!]_{\mathcal{M}}$: le domaine de satisfaisabilité de φ dans $\mathcal{M}$ en considérant la* *sémantique de* $\mathcal{L}_1$ *est le même que le domaine de satisfaisabilité de* ψ *dans* $\mathcal{M}$ *en considérant la sémantique de* $\mathcal{L}_2$*.*

Voici deux façons standards de déterminer que $\mathcal{L}_1$ *est* au moins aussi expressif *que* $\mathcal{L}_2$*:*

- $\mathcal{L}_2$ *forme un sous-langage de* $\mathcal{L}_1$
- *il existe une traduction telle que toute formule de* $\mathcal{L}_2$ *est logiquement équivalente à sa traduction dans* $\mathcal{L}_1$*.*

Le langage $\mathcal{L}_1$ *est dit* plus expressif *que* $\mathcal{L}_2$ *par rapport à* $\mathcal{C}$ *si* $\mathcal{L}_1$ *est au moins aussi expressif que* $\mathcal{L}_2$ *et* $\mathcal{L}_2$ *n'est pas au moins aussi expressif que* $\mathcal{L}_1$ *(cette notion est un ordre partiel).*

Une manière standard de déterminer que $\mathcal{L}_2$ *n'est pas au moins aussi expressif que* $\mathcal{L}_1$ *est de mettre en évidence une formule* φ *de* $\mathcal{L}_1$ *et deux modèles de* $\mathcal{C}$ $(\mathcal{M}, s)$ *et* $(\mathcal{M}', s')$ *tels que* φ *est vraie dans* $(\mathcal{M}, s)$ *et fausse dans* $(\mathcal{M}', s')$*, alors que toute formule* ψ *de* $\mathcal{L}_2$ *est vraie dans* $(\mathcal{M}, s)$ *ssi* ψ *est vraie dans* $(\mathcal{M}', s')$*. On dit dans ce cas que le langage* $\mathcal{L}_1$*, mais pas le langage* $\mathcal{L}_2$*, peut* distinguer *les modèles* $(\mathcal{M}, s)$ *et* $(\mathcal{M}', s')$*.*

Une dernière définition utile concernant les langages:

Definition 0.9 (Substitution) *Soit* $\mathcal{L}(PROP)$ *un langage récursivement énumérable basé sur un ensemble dénombrable d'atomes propositionnels* $PROP$*, soient* $\varphi, \psi, \psi_1, \psi_2, \ldots \in \mathcal{L}(PROP)$ *et soient* $p, p_1, p_2, \ldots \in PROP$

- *On note* $\varphi(\psi/p)$ *la formule de* $\mathcal{L}$ *obtenue en remplaçant dans* φ *toute occurence de* p *par* ψ*.*
- *On étend la notation précédente à la substitution simultanée d'une suite (finie ou infinie)* $p_1, p_2, \ldots$*:* $\varphi(\psi_1/p_1, \psi_2/p2, ...)$

Propriétés des modèles

Voyons quelques propriétés sémantiques de la logique modale, autrement dit quelques propriétés des modèles que l'on considère. D'abord, tous les modèles considérés ici sont des modèles de Kripke, auxquels vient s'ajouter évenutellement une relation supplémentaire (qui peut être entre des mondes et des ensembles de mondes, ou bien entre des mondes et des relations). Rappelons que de tels modèles, présentés dans la définition 0.3, sont composés d'un ensemble d'états, de relations binaires sur cet ensemble et d'une valuation qui attribue à chaque proposition un sous-ensemble de mondes (ceux où la proposition est considérée vraie). On peut donc les voir comme des graphes orientés ayant pour noeuds des modèles booléens (c'est à dire une liste des valeurs de vérité des différents atomes propositionnels). Rappelons également que cette définition impose la validité de la formule K pour tout opérateur qui suit la sémantique présentée dans la définition 0.4. Un *sous-modèle* d'un modèle $\mathcal{M}$ donné est composé d'un sous-ensemble de l'ensemble d'états de $\mathcal{M}$, et d'une structure qui est la restriction de la structure initiale sur ce sous-ensemble. Plus formellement:

Definition 0.10 (Sous-modèle) *Soit $\mathcal{M} = (S, V, \langle R_i \rangle_{i \in AG})$ un modèle de Kripke. On dit alors que le modèle $\mathcal{M}' = (S', V', \langle R'_i \rangle_{i \in AG})$ est un* sous-modèle *de $\mathcal{M}$ s'il satisfait les conditions suivantes:*

- $S' \subseteq S$
- *pour tout $p \in PROP$ et tout $s' \in S'$, $s' \in V'(p)$ ssi $s' \in V(p)$*
- *pour tout $i \in AG$ et tous $(s_1, s_2) \in S' \times S'$, $s_1 R'_i s_2$ ssi $s_1 R_i s_2$*

Rappelons la notion de clôture transitive d'un ensemble de relations, dans le contexte des modèles de Kripke.

Definition 0.11 (Clôture transitive) *Soit $\mathcal{R} = \{R_i\}_{i \in AG}$ un ensemble de relations binaires sur un ensemble donné S. On appelle* clôture transitive *de $\mathcal{R}$ la relation binaire $\mathcal{R}^*$ telle que pour tous $s, s' \in S$ il existe $n \in \mathbb{N}$ et $s_0, s_1, \ldots, s_n \in S$ tels que:*

- *$s_0 = s$ et $s_n = s'$*
- *pour tout $k \in \{0, \ldots, n-1\}$ il existe $i \in AG$ tel que $s_k R_i s_{k+1}$.*

On peut constater qu'un modèle peut ne pas être *connexe*, c'est à dire qu'il peut arriver qu'un sous-ensemble d'états du modèle n'ait aucune relation avec un autre. Dans ces conditions, un sous-modèle particulier peut se révéler utile:

Definition 0.12 (Composante connexe - sous-modèle engendré) *Soit $\mathcal{M} = (S, V, \mathcal{R})$ un modèle et $s \in S$. On appelle* composante connexe induite par s dans $\mathcal{M}$ *l'ensemble $S' = \{s' \in S \mid s\mathcal{R}^* s'\}$. On appelle* sous-modèle engendré de $\mathcal{M}, s$ *le sous-modèle $\mathcal{M}'$ de $\mathcal{M}$ basé sur la composante connexe induite par s dans $\mathcal{M}$.*

La composante connexe de s dans $\mathcal{M}$ est donc l'ensemble des états qui sont reliés à s dans le modèle. Cette notion est utile car le sous-modèle engendré d'un modèle pointé $\mathcal{M}, s$ est équivalent à $\mathcal{M}, s$ par rapport au langage $\mathcal{L}_{el}$: une formule qui est vraie dans l'un est aussi vraie dans l'autre. C'est ce qu'affirme la proposition 0.15, en disant que les deux modèles sont bisimilaires.

La *bisimulation* est une notion classique, en logique modale, de similarités entre structures (voir [Blackburn et al., 2001]). On l'utilise souvent dans cet essai, sur des exemples ou pour des preuves. Présentons-la en détails:

Definition 0.13 (Bisimulation) *Soient deux modèles $\mathcal{M} = (S, \mathcal{R}, V)$ et $\mathcal{M}' = (S', \mathcal{R}', V')$. Une relation non-vide $\mathfrak{R} \subseteq S \times S'$ est appelée bisimulation entre $\mathcal{M}$ et $\mathcal{M}'$ si pour tous $s \in S$ et $s' \in S'$ tels que $(s, s') \in \mathfrak{R}$ on a :*

atoms *pour tout* $p \in PROP$*:* $s \in V(p)$ *ssi* $s' \in V'(p)$*;*

forth *pour tout* $i \in AG$ *et tout* $t \in S$*: si* sR_it *alors il existe un* $t' \in S'$ *tel que* $s'R'_it'$ *et* $(t,t') \in \mathfrak{R}$*;*

back *pour tout* $i \in AG$ *et tout* $t' \in S'$*: si* $s'R'_it'$ *alors il existe un* $t \in S$ *tel que* sR_it *et* $(t,t') \in \mathfrak{R}$*.*

On note $(\mathcal{M}, s) \leftrightarrows (\mathcal{M}', s')$ *ssi il existe une bisimulation entre* $\mathcal{M}$ *et* $\mathcal{M}'$ *reliant* s *et* s'*, et on dit alors que les structures de Kripke pointées* $(\mathcal{M}, s)$ *et* $(\mathcal{M}', s')$ *sont bisimilaires.*

Notons que la bisimulation est une relation d'équivalence. C'est une notion importante car elle caractérise le fait que deux modèles sont modalement équivalent, c'est à dire qu'ils satisfont les mêmes formules de $\mathcal{L}_{el}$:

Proposition 0.14 *Soient deux modèles* $\mathcal{M} = (S, \mathcal{R}, V)$ *et* $\mathcal{M}' = (S', \mathcal{R}', V')$*, et soit une formule* $\varphi \in \mathcal{L}_{el}$*. Pour tout* $s \in S$ *et tout* $s' \in S'$*, si* $(\mathcal{M}, s) \leftrightarrows (\mathcal{M}', s')$ *alors* $\mathcal{M}, s \models \varphi$ *ssi* $\mathcal{M}', s' \models \varphi$*.*

La preuve de cette proposition apparaît par exemple dans [Fagin et al., 1995]. En particulier, la proposition suivante implique qu'un modèle satisfait les mêmes formules que son sous-modèle engendré.

Proposition 0.15 *Soit* $\mathcal{M}, s_0$ *un modèle pointé. Il est bisimilaire à son sous-modèle engendré.*

Proof Soit $\mathcal{M}' = (S', \mathcal{R}', V')$ le sous-modèle engendré de $\mathcal{M}, s_0$. Soit $\mathfrak{R}$ la relation binaire entre S et S' définie de la manière suivante: $s\mathfrak{R}s'$ ssi $s = s'$.

On va montrer que $\mathfrak{R}$ est une bisimulation entre $\mathcal{M}, s_0$ et $\mathcal{M}', s_0$. D'abord, il est clair que $s_0\mathfrak{R}s_0$. Pour tout $s \in S'$ on a que

atoms pour tout $p \in PROP$: $s \in V(p)$ ssi $s \in V'(p)$ (par la définition 0.10);

forth pour tout $i \in AG$ et tout $t \in S$: si sR_it, alors $t \in S'$ et sR'_it par la définition 0.10, et $t\mathfrak{R}t$;

back pour tout $i \in AG$ et tout $t \in S'$: si sR_it, alors sR_it par la définition 0.10, et $t\mathfrak{R}t$.

□

Introduisons maintenant un autre type de relation d'équivalence, qui est une sorte de généralisation de la bisimulation. L'idée est de considérer comme équivalents deux états d'un modèle donné qui satisfont les formules d'un sous-ensemble donné du langage. Nous obtenons la notion de filtration:

Definition 0.16 (Filtration) *Soit $\mathcal{M} = (S, \sim_i, V)$ un modèle et Γ un ensemble de formules clos pour la sous-formule (* id est *si une formule appartient à l'ensemble, toutes ses sous-formules y apparaissent également). Soit $\leftrightsquigarrow_\Gamma$ la relation binaire sur S définie, pour tous $s, t \in S$, par:*

$$s \leftrightsquigarrow_\Gamma t \quad \textit{ssi} \quad \textit{pour tout} \quad \varphi \in \Gamma : (\mathcal{M}, s \models \varphi \textit{ ssi } \mathcal{M}, t \models \varphi)$$

Notons que $\leftrightsquigarrow_\Gamma$ est une relation d'équivalence. On appelle filtration de $\mathcal{M}$ à travers Γ *(ou simplement* filtration de $\mathcal{M}$*) le modèle $\mathcal{M}^\Gamma = (S^\Gamma, \sim_i^\Gamma, V^\Gamma)$ où:*

- $S^\Gamma = S/_{\leftrightsquigarrow_\Gamma}$
- *pour tous $|s|, |t| \in S^\Gamma$, $|s| \sim_i^\Gamma |t|$ ssi pour tout $K_i\varphi \in \Gamma$, $(\mathcal{M}, s \models K_i\varphi$ iff $\mathcal{M}, t \models K_i\varphi)$*
- $V^\Gamma(p) = \begin{cases} \emptyset \textit{ si } p \notin \Gamma \\ V(p)/_{\leftrightsquigarrow_\Gamma} \textit{ si } p \in \Gamma) \end{cases}$

Une dernière remarque qui a son importance: dans la classe de tous les modèles de Kripke, il se peut que des sous-classes particulières soient utiles. On les définit en fonction des propriétés de ses relations binaires (réflexivité, transitivité, symétrie, sérialité, euclidianité, équivalence). Rappelons qu'une relation binaire R sur un ensemble S est dite:

- réflexive si pour tout $s \in S$, $(s, s) \in R$
- transitive si pour tous $s, t, u \in S$, $((s, t) \in R$ et $(t, u) \in R)$ implique que $(s, u) \in R)$
- symétrique si pour tous $s, t \in S$, $((s, t) \in R$ implique que $(t, s) \in R)$
- sérielle si pour tout $s \in S$, il existe $t \in S$ tel que $(s, t) \in R$
- euclidienne si pour tous $s, t, u \in S$, $((s, t) \in R$ et $(s, u) \in R)$ implique que $(t, u) \in R)$
- d'équivalence si elle est réflexive, transitive et symétrique.

Comme nous le verrons dans le paragraphe suivant, ces propriétés des modèles correspondent à des propriétés de la sémantique que l'on a présenté au paragraphe 0.2.1. Plus précisément:

- L'axiome de vérité (noté T) correspond à la réflexivité
- L'axiome de cohérence (noté D) correspond à la sérialité
- L'introspection positive (notée 4) correspond à la transitivité
- L'introspection négative (notée 5) correspond à l'euclidianité

Rappelons que tous les modèles satisfont l'axiome noté K. On appelle donc KT (resp. KD, $K45$, *etc.*) la classe des modèles reflexifs (resp. seriels, transitifs et euclidiens, *etc.*). $S5$ est une abréviation de $KT45$ et correspond à la classe des modèles pour lesquels les relations binaires sont des relations d'équivalence. Pour tout $n \in \mathbb{N}$ et pour toute classe de modèles $\mathcal{C}$, on note $\mathcal{C}_n$ la classe de modèles qui appartiennent à $\mathcal{C}$ et contiennent exactement n relations binaires.

En particulier, $S5_n$ est la classe des modèles de Kripke qui ont n relation qui sont toutes d'équivalence:

Definition 0.17 (Modèle épistémique) *Un modèle de Kripke $\mathcal{M} = (S, \{R_i\}_{i \in AG}, V)$ est dit* épistemique *si pour tout $i \in AG$, R_i est une relation d'équivalence sur S.*

Axiomatisation

Etant donnée une syntaxe (*id est* une langage), une classe de modèle où on souhaite l'interpréter (c'est à dire une classe de situations, un contexte), on aimerait pouvoir caractériser les formules qui sont vraies dans ce contexte. Autrement dit, quelles sont les propriétés d'une formule qui garantissent qu'elle va être vraie dans n'importe quelle situation d'un contexte donné? La question n'est pas seulement d'être capable de déterminer quelles sont les formules valides, mais aussi d'avoir une justification du fait qu'elles le sont, une preuve.

La notion d'*axiomatisation* a été développée dans ce but. Informellement, une axiomatisation est une description finie de schémas d'axiomes (considérés comme *théorèmes*, donc prouvés) et de règles qui permettent de déduire des nouveaux théorèmes à partir d'anciens. Plus précisément, une axiomatisation $\mathcal{A}$ est un ensemble de schémas d'axiomes (toute formule ayant la même structure que le schéma est un axiome et est donc –par principe– un théorème) et un ensemble de règles (appelées *règles d'inférences*).

$\top$	Vérité
$(A \wedge B) \rightarrow A \;\; ; \;\; (A \wedge B) \rightarrow B$	Simplification
$A \rightarrow (A \vee B) \;\; ; \;\; B \rightarrow (A \vee B)$	Addition
$A \rightarrow (B \rightarrow A)$	Conservation
$(A \rightarrow B) \rightarrow ((A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow C))$	Syllogisme hypothétique (SH)
$(A \rightarrow B) \rightarrow ((A \rightarrow C) \rightarrow (A \rightarrow (B \wedge C)))$	Composition
$(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow ((A \vee B) \rightarrow C))$	Disjonction
$\neg\neg A \rightarrow A$	Tiers exclus
$(A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A)$	Cohérence
A partir de A et de $A \rightarrow B$, déduire B	Modus Ponens (MP)

Table 1: Règle et axiomes de la logique propositionnelle

La table 1 est un exemple d'axiomatisation avec neuf schémas d'axiomes et une règle d'inférence (le modus ponens). Elle axiomatise la logique propositionnelle. On appelle *preuve de φ* une séquence finie de formules $\{\psi_1, \ldots, \psi_n\}$ telle que $\psi_n = \varphi$ et pour tout $i \in \{1, \ldots, n\}$ ou bien ψ_i est une instance d'un schéma d'axiome ou bien elle est obtenue à partir de $\{\psi_1, \ldots, \psi_{i-1}\}$ en utilisant une règle d'inférence. S'il existe une preuve de φ on dit que φ est un théorème et l'on note $\vdash_{\mathcal{A}} \varphi$, ou simplement $\vdash \varphi$. Plus généralement, si φ peut être prouvée en ajoutant à $\mathcal{A}$ un ensemble S de formules considérées comme des axiomes supplémentaires, on note alors $S \vdash_{\mathcal{A}} \varphi$.

Notons qu'une preuve peut être longue. La table 2 donne un exemple de preuve d'un théorème de la logique propositionnelle en utilisant l'axiomatisation $\mathcal{A}$.

$L1:$	$\vdash (\psi \wedge \varphi) \rightarrow \psi$	*Simplification*
$L2:$	$\vdash (\psi \wedge \varphi) \rightarrow \varphi$	*Simplification*
$L3:$	$\vdash \psi \rightarrow (\psi \vee \theta)$	*Addition*
$L4:$	$\vdash \varphi \rightarrow (\varphi \vee \theta)$	*Addition*
$L5:$	$\vdash (\psi \rightarrow (\psi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta)))$	*Conservation*
$L6:$	$\vdash (\varphi \rightarrow (\varphi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta)))$	*Conservation*
$L7:$	$\vdash (\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta))$	*L3, L5,* ***MP***
$L8:$	$\vdash (\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta))$	*L4, L6,* ***MP***
$L9:$	$\vdash ((\psi \wedge \varphi) \rightarrow \psi) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \vee \theta)))$	SH
$L10:$	$\vdash ((\psi \wedge \varphi) \rightarrow \varphi) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)))$	SH
$L11:$	$\vdash (\psi \wedge \varphi) \rightarrow (\psi \vee \theta)$	*L1, L7, L9,* ***MP***
$L12:$	$\vdash (\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)$	*L2, L8, L10,* ***MP***
$L13:$	$\vdash \theta \rightarrow (\psi \vee \theta)$	*Addition*
$L14:$	$\vdash \theta \rightarrow (\varphi \vee \theta)$	*Addition*
$L15:$	$\vdash ((\psi \wedge \varphi) \rightarrow (\psi \vee \theta)) \rightarrow ((\theta \rightarrow (\psi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)))$	*Disjonction*
$L16:$	$\vdash ((\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)) \rightarrow ((\theta \rightarrow (\varphi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)))$	*Disjonction*
$L17:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)$	$L11,\ L13,\ L15,\ MP$
$L18:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)$	$L12,\ L14,\ L16,\ MP$
$L19:$	$\vdash (((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)) \rightarrow$ $((((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow ((\psi \vee \theta) \wedge (\varphi \vee \theta)))$	*Composition*
$L20:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow ((\psi \vee \theta) \wedge (\varphi \vee \theta))$	$L17,\ L18,\ L19,\ MP$

Table 2: Preuve de la distributivité de $\vee$ sur $\wedge$ en utilisant une axiomatisation de type Hilbert

Etant donnée une axiomatisation, nous aimerions prouver qu'elle correspond à l'intuition que les théorèmes sont exactement les formules vraies. Plus précisément, on aimerait prouver qu'elle est correcte (c'est à dire que tout théorème est valide) et qu'elle est complète (c'est à dire que toute formule valide est un théorème).

La complétude est une propriété puissante qui garantit qu'on peut prouver tout ce qui est vrai. C'est cette propriété que [Gödel, 1951] a montré être fausse dans le cas de langages plus expressifs, en particulier l'arithmétique. Toutes les vérités mathématiques ne sont donc pas démontrables! Mais un tel résultat peut être satisfait pour les logiques modales. De fait, l'axiomatisation $\mathcal{A}$ est correcte et complète pour la logique booléenne, par rapport à la classe des modèles booléens.

De plus, il a été prouvé que l'axiomatisation présentée dans la table 3 est correcte et complète pour la logique K par rapport à la classe de tous les modèles de Kripke, autrement dit toute formule du langage modal est un théorème ssi elle est valide dans tous les modèles de Kripke.

Qui plus est, en utilisant la même axiomatisation agrémentée des axiomes additionnels D, T, 4 ou 5 (ou toute combinaison de ceux-ci) on obtient une axiomatisation correcte et complète par rapport à la classe de modèles $KD, KT, K4, K5$ (ou celle issue de la combinaison

PL	Axiomes de la logique propositionnelle
$K(A \longrightarrow B) \longrightarrow (KA \longrightarrow KB)$	Axiome K
A partir de A déduire KA	Nécessitation
A partir de A et de $A \rightarrow B$, déduire B	Modus Ponens

Table 3: Axiomes et règles de la logique modale

correspondante). Voir [Chellas, 1980] pour plus de détails.

Pour prouver la complétude d'une axiomatisation on utilise souvent la notion de *modèle canonique.* Nous sortirions du cadre de cette thèse en essayant de donner une définition générale de ce concept. La définition 0.19 donne donc une définition locale, suffisante dans le contexte de cet essai. Nous définissons d'abord la notion d'ensemble maximal consistant:

Definition 0.18 (Ensemble maximal consistant) *Soit $\mathcal{L}$ un langage et $\mathcal{A}$ une axiomatisation de ce langage. Un ensemble $\mathcal{S} \subseteq \mathcal{L}$ est dit:*

- inconsistant *si à partir des formules de $\mathcal{S}$ il est possible de déduire $\perp$ en utilisant $\mathcal{A}$ (*i.e. *$\mathcal{S} \vdash_{\mathcal{A}} \perp$)*
- consistant *sinon*
- maximal consistant *s'il est consistant et pour tout $\varphi \in \mathcal{L} \setminus \mathcal{S}$, $\mathcal{S} \cup \{\varphi\}$ est inconsistant*

Definition 0.19 (Modèle canonique pour les logiques épistémiques) *Soit $\mathcal{A}$ une des logiques qui peuvent être trouvées dans cette thèse, définie à partir d'un ensemble dénombrable d'atomes propositionnels $PROP$ et un ensemble dénombrables d'agents AG, dont le langage $\mathcal{L}_{\mathcal{A}}$ est basé sur des modalités épistémiques $\{K_i\}_{i \in AG}$ (dans le chapitre 4, 'K' est remplacé par 'B'). Le* modèle canonique *de $\mathcal{A}$ est le modèle de Kripke $\mathcal{M}^c = (S^c, \mathcal{R}^c, V^c)$ défini de la façon suivante:*

- $S^c = \{x \mid x$ *est un ensemble maximal consistant pour l'axiomatisation* $\mathcal{A}\}$
- $\mathcal{R}^c = \{R_i^c\}_{i \in AG}$ *où pour tout $i \in AG$, R_i^c est la relation binaire sur S^c suivante:* $R_i^c = \{(x, y) \in S^c \times S^c \mid K_i(x) \subseteq y\}$ *en notant* $K_i(x) = \{\varphi \mid K_i\varphi \in x\}$
- *pour tout $p \in PROP$, $V^c(p) = \{x \mid p \in x\}$.*

La plupart des preuves de complétude par rapport à une classe de modèles donnée utilise cette double nature du modèle canonique. En effet, nous souhaitons prouver dans ces situations que toute formule valide est un théorème de la logique. On commence donc par définir le modèle canonique de façon analogue à la dénfinition 0.19, on prouve qu'il s'agit bien d'un modèle, et qu'il appartient à la classe de modèles considérée. Si l'on prend alors une formule valide, elle est valide en particulier dans le modèle canonique (car on vient de prouver que c'est un modèle). Mais une formule valide du modèle canonique est un théorème. Pourquoi?

Car sa négation n'appartient à aucun ensemble maximal consistant, ce qui implique, comme nous aurons à le prouver, qu'elle n'appartient à aucun ensemble consistant. Si elle n'est consistante avec rien, c'est que cette formule est une contradiction, ce qui signifie que la formule initiale, qui est la négation d'une contradiction intrinsèque, est un théorème. En résumé, ceci démontrerait que toute formule valide est un théorème, autrement dit que toute formule vraie est démontrable.

Décidabilité et classes de complexité

Certains résultats de cette thèse relevant de la notion de *complexité*, nous en présentons dans ce paragraphe les éléments basiques.

Commençons par la décidabilité: on dit qu'un problème donné est *décidable* s'il existe une méthode automatique pour obtenir la réponse correcte à toute instance du problème. Nous pourrions appliquer cette notion à des problèmes de la vie courante. Par exemple 'A est-il plus grand que B' est un problème décidable: il est possible de mesurer. En effet si je veux tester si Alex (A) est plus grand que Brune (B) je peux appliquer ma méthode et obtenir la bonne réponse. Au contraire 'A est plus chanceux que B' semble être un problème indécidable.

Plus formellement, en informatique, un problème est dit décidable s'il existe une algorithme déterministe qui termine en répondant correctement oui ou non à toute instance du problème. Certains problèmes particuliers sont connus pour être indécidables, l'exemple le plus connu étant probablement le problème du domino (ou problème du pavage). L'objectif y est de savoir s'il est possible de paver une grille infinie en utilisant un ensemble fini donné de pavé colorés (dont chacun peut être utilisé autant de fois que l'on veut), en suivant les règles du domino. Un brique de Wang (Cf. [Wang, 1961]) est un carré dont chaque côté a une couleur choisie dans un ensemble fini de couleurs. On dit qu'un ensemble S de briques de Wang peut paver le plan si des copies de briques de S peuvent être placées, chacune à une position de la grille, de telle sorte que les côtés contigus de deux briques adjacentes soient de la même couleur. On peut utiliser de multiples copies de chaque brique, sans limitation sur le nombre. Si l'on accepte de pouvoir pivoter ou réfléchir les briques alors n'importe quelle brique de Wang peut à elle seule paver le plan. La question de savoir si un pavage existe pour un ensemble de briques de Wang donné n'est intéressant que dans le cas où nous n'autorisons aucune rotation ni réflexion, donc lorsque l'orientation de la brique est fixée. Par exemple, pensez-vous qu'il est possible de paver le plan avec l'ensemble de briques de Wang présenté Figure 0.2.2?

Ce problème de décision a été posé pour la première fois en 1961 par Wang dans [Wang, 1961] où il prouve qu'il est indécidable. Ca ne veut pas dire que l'on ne peut jamais savoir si un ensemble de briques donné permet de paver le plan. De fait, il a été prouvé qu'il existe un algorithme qui dit 'oui' en temps fini si le pavage est possible. Mais il est

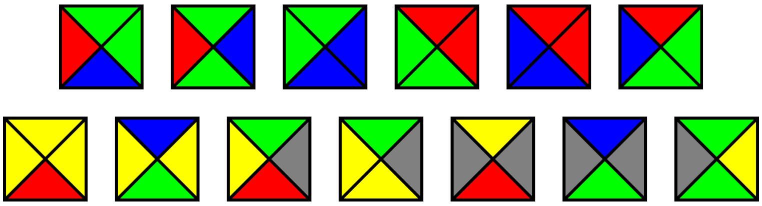


Figure 6: Une instance du problème du pavage

impossible d'être sûr d'avoir une réponse en temps fini: l'algorithme proposé peut ne jamais répondre lorsque le pavage est impossible.

Il est possible de prouver qu'un problème donné est indécidable en codant le problème du pavage. On traduit alors notre problème de telle sorte que pour chaque instance de celui-ci, sa traduction est une instance du problème du pavage (c'est à dire un ensemble de briques). Ainsi, si le problème était décidable alors le problème du pavage le serait lui aussi.

Mais à vrai dire les problèmes décidables nous intéressent davantage! On les classifie par la complexité de l'algorithme correspondant. En effet, on dit qu'un problème est dans P s'il peut être décidé par un algorithme déterministe dont l'exécution requiert un temps polynomial en la longueur de l'instance du problème. On dit qu'un problème est dans $EXPTIME$ si l'exécution de l'algorithme correspondant requiert un temps exponentiel en la taille de l'instance. On dit qu'il est dans NP (resp. $NEXPTIME$) si l'algorithme correspondant est non-déterministe, et qu'il est dans $PSPACE$ s'il requiert un temps exponentiel mais n'utilise qu'un espace polynômial. Parler d'algorithme est ici abusif (surtout dans le cas non déterministe), la définition rigoureuse utilise la notion de machine de Turing (voir [Papadimitriou, 1994]).

On appelle $EXPTIME$ la classe des problèmes qui sont dans $EXPTIME$, et ainsi de suite pour les autres classes de complexité. Nous savons que $P \subseteq NP \subseteq PSPACE \subseteq EXPTIME \subseteq NEXPTIME$. Nous savons également que $P \neq EXPTIME$, mais le fait de savoir si oui ou non $P = NP$ est un problème resté non-résolu qui pourrait vous rapporter un million de dollars (si toutefois vous le résolviez).

On dit qu'un problème P est NP-difficile si tout problème dans NP peut-être réduit à P. Plus formellement, P est NP-difficile si pour tout problème Q dans NP il existe une traduction tr exécutable en temps polynomial telle que pour toute instance i de Q, Q réponde oui à i si et seulement si P répond oui à $tr(i)$. On définit de la même manière les notions $PSPACE$-difficile , $EXPTIME$-difficile, $NEXPTIME$-difficile, *etc.* Si un problème est dans NP et est NP-difficile, on dit qu'il est NP-complet (et de même pour les autres classes de complexité).

Pour une définition plus rigoureuse de cette classification (qui implique d'expliquer en détails ce qu'est une machine de Turing), se reporter à [Papadimitriou, 1994].

En logique en général, et dans cette thèse en particulier, étant donnés un langage $\mathcal{L}$ avec sa sémantique et une classe de modèles $\mathcal{C}$ où l'on souhaite l'interpréter, on étudie deux problèmes classiques: le problème de satisfaisabilité (SAT) et le problème du *model checking* (MC). Ils peuvent être définis de la façon suivante:

MC: Etant donnés une formule φ de $\mathcal{L}$, un modèle fini $\mathcal{M} \in \mathcal{C}$ et un monde s de $\mathcal{M}$, φ est-elle satisfaite dans s?

SAT: Etant donnée une formule φ de $\mathcal{L}$, φ est-elle satisfaite dans un certain modèle $\mathcal{M}$ de la classe $\mathcal{C}$?

Une sémantique raisonnable assure que le problème du model checking est décidable (et appartient à une classe de complexité assez faible). En effet, c'est un minimum que d'être capable d'évaluer en temps fini si une formule est satisfaite ou non dans une situation donnée.

Le problème SAT pour la logique propositionnelle et celui pour la logique épistémique (avec les axiomes de $S5$) sont NP-complets, et ce problème est $PSPACE$-complet pour d'autres logiques modales. Il est indécidable pour certains langages plus expressifs, comme la logique du premier ordre par exemple, ou le langage $\mathcal{L}_{apal}$ présenté dans le paragraphe 3.1.2.

Etant donnée une logique, le problème SAT est important pour des motifs théoriques: il répond à la question de savoir si une formule fait sens, s'il existe une situation où elle est vraie. On dira donc qu'un langage est décidable (resp. NP-complet, $EXPTIME$-difficile, etc.) si son problème SAT est décidable (resp. NP-complet, $EXPTIME$-difficile, etc.).

La réponse à ce problème n'impose pas que soit donné un modèle qui satisfait la formule φ, mais dans certains cas nous aimerions également pouvoir construire ce modèle. On répond alors au problème de la construction du modèle:

Construction du modèle: Etant donnée une formule φ de $\mathcal{L}$, exhiber un modèle $\mathcal{M}$ satisfaisant φ.

Une méthode célèbre de construction de modèles utilise la notion de tableaux analytiques. Une telle méthode est présentée dans le paragraphe 6.6 pour le langage exprimant la connnaissance, les annonces publiques et la permission de donner une information.

Chapter 1

Introduction

Many situations involve communication and some kind of restrictions on this communication. This is the case when we think about military information, medical communication, moral norms, games, *etc.* In some situations, we may have structures to think about and organize the right to communicate in such situations. In the army, for example, such a structure is quite simple and easy to understand: the higher you are in the hierarchy, the more you may know and the less you are allowed to say. Indeed, a general can know any secret information but have no right to reveal it to his soldiers, while a soldier can give any information he wants (and may *have to* give the information he has) without having the right to access most of the information. As another example, in the medical field, more subtle restrictions prevent a boss from getting one of his workers' medical information, while a doctor may have access to it. Often such structures are presented as an informal and incomplete set of rules, that may be contradictory (and let the justice decide what should be done in case of conflict).

But we have no general framework to analyze such situations. The aim of this dissertation is to make some progress, in the field of logic, in the understanding of the notion of 'right to say', progress that may help us understand and answer problems that involve such a notion. We focus on the informative part of communication (and not on its form) leading our topic to the notion of 'right to give a piece of information'.

1.1 What is *Logic*?

Logic is the formal study of human arguments. In a way it can be considered as the study of human reasoning (if we consider arguments as the translation of internal reasoning in a human communication). Its aim is to get formal unambiguous results about it. Yet, natural language (in which are formed arguments) is particularly ambiguous, every word having different possible meanings and each concept having different interpretations in a same language. To form a logical theory, it is thus necessary to model a part of the reasoning, following predetermined conditions. This is what happens in the following famous syllogism, attributed to Aristotle[1]: *"All men are mortal. Socrates is a man. Therefore, Socrates is mortal."* Indeed, he supposes that the notion of mortality, man and 'being' have no ambiguity. It may be acceptable, but another sentence with the same structure, which is also quite

[1]We have no references to affirm that this sentence is actually from Aristotle, and the notion of syllogism in Aristotle's writings is nearer to: '*If all men are mortal and all Greeks are men, then all Greeks are mortal*'.

famous, shows the reader that it is not that obvious: "*All birds fly. Tweety the penguin is a bird. Therefore Tweety flies*". 'It is wrong that all birds fly', may answer the reader, and we could agree. But the interesting point is that there is an ambiguity in natural language on concepts: when we say that all birds fly, do we mean 'generally'? 'In every normal condition'? 'In every condition'? If an eagle breaks one of its wings, does that make wrong the fact that 'eagles fly'?

Therefore, to form a logical theory, we need to define a formal unambiguous language and a deterministic interpretation of its formulas. 'To interpret' a formula means here 'to say if it is true or false in the given context'. Nothing obliges us to consider the truth value as a binary function: true or false. Indeed, in our conception of reality some concepts are not binary: I am 1m76 tall, am I tall? Some would agree, others wouldn't, but it would be nearer to our common comprehension to say that I am rather tall, but not that much. Some logical theories (see for example [Dubois and Prade, 1988]) allow to consider this kind of concepts, which truth value is at the same time qualitative and quantitative. In this thesis, all the (abstract) concepts we consider can only be true or false (and not even both at the same time).

A representation of the world can thus be a list of all what is true. Such a big list may be impossible to get if we want to consider all the properties of the world (how many are they?), but in actual situations we can limit them to properties of interest and consider only these ones.

As an example, here are four children, Alex, Brune, Cha and Dan. We are interested only in their emotions, considered as binary: they feel good or bad. This may not be static: they are *moody* children, so these emotions are always changing.

We base our language on the following propositions: $Alex\_feels\_good$ (G_A), $Brune\_feels\_good$ (G_B), Cha_feels_good (G_C) and Dan_feels_good (G_D). Together they forms the set of *atomic propositions* of the language, noted $PROP = \{G_A, G_B, G_C, G_D\}$.

We then represent the actual world as a list of the truth values (true or false) of the propositions (taken from the set $PROP$). Figure 1.1 gives an example of such a representation, called *Boolean model*.

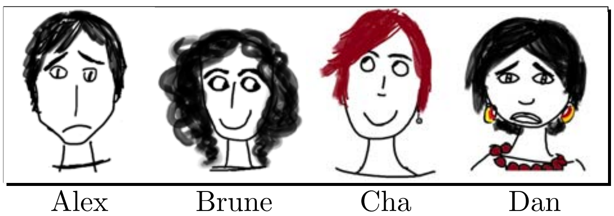


Figure 1.1: A boolean model for the moody children
The following propositions are true in this model : G_B, G_C

As they are moody, there are many possible worlds, exactly sixteen of them. They are represented in Figure 1.2.

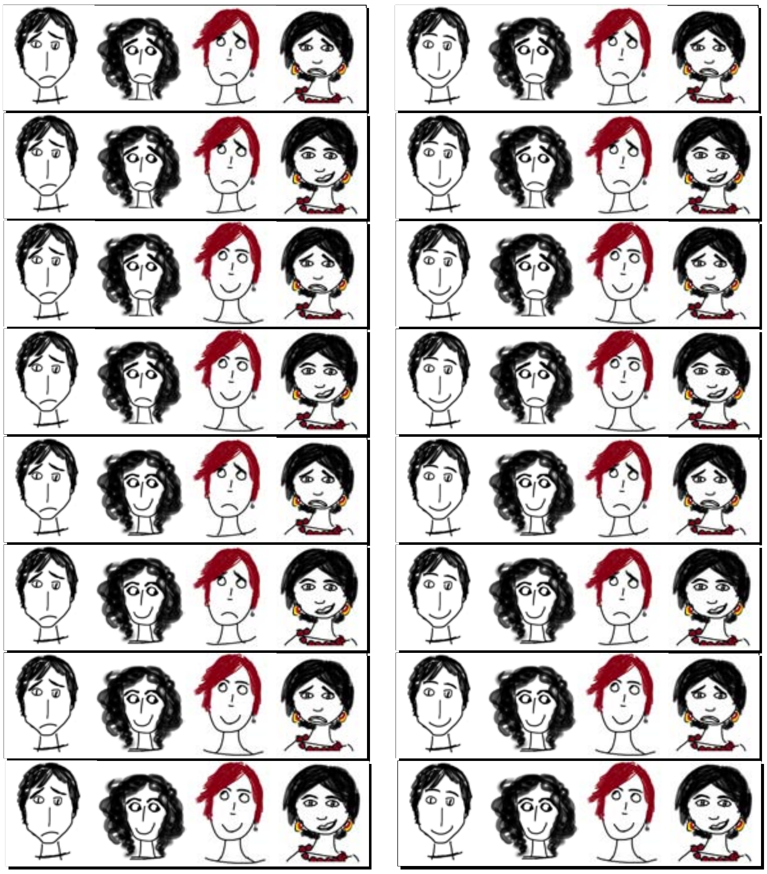


Figure 1.2: Moody children: all the possible worlds

These possible worlds are the basis of the representation of the actual world with modalities (such as time, belief, knowledge, result of actions, *etc.*). Such representations are introduced in the following chapter, using the notion of knowledge.

The link between logic and computer science is twofold. On one hand, computer science gives to logical theories important technical results, as deterministic algorithms that prove that a formula is true in a given context, or in every context. Also the time necessary to get such answer, in function of the size of the formula, can be proved. We will present in this work such kind of results. An introduction to them for the novice reader is proposed in Section 2.2.

On the other hand, logical theories give to computer scientists useful methods to solve actual problems. A fashionable example is the SUDOKU game: a classical algorithm may be extremely long to write, but a procedure formalising in a logical language the properties that have to be satisfied is quite easy to develop.

1.2 Outline

I think that scientific research should be permanently concerned about being accessible to a large majority of people. Clearly, not every scientific work can be understood by everybody, but every researcher can do her possible to give the elements of comprehension that allow someone out of his field to understand at least part of the work. It is particularly true for a PhD thesis that synthesizes years of work, with a non-imposed final size and that may be read by novice readers (friends, family,...). Chapter 2 is thus dedicated to present the basic notions of modal logic, in the context of the study of knowledge. Yet, some of these notions are much more general and can be used for any kind of modal logic.

More advanced frameworks of epistemic, dynamic and deontic logics are presented in Chapter 3. In this chapter we situate our work in the current research world and present some resources that we use in our proposals. We also discuss some principles that we may follow to correctly understand the notions linked to the 'right to say'.

While working during months on a given topic, many parallel questions rise and require an answer. Chapters 4 and 5 present the work that followed this process. Indeed, Chapter 4 deals with the concept of objective belief, a notion between knowledge and belief. It also presents technical results that complete a work proposed by [Hommersom et al., 2004]. As for Chapter 5, it presents a common work (published in [Ågotnes et al., 2010]) on the capacity of a group of agents to communicate information.

Chapters 6 and 7 present the most important result of this dissertation: a formalization of the 'right to say'. The former presents this notion in the context of public communication, *i.e.* in situations in which every communication is made publicly, and in which the restrictions to these communications are not dependent on the nature of the agent communicating. Indeed, in this framework, there is no agency that would allow us to say who is speaking, the only thing that matters is what is said. The report is based on an example, namely the french card game '*la Belote*'. The latter generalizes the first proposal, giving a framework including individual permissions for the agents communicating, and considering private communications, as well as public ones.

The last chapter concludes and opens toward further work. Indeed, this thesis is a work in progress that may be continued, in generalizing the framework or in analyzing different situations using such concepts.

Chapter 2

Modal Logic for the Representation of Knowledge

What does it mean that someone knows something? Can anything be known? These questions are not new, have been studied at least since the Greek philosophers (see [Plato, BC]) and form the field of Epistemology, the study of Knowledge. Some centuries after Plato, [Hintikka, 1962] proposed a formal logical analysis of knowledge in a multi-agent situation. His formalism, as we will see, uses the semantics of the possible worlds. Since then, epistemic logics have been used in various fields, such as artificial intelligence, economics, linguistics or theoretical computer sciences, focusing on the multi-agent aspects (interaction between agents, that can be human or computing systems) much more than on the philosophical understanding of knowledge.

You can then reason about what you know, about what your adversary knows, or about what makes part of the set of knowledge that is shared by all the agents. But how would Hintikka's formalism represent this knowledge?

2.1 Representation of Knowledge

2.1.1 The Language of Epistemic Logic

First of all, we need to define properly our language of *epistemic logic*, noted $\mathcal{L}_{el}$, starting from a countable set of agents AG and a countable set of propositional atoms $PROP$. In the example presented in Chapter 1, we consider $AG = \{a, b, c, d\}$ for *Alex*, *Brune*, *Cha* and *Dan*, and $PROP = \{G_A, G_B, G_C, G_D\}$. Here are some examples of formulas we can express in our language:

$K_b(G_A)$: "Brune knows Alex feels good"

$(G_C) \longrightarrow K_c(G_C)$: "If Cha feels good, she knows it"

$K_b(G_D \vee \neg G_D)$: "Brune knows Dan feels good or bad"

$\neg G_C \wedge K_a(G_C)$: "Cha feels bad and Alex knows Cha feels good".

More formally, here is the way in which the formulas are constructed:

Definition 2.1 (The Language $\mathcal{L}_{el}$) *The set $\mathcal{L}_{el}(AG, PROP)$ of epistemic formulas is obtained from AG and $PROP$ by iterating indefinitely the following operations:*

- *for all $p \in PROP$, p is a formula,*
- *$\bot$ ("falsum") is a formula,*
- *if φ is a formula then $\neg\varphi$ ("not φ") is a formula,*
- *if φ is a formula and ψ is a formula then $(\varphi \vee \psi)$ ("φ or ψ") is a formula,*
- *if φ is a formula then for all agent $i \in AG$, $K_i\varphi$ ("i knows that φ") is a formula.*
- *Nothing else is a formula except what can be constructed using these rules finitely many times.*

We often consider that the sets of atoms ($PROP$) and agents (AG) are clear or irrelevant, and we omit them. This definition can be written in the following shorter form[1]:

Definition 2.2 (The Language $\mathcal{L}_{el}$) *The language $\mathcal{L}_{el}$ over a countable set of agents AG and a countable set of propositional atoms $PROP$ is defined as follows:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid (\varphi_1 \vee \varphi_2) \mid K_i\varphi$$

where $i \in AG$ and $p \in PROP$.

We add some abbreviations:

- $\top$ ("true") abbreviates $\neg\bot$
- $(\varphi \wedge \psi)$ ("φ and ψ") abbreviates $\neg(\neg\varphi \vee \neg\psi)$
- $(\varphi \longrightarrow \psi)$ ("φ implies ψ") abbreviates $(\neg\varphi \vee \psi)$
- $(\varphi \longleftrightarrow \psi)$ ("φ is equivalent to ψ") abbreviates $((\varphi \longrightarrow \psi) \wedge (\psi \longrightarrow \varphi))$
- $\hat{K}_i\varphi$ ("i considers φ possible") abbreviates $\neg K_i\neg\varphi$. We say that $\hat{K}_i$ is the dual of K_i.

As we see, $\neg G_C \wedge K_a(G_B)$ (read "Cha feels bad and Alex knows she feels good") is a formula of the language, thus not all the formulas belonging to the language are intuitively true. But nobody said that every expressible formula had to be true. In fact, we did not define how to evaluate the truth value of an epistemic formula. Hence the following.

[1]This formal way to describe languages is inspired by the classical grammar in computer sciences called Backus-Naur Form (BNF). Hereafter we shall use this notation to describe formal languages.

2.1.2 The *Possible Worlds* Semantics

First, we assume that there exists an interpretation of the *real* world, independently of who is looking at it. This interpretation is a list of the truth values of all objective facts in the current state of affairs. If we call *propositions* these objective facts, we understand easily that this representation of the world is nothing else than a Boolean *propositional model*, as introduced in Chapter 1. In our example, it would be a list of the emotions of the moody children.

The lack of knowledge can thus be seen as an uncertainty about which is the current state of affairs. Hintikka represents thus the *epistemic world* (i.e. the world and the knowledge of all the agents) as a graph where a node is a representation of a possible world (i.e. a propositional model) and an edge, which is indexed by an agent a, represents the fact that agent a *does not know* if the current state of affairs is one or the other node linked by the edge. Reciprocally, we say that a *knows* a sentence φ if φ is true in all the state of affairs linked by edges to the current one. Here is a representation of a situation in which Brune does not know Cha's mood.

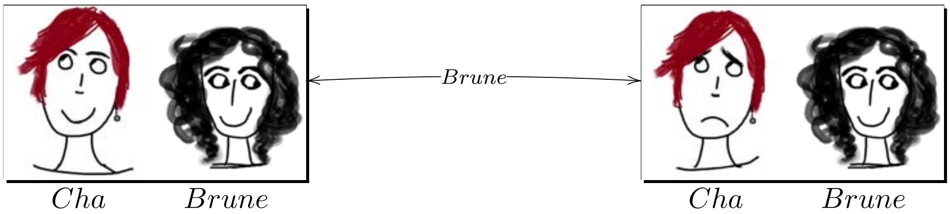


Figure 2.1: Epistemic model

Figure 2.2 gives a more complete representation of this kind of epistemic situations: Alex knows his own feelings and does not know Dan's ones, and he knows that Dan knows her feelings but not his ones. And Dan knows this fact, *etc.*

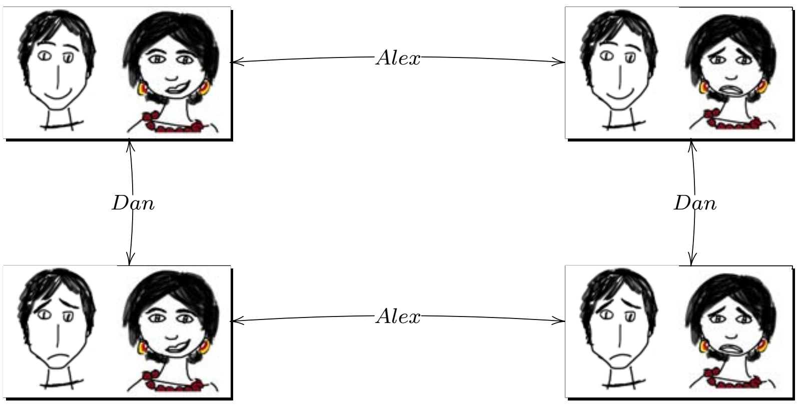


Figure 2.2: Epistemic model

We omit here the reflexive arrows that represent the fact that the children consider possible the actual world, which is considered always true.

Before defining these notions properly later, we define more precisely what a model is:

Definition 2.3 (Kripke model) *Given a countable set of agents AG and a countable set of propositional atoms $PROP$, a Kripke model is a tuple $\mathcal{M} = (S, \mathcal{R}, V)$ where:*

- *S is a set, each of its elements being called "world" or "state",*
- *$V : PROP \longrightarrow 2^S$ is a valuation function that assigns to any propositional atom p the set of worlds $V(p)$ in which p is considered true, and*
- *$\mathcal{R} = \{R_i\}_{i \in AG}$ with for all $i \in AG$, $R_i \subseteq S \times S$ is a binary relation on S.*

$(\mathcal{M}, s)$, a Kripke model joint with one of its states is called pointed *model.*

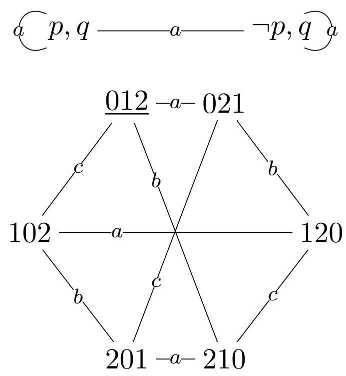


Figure 2.3: Some examples of Kripke models

Figure 2.3 gives two more classical representations of epistemic situations with Kripke models. The first model represents the uncertainty of agent a about p while q is known. The second one, explained in page 45, is a representation of a card deal, with a set of three cards 0, 1 and 2 dealt to three players a, b and c. In the second one, the reflexive arrows are omitted again.

These models allow us to interpret sentences that speak about truth of an objective fact, knowledge of agents about these facts, and knowledge of agents about this kind of sentences.

Definition 2.4 (satisfiability relation for $\mathcal{L}_{el}$) *Let $\mathcal{M}$ be a model. We define the satisfiability relation $\models : S \times \mathcal{L}_{el} \longrightarrow \{0, 1\}$ inductively on the structure of φ*[2] *in the following way: (We note $\mathcal{M}, s \models \varphi$, read "$\varphi$ is true in the state s of the model $\mathcal{M}$", if $\models (s, \varphi) = 1$ and $\mathcal{M}, s \not\models \varphi$, read "$\varphi$ is false in the state s of the model $\mathcal{M}$", if $\models (s, \varphi) = 0$)*

for all $s \in S$, $\mathcal{M}, s \models p$ iff $s \in V(p)$

for all $s \in S$, $\mathcal{M}, s \not\models \bot$

for all $s \in S$, $\mathcal{M}, s \models \neg\psi$ iff $\mathcal{M}, s \not\models \psi$

[2]This notion is presented in detail in Remark 2.7

for all $s \in S$, $\mathcal{M}, s \models \psi_1 \vee \psi_2$ *iff* $(\mathcal{M}, s \models \psi_1$ *or* $\mathcal{M}, s \models \psi_2)$

for all $s \in S$, $\mathcal{M}, s \models K_i\psi$ *iff for all* t *such that* sR_it, $\mathcal{M}, t \models \psi$

We say that φ *is valid in the model* $\mathcal{M}$*, noted* $\mathcal{M} \models \varphi$ *if for all* $s \in S$, $\mathcal{M}, s \models \varphi$*. We say that* φ *is valid if for all models* $\mathcal{M}$, $\mathcal{M} \models \varphi$*, i.e. if* φ *is valid in any state of any model. We note* $[\![\varphi]\!]_{\mathcal{M}}$ *the subset of* S *composed by the states* s *such that* $\mathcal{M}, s \models \varphi$*.*

2.1.3 Characterisation of Knowledge

We claimed, in Definition 2.1, that we could read $K_a\varphi$ as "agent a knows that φ". As said in Section 2, this claim, to be reasonable, needs to be followed by some arguments that make this semantics appropriate to speak about knowledge. Let us examine the validities given by the semantics, and the validities we should enforce to model a maybe idealized conception of knowledge.

First of all, the Kripke semantics, presented in Definition 2.4, imposes that our agents, who are able to know, have an absolute capacity of deduction. Why? Suppose that, in a given state of a given model , an agent i knows ψ and knows that ψ implies φ. Then ψ and $\psi \longrightarrow \varphi$ are satisfied in any state that i considers possible, thus φ is satisfied there also. In other words, the following formulas are valid, for all ψ, φ in the language:

$$(K_i\psi \wedge K_i(\psi \longrightarrow \varphi)) \longrightarrow K_i\varphi. \tag{K}$$

This could appear a reasonable simulation of the capacity of deduction of a rational agent. But it implies, for example, that every agent "knows" every boolean tautology, *i.e.* every sentence that is always true. But even if the reader is probably rational, could he say that he 'knows' that the following formula is a tautology?

$$(((p \vee t) \rightarrow s) \wedge (q \leftrightarrow u)) \vee ((p \vee t \wedge \neg s) \wedge ((q \wedge u) \vee (\neg q \wedge \neg u))) \longrightarrow (q \vee (v \wedge u \wedge (v \longrightarrow \bot)))$$

It is also widely accepted that if an agent knows something then it is true. Thus the following formulas are valid:

$$K_i\varphi \longrightarrow \varphi. \tag{T}$$

There clearly lies one difference between "knowledge" and "belief", though we cannot reduce knowledge to true belief (see [Burnyeat and Barnes, 1980]). Logics of belief usually avoid this principle of truth of belief, but use a deeper one: coherence. Indeed, we usually consider that if you believe something you do not believe its negation. This would be translated into one of the following equivalent principles (that remain true in the case of knowledge):

$$K_i\varphi \longrightarrow \neg K_i\neg\varphi \quad ; \quad \neg K_i\bot \tag{D}$$

We also accept that knowledge obeys positive and negative introspection. In other words, we consider that if an agent knows something, she knows that she knows it, and, which is more, if she does not know something, then she knows she does not. This implication is very strong: do you actually know what is your "knowledge base"? And do you know the entire list of what you do not know? If we accept these properties, we accept the validity of the following formulas, for all φ in the language:

$$K_i\varphi \longrightarrow K_iK_i\varphi \tag{4}$$

$$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi. \tag{5}$$

A last notion that we may introduce is the notion of *common knowledge.* Alex and Cha are used to play Poker (as you shall see, our moody children like to play cards). Alex knows the rules of the game. He also knows Cha knows the rules – if it were not the case, he could try to cheat. But he also knows Cha knows that he knows the rules – and therefore he can suppose she will not try to cheat. We could continue making sentences of this form... In fact, the rules of the game are common knowledge.

Formally, the common knowledge of φ is the syntactic abbreviation of an infinite conjunction of formulas. Let G be a set of agents, therefore $CK_G\varphi$ abbreviates $\bigwedge_{n\in\mathbb{N}}\bigwedge_{i_1,\dots,i_n\in G}K_{i_1}\dots K_{i_n}\varphi$.

As we shall see, this notion is very important when considering public learning: if Brune learns something together with Alex, and if each one can see that this learning is mutual, therefore the information learned becomes common knowledge. See [van Ditmarsch et al., 2009] for more details.

2.2 Classical Technical Notions in Modal Logic

This chapter may be harder to understand for the novice reader and may again be perfectly redundant for the expert. But it seems important to define and explain correctly the computer-science notions that are relevant in studies of logic. These notions are not restricted to the study of knowledge, on the contrary the majority of them are very common in all the fields of modal logic. Nevertheless, we shall present these notions using the language and semantics of epistemic logic.

2.2.1 Properties of the Language

Let us start with very basic notions regarding syntax in logical languages.

Definition 2.5 (Length of a formula) *Given a formula φ of a language $\mathcal{L}$ we call length of φ, noted $|\varphi|$, the number of symbols that constitute φ.*

Definition 2.6 (Subformula) *For all formula $\varphi \in \mathcal{L}_{el}$ we define $Sub(\varphi)$ the set of subformulas of φ depending of the form of φ:*

- $Sub(p) = \{p\}$
- $Sub(\bot) = \{\bot\}$
- $Sub(\neg\psi) = \{\neg\psi\} \cup Sub(\psi)$
- $Sub(\varphi_1 \vee \varphi_2) = \{\varphi_1 \vee \varphi_2\} \cup Sub(\psi_1) \cup Sub(\varphi_2)$
- $Sub(K_i\psi) = \{K_i\psi\} \cup Sub(\psi)$.

If $\psi \in Sub(\varphi)$ we say that ψ is a subformula of φ.

We can prove that $Sub(\varphi)$ is well defined by induction on the length of φ.

Remark 2.7 (Induction "on the structure of φ") *From now on, "prove* (resp. define) *a property $P(\varphi)$ by induction on the structure of φ" means "prove* (resp. define) *$P(\psi)$ for all $\psi \in PROP \cup \{\bot\}$ and prove* (resp. define) *$P(\varphi)$ admitting the following Induction Hypothesis (IH): $P(\psi)$ is true* (resp. defined) *for all subformula ψ of φ".*

The language studied here can express limited notions. If you add a modal operator (*i.e.* a new symbol) to a given language, without changing the semantics of the previous symbols, the language you obtain can clearly express *at least* the concepts that could be expressed by the previous language, and maybe more. Let us precisely describe this concept of expressivity of a language.

Definition 2.8 (Expressivity of a language) *Given languages $\mathcal{L}_1$ and $\mathcal{L}_2$ and a model class $\mathcal{C}$, $\mathcal{L}_1$ is* at least as expressive as *$\mathcal{L}_2$ with respect to $\mathcal{C}$ iff for every $\mathcal{L}_1$-formula φ there is a equivalent $\mathcal{L}_2$-formula ψ. In other words, for every $\mathcal{C}$-model $\mathcal{M}$, $[\![\varphi]\!]_{\mathcal{M}} = [\![\psi]\!]_{\mathcal{M}}$: the denotation of φ in $\mathcal{M}$ with respect to the $\mathcal{L}_1$-semantics is the same as the denotation of ψ in $\mathcal{M}$ with respect to the $\mathcal{L}_2$-semantics.*

Two standard ways to determine that $\mathcal{L}_1$ is at least as expressive as *$\mathcal{L}_2$ are:*

- *$\mathcal{L}_2$ form a sublanguage of $\mathcal{L}_1$*
- *there is a translation (reduction) from $\mathcal{L}_2$ to $\mathcal{L}_1$ such that every $\mathcal{L}_2$-formula is logically equivalent to its transalation in $\mathcal{L}_1$.*

The language $\mathcal{L}_1$ is more expressive than *$\mathcal{L}_2$ with respect to $\mathcal{C}$ if $\mathcal{L}_1$ is at least as expressive as $\mathcal{L}_2$, but $\mathcal{L}_2$ is not at least as expressive as $\mathcal{L}_1$ (the notion is a partial order).*

The standard way to determine that $\mathcal{L}_2$ is not at least as expressive as $\mathcal{L}_1$ is that there are an $\mathcal{L}_1$-formula φ and two $\mathcal{C}$-models $(\mathcal{M}, s)$ and $(\mathcal{M}', s')$ such that φ is true in $(\mathcal{M}, s)$ and false in $(\mathcal{M}', s')$, but any $\mathcal{L}_2$-formula ψ is true in $(\mathcal{M}, s)$ iff ψ is true in $(\mathcal{M}', s')$. We then also say that the language $\mathcal{L}_1$, but not $\mathcal{L}_2$, can distinguish *between the models $(\mathcal{M}, s)$ and $(\mathcal{M}', s')$.*

A last useful definition about languages:

Definition 2.9 (Substitution) *Let $\mathcal{L}(PROP)$ be a recursively enumerable language based on a countable set of atomic propositions $PROP$, $\varphi, \psi, \psi_1, \psi_2, \ldots \in \mathcal{L}(PROP)$ and $p, p_1, p_2, \ldots \in PROP$*

- *We denote as$\varphi(\psi/p)$ the $\mathcal{L}$ formula obtained from φ by replacing every occurrence of p in φ by ψ.*
- *We extend the previous notation to simultaneous substitution for the infinite sequences $p_1, p_2, \ldots$: $\varphi(\psi_1/p_1, \psi_2/p2, ...)$*

2.2.2 Properties of Models

Let us now see some semantical properties of modal logic, in other words some properties of the models we consider. First of all, the models we consider in all this work are Kripke models, possibly augmented with an additional relation (that can be between a world and set of worlds or between a world and relations). Let us recall that such models, defined in Definition 2.3, are composed of a set of states, binary relations on this set, and a valuation that assigns to any propositional atom a subset of states (those in which the proposition is true). Thus we can see them as oriented graphs where the nodes are boolean models (*i.e.* a truth value for any propositional atom). Note that this definition imposes the validity of the formula K for any operator which follows the semantics presented in Definition 2.4. Indeed, if $K_i\psi \wedge K_i(\psi \longrightarrow \varphi)$ is true in a state s of a model $\mathcal{M}$, then it means that for all states t linked to s by R_i, they satisfy both ψ and $\psi \longrightarrow \varphi$. As they are boolean models, we conclude that they all satisfy φ, QED.

A *submodel* of a given model $\mathcal{M}$ is composed of a subset of the states of $\mathcal{M}$, and a structure that is the restriction of the initial structure on the obtained subset. More formally:

Definition 2.10 (Submodel) *Let $\mathcal{M} = (S, V, \langle R_i \rangle_{i \in AG})$ be a Kripke model. Then we call a* submodel *of $\mathcal{M}$ a model $\mathcal{M}' = (S', V', \langle R'_i \rangle_{i \in AG})$ satisfying the following:*

- *$S' \subseteq S$*
- *for all $p \in PROP$ and all $s' \in S'$, $s' \in V'(p)$ iff $s' \in V(p)$*
- *for all $i \in AG$ and all $(s_1, s_2) \in S' \times S'$, $s_1 R'_i s_2$ iff $s_1 R_i s_2$*

Recall the notion of transitive closure of a set of relations in the context of Kripke model.

Definition 2.11 (transitive closure) *Let $\mathcal{R} = \{R_i\}_{i \in AG}$ be a set of binary relations over a set S. We call* transitive closure *of $\mathcal{R}$ the binary relation $\mathcal{R}^*$ such that for all $s, s' \in S$ there exist $n \in \mathbb{N}$ and $s_0, s_1, \ldots, s_n \in S$ satisfying:*

- *$s_0 = s$ and $s_n = s'$*

- *for all $k \in \{0, \dots, n-1\}$ there exists $i \in AG$ such that $s_k R_i s_{k+1}$.*

We can now observe that a model may not be *connected*, *i.e.* it may happen that one subset of states has no relation with another. In these conditions, a particular submodel comes to be relevant:

Definition 2.12 (Connected component - Generated submodel) *Let $\mathcal{M} = (S, V, \mathcal{R})$ be a model and $s \in S$. We call* connected component induced by s in $\mathcal{M}$ *the set $S' = \{s' \in S \mid s\mathcal{R}^* s'\}$. We call* generated submodel of $\mathcal{M}, s$ *the particular submodel $\mathcal{M}'$ of $\mathcal{M}$ based on the connected component induced by s.*

The connected composant of s in $\mathcal{M}$ is thus the set of all the states linked to s in $\mathcal{M}$. This notion is useful as the generated submodel of a model $\mathcal{M}$ is equivalent to $\mathcal{M}$ with respect to the language $\mathcal{L}_{el}$: a formula that is true in one is also true in the other. This is statuted by Proposition 2.15, saying that the two models are bisimilar.

Bisimulation is a well-known notion of structural similarity (see [Blackburn et al., 2001]) that we use frequently in examples and proofs. It sometimes says that two models are modally equivalent. Let us present it in details:

Definition 2.13 (Bisimulation) *Let two models $\mathcal{M} = (S, \mathcal{R}, V)$ and $\mathcal{M}' = (S', \mathcal{R}', V')$ be given. A non-empty relation $\mathfrak{R} \subseteq S \times S'$ is a bisimulation between $\mathcal{M}$ and $\mathcal{M}'$ iff for all $s \in S$ and $s' \in S'$ with $(s, s') \in \mathfrak{R}$:*

atoms *for all $p \in PROP$: $s \in V(p)$ iff $s' \in V'(p)$;*

forth *for all $i \in AG$ and all $t \in S$: if $sR_i t$, then there is a $t' \in S'$ such that $s'R'_i t'$ and $(t, t') \in \mathfrak{R}$;*

back *for all $i \in AG$ and all $t' \in S'$: if $s'R'_i t'$, then there is a $t \in S$ such that $sR_i t$ and $(t, t') \in \mathfrak{R}$.*

We write $(\mathcal{M}, s) \underline{\leftrightarrow} (\mathcal{M}', s')$ iff there is a bisimulation between $\mathcal{M}$ and $\mathcal{M}'$ linking s and s', and we then say the pointed Kripke structures $(\mathcal{M}, s)$ and $(\mathcal{M}', s')$ are bisimilar.

Note that bisimuation is an equivalence relation. Bisimulation is an important notion because it characterizes the fact that two models are modally equivalent, *i.e.* satisfy the same formulas of $\mathcal{L}_{el}$:

Proposition 2.14 *Let two models $\mathcal{M} = (S, \mathcal{R}, V)$ and $\mathcal{M}' = (S', \mathcal{R}', V')$ be given. Let $\varphi \in \mathcal{L}_{el}$ be a formula. For all $s \in S$ and for all $s' \in S'$, if $(\mathcal{M}, s) \underline{\leftrightarrow} (\mathcal{M}', s')$ then $\mathcal{M}, s \models \varphi$ iff $\mathcal{M}', s' \models \varphi$.*

The proof of this proposition can be found for example in [Fagin et al., 1995]. In particular, we obtain with the following proposition that a model satisfies the same formulas as its generated submodel.

Proposition 2.15 *Let $\mathcal{M}, s_0$ be a pointed model. It is bisimilar to its generated submodel.*

Proof Let $\mathcal{M}' = (S', \mathcal{R}', V')$ be the generated submodel of $\mathcal{M}, s_0$. Let $\mathfrak{R}$ be the binary relation between S and S' defined in the following way: $s\mathfrak{R}s'$ iff $s = s'$.

We show that $\mathfrak{R}$ is a bisimulation between $\mathcal{M}, s_0$ and $\mathcal{M}', s_0$. First, clearly $s_0\mathfrak{R}s_0$. For all $s \in S'$

atoms for all $p \in PROP$: $s \in V(p)$ iff $s \in V'(p)$ (by Definition 2.10);

forth for all $i \in AG$ and all $t \in S$: if sR_it, then $t \in S'$ and sR'_it by Definition 2.10, and $t\mathfrak{R}t$;

back for all $i \in AG$ and all $t \in S'$: if sR_it, then sR_it by Definition 2.10, and $t\mathfrak{R}t$.

$\square$

We introduce another kind of equivalence relation, that is a form of generalization of bisimulation. The idea is to consider as equivalent two states of a given model that satisfy all the formulas of a particular subset of the language. We then introduce the useful notion of filtration:

Definition 2.16 (filtration) *Let $\mathcal{M} = (S, \sim_i, V)$ be a model and let Γ be a set of formulas closed under subformulas. Let $\leftrightsquigarrow_\Gamma$ be the relation on S defined, for all $s, t \in S$, by:*

$$s \leftrightsquigarrow_\Gamma t \ \ \textit{iff} \ \ \textit{for all } \varphi \in \Gamma : (\mathcal{M}, s \models \varphi \ \textit{iff} \ \mathcal{M}, t \models \varphi)$$

Note that $\leftrightsquigarrow_\Gamma$ is an equivalence relation. We call the filtration of $\mathcal{M}$ through Γ *(or simply* the filtration of $\mathcal{M}$*) the model $\mathcal{M}^\Gamma = (S^\Gamma, \sim_i^\Gamma, V^\Gamma)$ where:*

- $S^\Gamma = S/_{\leftrightsquigarrow_\Gamma}$
- *for all $|s|, |t| \in S^\Gamma$, $|s| \sim_i^\Gamma |t|$ iff for all $K_i\varphi \in \Gamma$, $(\mathcal{M}, s \models K_i\varphi$ iff $\mathcal{M}, t \models K_i\varphi)$*
- $V^\Gamma(p) = \begin{cases} \emptyset \ if \ p \notin \Gamma \\ V(p)/_{\leftrightsquigarrow_\Gamma} \ if \ p \in \Gamma) \end{cases}$

An important last remark: in the entire class of all Kripke models, some particular subclasses may be useful. We define them according to the properties of its binary relations (reflexivity, transitivity, symmetry, seriality, euclidianity, equivalence). Recall that a binary relation R over a set S is

- reflexive if for all $s \in S$, $(s, s) \in R$
- transitive if for all $s, t, u \in S$, $((s, t) \in R$ and $(t, u) \in R)$ implies $(s, u) \in R)$
- symmetric if for all $s, t \in S$, $((s, t) \in R$ implies $(t, s) \in R)$
- serial if for all $s \in S$, there exists $t \in S$ such that $(s, t) \in R$

- euclidean if for all $s, t, u \in S$, $((s,t) \in R$ and $(s,u) \in R)$ implies $(t,u) \in R)$
- an equivalence relation if it is reflexive, transitive and symmetric.

As we will see in the following section, these properties of the models 'correspond' to the axioms we presented in Section 2.1.3 . More precisely,

- Truth (noted T) corresponds to reflexivity
- Coherence (noted D) corresponds to seriality
- Positive introspection (noted 4) corresponds to transitivity
- Negative introspection (noted 5) corresponds to euclideanicity

Recall that all the models satisfy the implication noted K. We thus call KT (resp. KD, $K45$, *etc.*) the class of reflexive models (resp. serial models, transitive and euclidian models, *etc.*). $S5$ abbreviates $KT45$ and corresponds to the class of models for which R is an equivalence relation. For all $n \in \mathbb{N}$ and for $\mathcal{C}$ a class of models, we call $\mathcal{C}_n$ the class of models that belongs to $\mathcal{C}$ and contains exactly n binary relations.

In particular, $S5_n$ is the class of Kripke model that have n relations that are equivalence:

Definition 2.17 (Epistemic model) *A Kripke model $\mathcal{M} = (S, \{R_i\}_{i \in AG}, V)$ is called an* epistemic model *if for all $i \in AG$, R_i is an equivalence relation over S.*

2.2.3 Axiomatization

Given a syntax (*i.e.* a language), a class of models where to interpret it (*i.e.* a class of concrete situations, a context) and semantics (*i.e.* an interpretation of the language in the contexts), we would like to characterize formulas that are true in this context. What are the properties of a formula that guarantee that it will be true in every situation of a given context? The question is not only to be able to determine which are the valid formulas, but to get an explanation of why they are true, and a proof of it.

$\top$	Truth
$(A \wedge B) \rightarrow A$; $(A \wedge B) \rightarrow B$	Simplification
$A \rightarrow (A \vee B)$; $B \rightarrow (A \vee B)$	Addition
$A \rightarrow (B \rightarrow A)$	Conservation
$(A \rightarrow B) \rightarrow ((A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow C))$	Hypothetical syllogism (HS)
$(A \rightarrow B) \rightarrow ((A \rightarrow C) \rightarrow (A \rightarrow (B \wedge C)))$	Composition
$(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow ((A \vee B) \rightarrow C))$	Disjunction
$\neg\neg A \rightarrow A$	Excluded middle
$(A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A)$	Coherence
From A and $A \rightarrow B$, infer B	Modus Ponens (MP)

Table 2.1: Propositional logic, axioms and rule

The notion of *axiomatization* has been developed for this purpose. Informally, an axiomatization is a finite description of axiom schemata and rules that allows to derive deterministically all the formulas we consider as *theorems* of the logic. More precisely, an axiomatization $\mathcal{A}$ is a set of axiom schemata (*i.e.* each formula that has the same structure of the schemata is an axiom and then is –by principle– a theorem) and a set of rules (called *inference rules*). Table 2.1 is an example of axiomatization, with nine axiom schemata and one inference rule. It axiomatizes propositional logic. We call *proof for* φ a finite sequence $\{\psi_1, \dots, \psi_n\}$ of formulas such that $\psi_n = \varphi$ and for all $i \in \{1, \dots, n\}$ either ψ_i is an instantiation of an axiom or is obtained from $\{\psi_1, \dots, \psi_{i-1}\}$ using an inference rule. If there is a proof for φ we say that φ is a *theorem* and we note $\vdash_{\mathcal{A}} \varphi$, or simply $\vdash \varphi$. More generally, if φ can be proved adding to $\mathcal{A}$ a set of formulas S considered as additional atoms, we note it $S \vdash_{\mathcal{A}} \varphi$.

Note that a proof may be quite long. Table 2.2 gives an example of a theorem' proof in propositional logic.

$L1:$	$\vdash (\psi \wedge \varphi) \rightarrow \psi$	*Simplification*
$L2:$	$\vdash (\psi \wedge \varphi) \rightarrow \varphi$	*Simplification*
$L3:$	$\vdash \psi \rightarrow (\psi \vee \theta)$	*Addition*
$L4:$	$\vdash \varphi \rightarrow (\varphi \vee \theta)$	*Addition*
$L5:$	$\vdash (\psi \rightarrow (\psi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta)))$	*Conservation*
$L6:$	$\vdash (\varphi \rightarrow (\varphi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta)))$	*Conservation*
$L7:$	$\vdash (\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta))$	*L3*, *L5*, ***MP***
$L8:$	$\vdash (\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta))$	*L4*, *L6*, ***MP***
$L9:$	$\vdash ((\psi \wedge \varphi) \rightarrow \psi) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \rightarrow (\psi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\psi \vee \theta)))$	*HS*
$L10:$	$\vdash ((\psi \wedge \varphi) \rightarrow \varphi) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \rightarrow (\varphi \vee \theta)) \rightarrow ((\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)))$	*HS*
$L11:$	$\vdash (\psi \wedge \varphi) \rightarrow (\psi \vee \theta)$	*L1*, *L7*, *L9*, ***MP***
$L12:$	$\vdash (\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)$	*L2*, *L8*, *L10*, ***MP***
$L13:$	$\vdash \theta \rightarrow (\psi \vee \theta)$	*Addition*
$L14:$	$\vdash \theta \rightarrow (\varphi \vee \theta)$	*Addition*
$L15:$	$\vdash ((\psi \wedge \varphi) \rightarrow (\psi \vee \theta)) \rightarrow ((\theta \rightarrow (\psi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)))$	*Disjunction*
$L16:$	$\vdash ((\psi \wedge \varphi) \rightarrow (\varphi \vee \theta)) \rightarrow ((\theta \rightarrow (\varphi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)))$	*Disjunction*
$L17:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)$	*L11*, *L13*, *L15*, *MP*
$L18:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)$	*L12*, *L14*, *L16*, *MP*
$L19:$	$\vdash (((\psi \wedge \varphi) \vee \theta) \rightarrow (\psi \vee \theta)) \rightarrow$ $((((\psi \wedge \varphi) \vee \theta) \rightarrow (\varphi \vee \theta)) \rightarrow (((\psi \wedge \varphi) \vee \theta) \rightarrow ((\psi \vee \theta) \wedge (\varphi \vee \theta)))$	*Composition*
$L20:$	$\vdash ((\psi \wedge \varphi) \vee \theta) \rightarrow ((\psi \vee \theta) \wedge (\varphi \vee \theta))$	*L17*, *L18*, *L19*, *MP*

Table 2.2: Proof for the distributivity of $\vee$ over $\wedge$ using Hilbert-style axiomatization

Given an axiomatization, we may want to prove that it corresponds to the intuition that the theorems are the always true formulas. More precisely, we may want to prove that it is sound (*i.e.* every theorem is a validity) in all considered models and complete (*i.e.* every valid formula is a theorem).

Completeness is a powerful property that guarantees that all that is true can be proved. This is the property that [Gödel, 1951] proved to be wrong for more expressive languages, in particular arithmetic. But such a result may be satisfied for modal logics. In fact, Hilbert-style axiomatization for propositional logic is proved to be sound and complete in all Boolean models.

What is more, it has been proved that the axiomatization presented in Table 2.3 is sound and complete with respect to all Kripke models, *i.e.* that a formula is a theorem of modal logic iff it is valid in all Kripke models.

PL	Axioms of Prop. Logic as in Table 2.1
$K_i(A \longrightarrow B) \longrightarrow (KA \longrightarrow KB)$	Axiom K
From A infer K_iA	Necessitation
From A and $A \rightarrow B$, infer B	Modus Ponens

Table 2.3: Modal logic, axioms and rules

Moreover, using the same axiomatization with additional axioms D, T, 4 or 5 (or any combination of them) we obtain a sound and complete axiomatization with respect to the corresponding class of models $KD, KT, K4, K5$ (or the corresponding combination). See [Chellas, 1980] for details.

To prove completeness, we often use the notion of *canonical model*. Trying to give a general definition of this concept would lead us out of our purpose. Definition 2.19 gives thus a local definition, that is sufficient in the context of this thesis. We first introduce the notion of maximal consistent set:

Definition 2.18 (Maximal consistent set) *Let a language $\mathcal{L}$ and an axiomatization for this language $\mathcal{A}$ be given. A set $\mathcal{S} \subseteq \mathcal{L}$ is said to be:*

- inconsistent *if from the formulas of $\mathcal{S}$ it is possible to derive $\bot$ using $\mathcal{A}$ (*i.e. $\mathcal{S} \vdash_{\mathcal{A}} \bot$*)*
- consistent *otherwise*
- maximal consistent *if it is consistent and for all $\varphi \in \mathcal{L} \setminus \mathcal{S}$, $\mathcal{S} \cup \{\varphi\}$ is inconsistent*

Definition 2.19 (Canonical model for epistemic logics) *Let $\mathcal{A}$ be one of the logics that can be found in this thesis, defined over a countable set of propositional atoms $PROP$ and a countable set of agents AG, which language $\mathcal{L}_{\mathcal{A}}$ is based on the epistemic modalities $\{K_i\}_{i \in AG}$ (in chapter 4, symbol 'K' is replaced by 'B'). The* canonical model *of $\mathcal{A}$ is the Kripke model $\mathcal{M}^c = (S^c, \mathcal{R}^c, V^c)$ defined as follows:*

- $S^c = \{x \mid x$ *is a maximal consistent set of the axiomatization of* $\mathcal{A}\}$
- $\mathcal{R}^c = \{R_i^c\}_{i \in AG}$ *with for all $i \in AG$, R_i^c is the following binary relation over S^c:* $R_i^c = \{(x, y) \in S^c \times S^c \mid K_i(x) \subseteq y\}$ *where* $K_i(x) = \{\varphi \mid K_i\varphi \in x\}$
- *for all $p \in PROP$, $V^c(p) = \{x \mid p \in x\}$.*

Most of the proofs of completeness with respect to a given class of models use the double nature of the canonical model. Indeed, in such proofs we want to show that any valid formula is a theorem of the logic. Therefore, we first define the canonical model in a similar way as in

Definition 2.19, prove that it *is* a model, and that it belongs to the corresponding class. We then take a valid formula which, as valid in every model, is valid in the canonical model. But a valid formula of the canonical model is a theorem. Why? Because, its negation does not belong to any maximal consistent set, which implies, as we would have to prove, that it does not belong to any consistent set. And if it cannot be consistent with anything, then it is an intrinsic contradiction, which means that the initial formula, its negation, is a theorem. To sum it up, this would prove that every valid formula is a theorem, which means that every true formula is provable.

2.2.4 Decidability and Classes of Complexity

Some of the results of this thesis being related with the notion of *problem complexity*, we present briefly in this section the basic notions of this concept.

Decidability is the first useful notion: informally we say that a given problem is *decidable* if there is an automatic method to find the correct answer for every instance of the problem. We could imagine a comparison with real-life problems. For example 'Is A taller than B' is a decidable problem: it is possible to measure. Indeed if I want to test if Alex (A) is taller than Brune (B) I can apply my method and get the good answer. But 'Is A more lucky than B' seems to be an undecidable problem.

More formally, in computer science, a problem is said decidable if there exists a deterministic algorithm which ends answering correctly yes or no to every instance of the problem. Some particular problems are known to be undecidable. The most famous example is probably the Domino problem (also known as the Tiling problem). The purpose is to know if it is possible to tile an infinite grid using a given finite set of reproducible coloured tiles, following the condition of the domino game. A Wang tile (cf. [Wang, 1961]) is a unit square with each edge colored from a finite set of colors. A set S of Wang tiles is said to tile a planar grid if copies of tiles from S can be placed, one at each grid position, such that abutting edges of adjacent tiles have the same color. Multiple copies of any tile may be used, with no restriction on the number. If we allow the tiles to be rotated or reflected, any single Wang tile can tile the plane by itself. The question of whether such a tiling exists for a given set of tiles is interesting only in the case where we do not allow rotation or reflection, thus holding tile orientation fixed. For example, do you think it is possible to tile the plane with the following set of Wang tiles?

This decision problem was first posed in 1961 by Wang in a seminal paper ([Wang, 1961]) and has been proved to be undecidable. That does not mean that you can never decide if a given set allow to tile the plane. In fact it has been proved that there is an algorithm that says yes in finite time if the tiling is possible. But the algorithm proposed before may never answer in the case where the tiling is not possible.

One way to show that a given problem is undecidable is by encoding the tiling problem

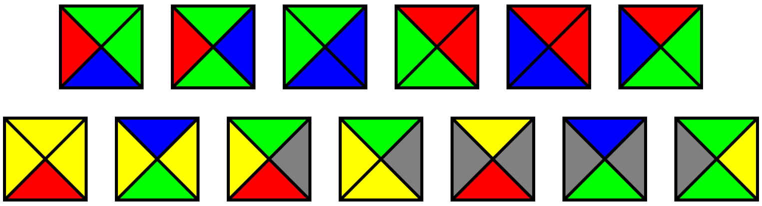


Figure 2.4: An instance of the domino problem

into it. We translate our problem so that for every instantiation of it, its translation is an instantiation of the Tiling problem (*i.e.* every finite set of tiles). Therefore if the given problem were decidable, so would be the Tiling problem.

But we are more interested in decidable problems! Those are classified by the complexity of the relative algorithm. Indeed, we say that a problem is in P if it can be decided by a deterministic algorithm which execution requires a time polynomial on the length of the given problem. We say that a problem is in $EXPTIME$ if the execution of the relative algorithm requires a time exponential on the length of the input. We say that it is in NP (resp. $NEXPTIME$) if the corresponding algorithm is non deterministic, and that it is in $PSPACE$ if it requires only polynomial space.

We call $EXPTIME$ the class of problems that are in $EXPTIME$, and so on for the other classes of complexity. It is a known fact that $P \subseteq NP \subseteq PSPACE \subseteq EXPTIME \subseteq NEXPTIME$. It is also known that $P \neq EXPTIME$, but to know if $P = NP$ or not is an open problem that may yield you one million dollars (if you solve it).

We say that a problem A is NP-hard if every NP problem can be reduced to A. More formally, A is NP-hard if for every problem B in NP there is a translation tr computable in polynomial time such that for all instances i of B, B answers yes to i if and only if A answers yes to $tr(i)$. We define in the same way $PSPACE$-hardness , $EXPTIME$-hardness, $NEXPTIME$-hardness, *etc.* If a problem is in NP and is NP-hard, we say that it is NP-complete (and so on for the other complexity classes).

For a more rigorous definition of this classification (which requires to explain what a Turing machine is), the reader may look at [Papadimitriou, 1994].

In logic in general and hereafter in particular, given a language $\mathcal{L}$ together with its semantics and a class $\mathcal{C}$ of models where to interpret it, two classical problems are studied: the problem of Satisfiability (SAT) and the problem of the model checking (MC). They can be defined in the following way:

MC: Given a formula φ of $\mathcal{L}$ and a finite model $\mathcal{M} \in \mathcal{C}$, is φ satisfied in a world of $\mathcal{M}$?

SAT: Given a formula φ of $\mathcal{L}$, is φ satisfied in a model $\mathcal{M}$ of the class $\mathcal{C}$?

Reasonable semantics ensure that the problem of the model checking is decidable (and with a low class of complexity). Indeed, a requirement we may demand to such semantics is to be able to evaluate in finite time if a formula is satisfied or not in a given finite model. The SAT problem for propositional logic and epistemic logic (with the $S5$ axioms and only one agent) is NP-complete, and it is $PSPACE$-complete for some other modal logics. It is undecidable for some more expressive languages, as first order logic for example, or the language $\mathcal{L}_{apal}$ presented in Section 3.1.2.

Given a logic, the SAT problem is important at a theoretical level: it answer to the question if a given formula makes sense. Therefore, we say that a language is decidable (resp. NP-complete, $EXPTIME$-hard, etc.) if the SAT problem is decidable (resp. NP-complete, $EXPTIME$-hard, etc.). The answer to this problem does not require to give a concrete model where the formula φ is satisfied, but in some cases we would like to construct such a model. In this case, we find an answer to the problem of the model construction:

Model Construction: Given a formula φ of $\mathcal{L}$, give a model $\mathcal{M}$ satisfying φ.

A well known method of model construction uses analytic tableaux. One such method is presented in Section 6.6 for a language expressing knowledge, announcements and permission to give a piece of information.

Chapter 3

State of Art

In this chapter we present some of the works that inspired us in developing our formalisms. Indeed, to understand the notion of "right to communicate a piece of information" some notions have to be clarified from the start. We would like to precisely define what we mean by "communication", which notion of "right" we shall use, and if previous works tried to merge these notions.

The first part of this chapter is thus dedicated to the logics of knowledge and communication, called *Dynamic Epistemic Logics* tackling also the concept of *Knowability*: what can be known? Some gaps appear and justify the works presented in the chapters 4 and 5.

We present in a second part the basic ideas of deontic logics, the logics of obligation, permission and prohibition, before seeing in a third section which works tried to merge both fields in a formalism of the right to communicate.

3.1 Dynamic Epistemic Logics

3.1.1 Public Announcement Logic

Consider an example involving our moody children: Brune and Alex reached the final round of the *Texas Hold'em* Poker game (Cha and Dan, in a bad mood, let them play alone). *Texas*

Figure 3.1: Texas Hold'em

Hold'em is a very fashion-conscious variant of classical poker in which each player gets five cards from a set of fifty-two. At the end of the deal each one will have two cards in his/her hand, and five cards will be face up on the table. Each player is then able to form her hand, which is composed by the best 5-cards combination using his own cards and the five cards shared with her adversaries. This variant of classical poker makes the game more interesting because the uncertainty about the cards of your adversary is not total, some of them being common knowledge to all players.

But not all the cards are dealt at the beginning, and the structure of the deal makes the game even more interesting. Indeed, each player first receives her two cards and a round of bets starts. Then three cards are dealt face up ('the *flop*'),then a fourth ('the *turn*') and a fifth ('the *river*'), the three of these deals being followed by a round of bets. Each of these deals can be seen as an announcement to all players that one card is part of the common final hand. This announcement is public, in the sense that everybody knows it takes places, and everybody knows everybody knows it, and so on... In fact, it is common knowledge (cf. Section 2.1.3) that it takes place. Clearly, this announcement will change the state of knowledge of the agents about the world, though it does not change the factual world (because the deal was fixed before the game started, and no actions that are part of the game can change it). After this announcement, the players will eliminate from the deals they considered possible the ones in which the revealed card wasn't on the table. But also, as the announcement is public, they know that nobody consider these deals possible anymore, so they will eliminate them from the representation they have of the state of knowledge of the other agents. And they know that the other agents know that no other agent consider those deals possible anymore... They are erased from the common representation of the world!

This is the basic idea of Public Announcement Logic, that was first introduced by [Plaza, 1989]. Let us introduce its language and semantics:

Definition 3.1 (Syntax of $\mathcal{L}_{pal}$) *The language of public announcement logic $\mathcal{L}_{pal}$ over a countable set of agents AG and a countable set of propositional atoms $PROP$ is defined as follows:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid K_i\varphi \mid [\varphi_1]\varphi_2$$

where $i \in AG$ and $p \in PROP$.

The boolean and epistemic parts of the language are read as usual (See Chapter 2) and $[\varphi_1]\varphi_2$ is read 'after the public announcement of ψ_1, ψ_2 is true'.

This language is interpreted in the same models than $\mathcal{L}_{el}$ in the following way.

Definition 3.2 (Semantics of $\mathcal{L}_{pal}$ and restricted model) *Let $\mathcal{M}$ be a model and let s be a state of S. The satisfiability relation $\mathcal{M}, s \models \varphi$ is defined inductively on the structure of φ as follows:*

$\mathcal{M}, s \models p$ *iff* $s \in V(p)$

$\mathcal{M}, s \not\models \bot$

$\mathcal{M}, s \models \neg\psi$ *iff* $\mathcal{M}, s \not\models \psi$

$\mathcal{M}, s \models \psi_1 \vee \psi_2$ *iff* $(\mathcal{M}, s \models \psi_1$ *or* $\mathcal{M}, s \models \psi_2)$

$\mathcal{M}, s \models K_i\psi$ *iff for all* t *such that* sR_it, $\mathcal{M}, t \models \psi$

$\mathcal{M}, s \models [\psi]\chi$ *iff* ($\mathcal{M}, s \models \psi$ *implies* $\mathcal{M}|\psi, s \models \chi$)

where $\mathcal{M}|\psi = (S', \mathcal{R}', V')$ *is the update (or restriction) of a model* $\mathcal{M}$ *after the public announcement of* ψ*, defined as:*

- $S' = [\![\psi]\!]_{\mathcal{M}} = \{s \in S | \mathcal{M}, s \models \psi\}$
- $V'(p) = V(p) \cap S'$ *for all* $p \in PROP$
- *for all* i, $R'_i \ = \ R_i \cap \ (S' \times S')$

Remark 3.3 *This definition looks incorrect, because it uses the satisfiability relation to define the restricted model, and* vice versa. *But to define the restricted model, we only need the definition of the satisfiability relation for a subformula* ψ *of the initial one, and thus for all subformulas of* ψ*. The definition is thus well-founded by induction on the structure of the formula.*

Here is, as an example, the representation of a simple deal of cards between three of our children:

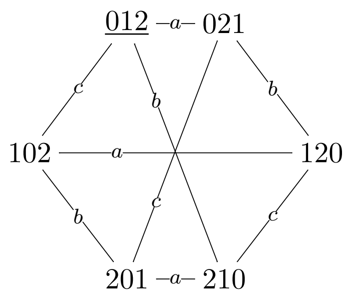


Figure 3.2: Deal of three cards between three agents
Actual state '012' represents the deal in which Alex has 0, Brune 1 and Cha 2.

Cha, Alex and Brune have one card each, dealt from a set of three cards only, $\{0, 1, 2\}$. They know that these are the dealt cards, though they don't know the actual deal in which Alex has 0, Brune 1 and Cha 2. The propositional atoms are of the form X_i with $X \in \{0, 1, 2\}$ and $i \in \{a, b, c\}$. We read for example 2_a as 'Alex has 2'. We labelled the states of the models so that the name of a state makes clear the deal it represents. The actual state, for example, is called '012'. The model of Figure 3.2 represents this epistemic situation. We call this model $\mathcal{M}_*$. We omitted reflexive arrows in each state for each of our three children.

What happens now if Alex announces publicly that he does not have the 1? The result appears in Figure 3.3.

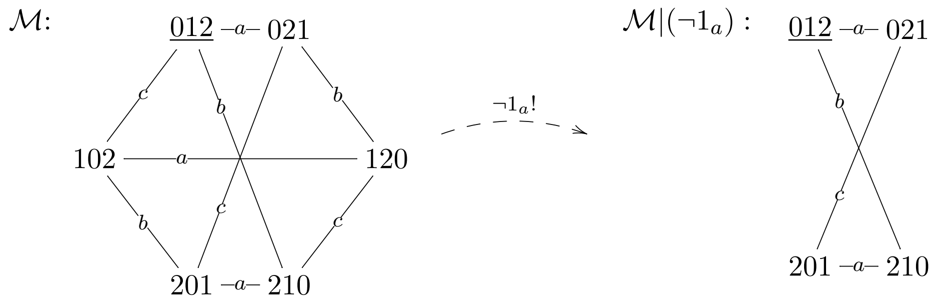


Figure 3.3: Alex announces: "*I don't have the 1*"

As we said before, the result of this public announcement is that the states in which 1_a was true are erased from the model. This comes from the strong property of public announcement: not only everybody learns its content, but the states for which it was false before the announcement are not conceivable anymore.

After the announcement, Cha knows what is the exact distribution, in other words $\mathcal{M}, 012 \models \langle \neg 1_a \rangle K_c(0_a \wedge 1_b \wedge 2_c)$. Brune does not know which it is, but she knows Cha knows it: $\mathcal{M}, 012 \models \langle \neg 1_a \rangle K_b(K_c(0_a \wedge 1_b \wedge 2_c) \vee K_c(2_a \wedge 1_b \wedge 0_c))$.

Remark 3.4 (Moore Sentences) *We said 'its content* was *true* before *the announcement', and this use of the past may appear as superfluous. Indeed, we could believe that after an announcement of φ, φ is still true and the receiver of the message knows that φ. But this is not true in general, because a formula may become false after it has been announced. We call this an* unsuccessful update*. Let us see an example.*

*Suppose the reader you are does not know what is a Moore sentence. If you are not used to epistemic logics this may be the case. Then let me say to you the following: ''*You don't know that Moore sentences are unsuccessful updates but they are!*".*

This sentence was true before my announcement, indeed Moore sentences were actually unsuccessful updates and you didn't know that. But now you do! (Though you may not see what precisely is a Moore sentence, if you believe me you know at least that it is an unsuccessful update, because I told you so). Therefore this sentence was true before I announced it, and it became false after. We call Moore sentences *this kind of formulas, of the form $p \wedge \neg K_i p$, presented by [Hintikka, 1962]. For more information about Moore sentences, see [van Ditmarsch and Kooi, 2006]*

Properties: One first (and surprising) property is that $\mathcal{L}_{pal}$ is not more expressive than $\mathcal{L}_{el}$. This comes from the following two principles:

- announcements cannot change the objective state of the world (they only change the knowledge of the agents about it)
- an announcement teaches something to an agent iff she knows that this announcement (if possible) would teach it to her

More formally, we define (by induction on the structure of φ) the following translation tr from $\mathcal{L}_{pal}$ to $\mathcal{L}_{el}$ such that any $S5_n$-model valid, for all $\varphi \in \mathcal{L}_{pal}$, $tr(\varphi) \longleftrightarrow \varphi$.

- $tr(p) = p$
- $tr(\bot) = \bot$
- $tr(\neg\varphi) = \neg tr(\varphi)$
- $tr(\psi \vee \varphi) = tr(\psi) \vee tr(\varphi)$
- $tr(K_i\varphi) = K_i tr(\varphi)$
- $tr([\psi]p) = tr(\psi) \longrightarrow p$
- $tr([\psi]\bot) = \neg tr(\psi)$
- $tr([\psi]\neg\varphi) = tr(\psi) \longrightarrow \neg tr([\psi]\varphi)$
- $tr([\psi](\varphi_1 \vee \varphi_2)) = tr([\psi]\varphi_1) \vee tr([\psi]\varphi_2)$
- $tr([\psi]K_i\varphi) = tr(\psi) \longrightarrow K_i tr([\psi]\varphi)$

The same idea gives us an axiomatization PAL for this language. Indeed, we need only to take the axiomatization for $S5_n$ augmented with the reduction axioms corresponding to the last five equalities (see table 3.1).

Nevertheless, Public Announcement Logic is not only a beautiful way of thinking about epistemic logic, it has some interesting properties. Indeed, $\mathcal{L}_{pal}$ is proved to be exponentially succinct on unrestricted structures ([Halpern and Moses, 1992, Lutz, 2006]), and interesting complexity results have been given, even when we consider common knowledge (see [Lutz, 2006] for details).

Public announcement logic have been widely studied and extended. The next two sections present two of those extensions that are useful to our work.

3.1.2 What Is Knowable? - Arbitrary Public Announcement Logic

Arbitrary public announcement logic [Balbiani et al., 2008] has been developed to tackle the problem of what is called 'knowability' in Philosophy. In [Fitch, 1963], Fitch addresses the question whether what is true can become known. As we saw with Moore sentences, this is not true in general. Furthermore Fitch's Paradox ([Fitch, 1963]) ensures that if every truth

instantiations of propositional tautologies	
$K_i(\varphi \longrightarrow \psi) \longrightarrow (K_i\varphi \longrightarrow K_i\psi)$	distribution (of knowl. over impl.)
$K_i\varphi \longrightarrow \varphi$	truth
$K_i\varphi \longrightarrow K_iK_i\varphi$	positive introspection
$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi$	negative introspection
$[\varphi]p \longleftrightarrow (\varphi \longrightarrow p)$	atomic permanence
$[\varphi]\neg\psi \longleftrightarrow (\varphi \longrightarrow \neg[\varphi]\psi)$	announcement and negation
$[\varphi](\psi \wedge \chi) \longleftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$	announcement and conjunction
$[\varphi]K_i\psi \longleftrightarrow (\varphi \longrightarrow K_i[\varphi]\psi)$	announcement and knowledge
$[\varphi][\psi]\chi \longleftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	announcement composition
From φ and $\varphi \longrightarrow \psi$, infer ψ	modus ponens
From φ, infer $K_i\varphi$	necessitation of knowledge
From φ, infer $[\psi]\varphi$	necessitation of announcement

Table 3.1: PAL axioms and rules

can be known then every truth is actually known. An overview of the different studies on Fitch's paradox can be found in [Brogaard and Salerno, 2004].

Public announcement logic is not expressive enough to face this notion. The main idea of the work presented in this chapter is to add a quantified modality to $\mathcal{L}_{pal}$ and to interpret 'knowable' as 'known after some announcement'.

Therefore, the language $\mathcal{L}_{apal}$ extends $\mathcal{L}_{pal}$ with an additional inductive construct $\Box\varphi$, read 'after any possible announcement, φ becomes true'. In other words:

Definition 3.5 (Syntax of $\mathcal{L}_{apal}$) *The language of arbitrary public announcement logic $\mathcal{L}_{apal}$ over a countable set of agents AG and a countable set of propositional atoms $PROP$ is defined as follows:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid K_i\varphi \mid [\varphi_1]\varphi_2 \mid \Box\varphi$$

where $i \in AG$ and $p \in PROP$.

We denote by $\Diamond$ the dual of $\Box$, *id est* $\Diamond\varphi := \neg\Box\neg\varphi$: 'there is some announcement after which φ is true'. Therefore φ would be 'knowable' if $\Diamond(K\varphi)$. We interpret this language in the same class of models, the epistemic models. The interpretation of this new modality $\Box$ (the other ones remaining identical) is :

$$\mathcal{M}, s \models \Box\varphi \text{ iff } \text{ for all } \psi \in \mathcal{L}_{el} : \mathcal{M}, s \models [\psi]\varphi.$$

This logic has many interesting properties, most of them proved in [Balbiani et al., 2007] or [Balbiani et al., 2008]:

- $\mathcal{L}_{apal}$ is strictly more expressive than $\mathcal{L}_{pal}$.
- $\Box$ has the $S4$ properties, that is for all $\varphi, \varphi \in \mathcal{L}_{apal}$:

1. $\models \Box(\varphi \wedge \psi) \longleftrightarrow \Box\varphi \wedge \Box\psi$
2. $\models \Box\varphi \longrightarrow \varphi$
3. $\models \Box\varphi \longrightarrow \Box\Box\varphi$
4. $\models \varphi$ implies $\models \Box\varphi$.

- Mac-Kinsey (MK) and Church-Rosser (CR) formulas are valid for $\Box$, that is for all $\varphi \in \mathcal{L}_{apal}$:

  MK: $\models \Box\Diamond\varphi \longrightarrow \Diamond\Box\varphi$

  CR: $\models \Diamond\Box\varphi \longrightarrow \Box\Diamond\varphi$

- For all $\varphi \in \mathcal{L}_{apal}$, $K_a\Box\varphi \longrightarrow \Box K_a\varphi$ is valid, but not the converse.

- There is a finite axiomatization for $\mathcal{L}_{apal}$ that is shown to be sound and complete. The axioms and inference rules involving arbitrary announcement are:

$$\begin{array}{ll} \Box\varphi \longrightarrow [\psi]\varphi & \text{where } \psi \in \mathcal{L}_{el} \\ \text{From } \varphi, \text{infer } \Box\varphi & \\ \text{From } \psi \longrightarrow [\theta][p]\varphi, \text{infer } \psi \longrightarrow [\theta]\Box\varphi & \text{where } p \notin \Theta_\psi \cup \Theta_\theta \cup \Theta_\varphi \end{array}$$

  where Θ_φ denotes the set of atoms occurring in a formula φ.

- Unfortunately, $\mathcal{L}_{apal}$ is shown to be undecidable by encoding the tiling problem into the SAT-problem for $APAL$ (see [French and van Ditmarsch, 2008]).

3.1.3 Not All Announcements Are Public and Made by an Omniscient Agent - Action Model Logic

Two gaps appear in the previous work to those willing to use logic to consider multi-agent systems, or only situations that involve different persons. First, if not any truth can necessarily be known, in the formalisms we presented until now any truth *can be announced.* But if announcements can be made by actual agents (and not only by an omniscient one), it appears clearly that for an agent to be able to say truthfully something, she has at least to *know* it, condition not always satisfied. Second, communication is not limited to public announcements: there are private announcements between two agents or inside a group of agents, other agents may see that a communication is taking place or not, etc...

[Baltag and Moss, 2004] propose the following formalism, called *action model logic*, that generalises public announcement logic, including many kinds of informative events[1]. An event is not only a formula publicly announced, but represents the uncertainty of the different agents about the announcement that is actually taking place. Follows an example.

[1] We use the terminology of "event" instead of "action" that supposes a notion of will that we do not have.

Return to the situation presented in Figure 3.2. Suppose that Cha shows her card to Brune. Alex does not know which card she is showing, but he can see she shows her card, so he knows that Brune and Cha know which card it is. Baltag et al. suggest the following *epistemic event model*. Each state is a deterministic epistemic event, with a precondition that ensures that it is executable. As for epistemic models, a link indexed by an agent i between two events means that agent i cannot distinguish these two events. In our example the representation of the event is the following:

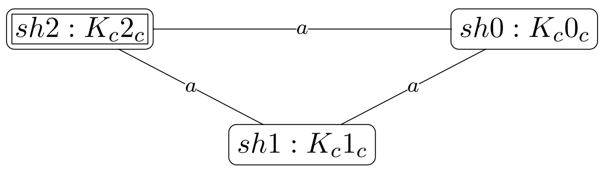


Figure 3.4: Cha shows her card to Brune

There are three possible events corresponding to the fact that Cha shows her card to Brune. The two girls know that Cha showed 2, and Alex consider two possible events (Cha showing 1 or Cha showing 2). The third one (Cha showing 0) is also in their collective imagination. Indeed, Cha can imagine that Alex does not have 0, and leading him to consider posible that Cha shows her 0. Each of those events has a precondition (that is Cha having 0, 1 or 2 respectively). In each case Brune and Cha know exactly which event is taking place. All of that is common knowledge between the three agents (because the event in itself is public). What would be the result of such a complex event on our initial model is presented in figure 3.5.

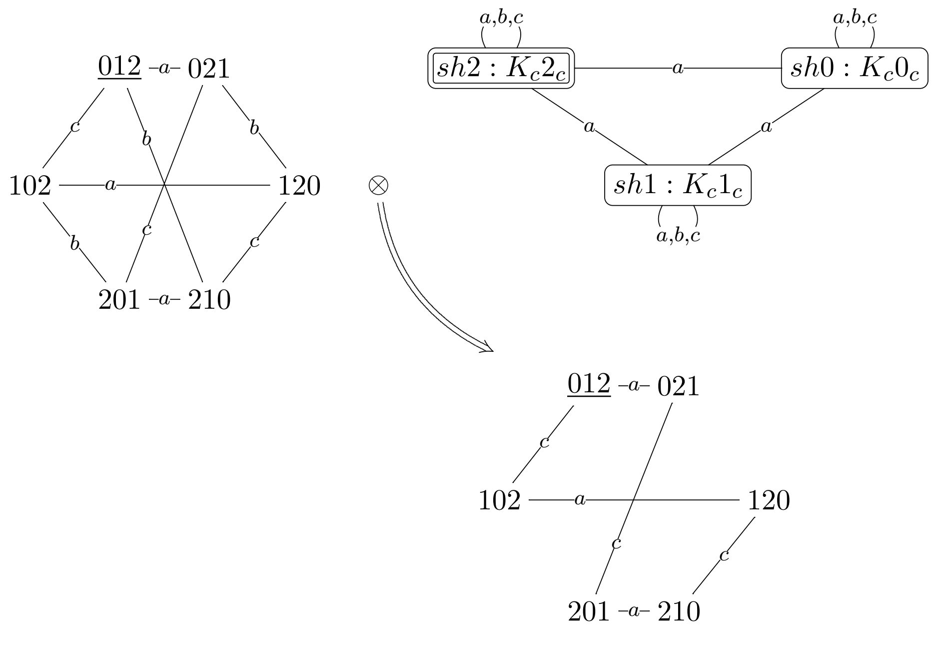


Figure 3.5: Result of the event

After this example, we shall define precisely this formalism.

Definition 3.6 (Event Model) *An epistemic event model over a countable set of agents AG is a triple $\mathcal{E} = (E, \longrightarrow_{AG}, pre)$ where E is a set (of* simple events*), $\longrightarrow_{AG} = \{\longrightarrow_i \mid i \in AG\}$ is a AG-indexed family of binary relations on E and $pre : E \longrightarrow \mathcal{L}_{el}$.*

The main idea is that 'simple' events are deterministic events, that is that any observer can deduce the impact of one event on one state. Uncertainty about which deterministic event is actually taking place is added to the concept of simple event, creating an event model. The event model thus created is independent from the actual state model, and the uncertainty about which event is taking place is thus independent from the uncertainty that agents may have about the actual state of the world. Each simple event can be executed only in the states where its precondition is true. As they are epistemic and deterministic events, we can consider that the precondition is the information carried by the event itself.

We define the model obtained from an initial one by executing an event model:

Definition 3.7 (Resulting model) *Let $\mathcal{M} = (S, \mathcal{R}, V)$ be a state model and $\mathcal{E} = (E, \longrightarrow_{AG}, pre)$ be an event model. The model resulting from $\mathcal{M}$ by application of the event model $\mathcal{E}$ is the following: $\mathcal{M} \otimes \mathcal{E} = (S', \mathcal{R}', V')$ with*

- $S' = \{(s, e) \mid s \in S, e \in E \text{ and } \mathcal{M}, s \models pre(e)\}$
- $(s_1, e_1) R'_i (s_2, e_2)$ *iff* $(s_1 R_i s_2 \text{ and } e_1 \longrightarrow_i e_2)$
- $(s, e) \in V'(p)$ *iff* $((s, e) \in S' \text{ and } s \in V(p))$

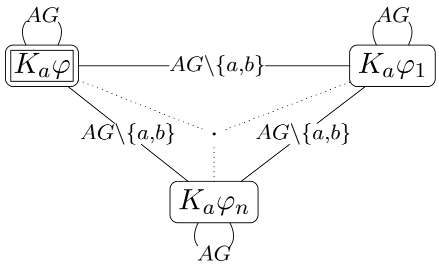
As we will see in further examples, event model logic allows us to add a notion of agency to announcements. Indeed, in public announcement logic, any announcement is made by an exterior agent, let us say the modeler herself. Such an announcement is commonly known to be truthful by the whole set of agents. Such a condition on announcements is a serious limit to model situations in which agents communicate.

Here are different kinds of private communication using event models. In these models, agent a is the agent speaking, and the information he gives is φ, as the information exchanged is of the form $K_a\varphi$ "agent a knows φ":

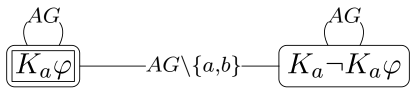
Public announcement: $\begin{tikzcd} \boxed{\boxed{K_a\varphi}} \arrow[loop above, "AG"] \end{tikzcd}$

Hidden announcement to b: $\begin{tikzcd} \boxed{\boxed{K_a\varphi}} \arrow[loop above, "AG"] \arrow[r, "AG\setminus\{a,b\}"] & \boxed{\top} \arrow[loop above, "AG"] \end{tikzcd}$

Visible private announcement:



Idem with Common Knowledge on subject:



Let us define precisely the syntax and semantics of DEL:

Definition 3.8 ($\mathcal{L}_{aml}$) *The language of action model logic $\mathcal{L}_{aml}$ over a countable set of agents AG and a countable set of atoms $PROP$, given an epistemic event model $\mathcal{E} = (E, \longrightarrow_{AG}, pre)$ is defined as follows*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid K_i\varphi \mid [a]\varphi$$

where $i \in AG$ and $p \in PROP$ and $a \in E$.

$[a]\varphi$ is read 'after any execution of action a, φ becomes true. We note $\langle a\rangle$ the dual of $[a]$, *id est* $\langle a\rangle\varphi := \neg[a]\neg\varphi$. We interpret this language in the class of all Kripke models, an event model $\mathcal{E}$ being given. Using Definition 3.7 we can precise the interpretation of this new modality $[a]$ (the other ones remaining identical):

$$\mathcal{M}, s \models [a]\varphi \text{ iff } (\mathcal{M}, s \models pre(a) \text{ implies } \mathcal{M} \otimes \mathcal{E}, (s, a) \models \varphi).$$

The following Definition details how to consider the succession of two epistemic events:

Definition 3.9 *Let $\mathcal{A} = (A, \longrightarrow^{\mathcal{A}}, pre^{\mathcal{A}})$ and $\mathcal{B} = (B, \longrightarrow^{\mathcal{B}}, pre^{\mathcal{B}})$ be two events models over a same set of agents AG. Then we define $\mathcal{A} \otimes \mathcal{B} = \mathcal{C} = (C, \longrightarrow^{\mathcal{C}}, pre^{\mathcal{C}})$ where*

- $C = A \times B$
- *for all $i \in AG$ and all $(a, b), (a', b') \in C$: $(a, b) \longrightarrow_i^{\mathcal{C}} (a', b')$ iff $(a \longrightarrow_i^{\mathcal{A}} a'$ and $b \longrightarrow_i^{\mathcal{B}} b')$*
- *for all $(a, b) \in C$, $pre(a, b) = pre(a) \wedge \langle a\rangle pre(b)$.*

3.1.4 Objective Beliefs

Turn back to the *Texas Hold'em* poker game of Alex and Brune. We saw how cards are dealt, until five shared cards are put on the table. Each player is then able to define her own hand, composed of the best 5-cards combination between the two cards of her game and the five shared cards on the table. The order of the combinations is the following, from the least to the best: high cards, one pair, two pairs, three of a kind, straight, flush, full house, four of a kind ('poker') and straight flush[2].

Though our two players do not know the value of the cards that will be on the table, this value is fixed from the beginning of the game and cannot be modified by any action of the players. They can only update their beliefs about these objective facts.

In such a situation, the player can be wrong about the belief of the other player ("Alex believes Brune believes she has the winning hand"), but their beliefs regarding to the cards dealt on the table are true beliefs. We present in this section a framework proposed by [Hommersom et al., 2004] that allows us to express this notion of objectivity of some beliefs, that ensures that such an objective belief is true. We first first define the language of this logic:

Definition 3.10 (Syntax of $\mathcal{L}_{lob}$) *The language of the logic of objective beliefs $\mathcal{L}_{lob}$ over a countable set of agents AG and a countable set of propositional atoms $PROP$ is defined as follows:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid B_i\varphi$$

where $i \in AG$ and $p \in PROP$.

$B_i\varphi$ is read "agent i believes that φ". We use the usual abbreviations, in particular $\hat{B}_i\varphi := \neg B_i\neg\varphi$ ("agent i considers possible that φ"). We say that a formula of $\mathcal{L}_{lob}$ is boolean if there is no occurrence of the operator of belief in it. We interpret semantically this language on Kripke models $\mathcal{M} = (S, R, V)$, the satisfiability relation $\mathcal{M}, s \models \varphi$ being defined just as usual. More precisely:

- $\mathcal{M}, s \models p$ iff $p \in V(s)$,
- $\mathcal{M}, s \not\models \bot$,
- $\mathcal{M}, s \models \neg\varphi$ iff not $\mathcal{M}, s \models \varphi$,
- $\mathcal{M}, s \models \varphi \vee \psi$ iff $\mathcal{M}, s \models \varphi$ or $\mathcal{M}, s \models \psi$,
- $\mathcal{M}, s \models B_i\varphi$ iff for all $t \in S$, sR_it implies $\mathcal{M}, t \models \varphi$.

The notion of validity upon a model or a class of models is defined as usual. These definitions are exactly the same as the equivalent definitions for epistemic logic. The difference appears when considering the properties of the models. Indeed, we do not interpret $\mathcal{L}_{lob}$-formulas in $S5_n$ models but in the class $\mathcal{C}_0$ of transitive, euclidian and o-serial models, defined as follows:

Definition 3.11 (O-serial model) *We say that a Kripke model $\mathcal{M} = (S, \mathcal{R}, V)$ is o-serial if for all $s \in S$, there exists $t \in S$ such that sR_it and $V(s) = V(t)$. We call $\mathcal{C}_0$ the class of models that are transitive, euclidian and o-serial.*

[2]The entire set of rules can be found, for example, in http://en.wikipedia.org/wiki/Texas_hold_'em.

Indeed, this class of models bring the property we expect for objective belief: an agent cannot believe a false propositional formula. Therefore, she may not know which is the actual state, but she always consider possible a state in which the objective facts have the same truth value as in the actual one, *i.e.* a state that has the same valuation.

3.1.5 Partial Conclusion

In this section, we presented different works that use epistemic logic, the logic of knowledge, to express dynamic situations. The most used is the notion of public announcement, exchange of information between the modeler and the whole set of agents. Though Action Model Logic gives us a framework that allows to speak about private announcement, some more notions should be improved in order to tackle the problem of 'the right to say':

- What kind of information is given? We presented the distinction between beliefs and objective beliefs. What kind of technical result can we get? Can we add to this logic the notion of arbitrary announcement, like in $\mathcal{L}_{apal}$? We explore these questions in Chapter 4.

- Who is speaking? Is information given by an individual agent or is it given by a group of agents? How can we formalize the notion of group announcement? This is the topic of Chapter 5.

3.2 Deontic Logics

Dynamic epistemic logics are a good starting point to understand the notion of 'right to say something': if the content of the speech is the relevant element that determines the right, then 'to say something' can be interpreted as 'to give a piece of information'. But the notion of 'right' still has to be interpreted formally.

Deontic logic gives formal interpretations to the notions linked to permission and obligation. 'Must', 'permitted', 'optional', 'ought', 'should', 'obligatory', 'might', or 'forbidden' are classical notions that deontic logic tries to formalize. Obviously, one cannot pretend to give a unique interpretation of these notions: they may in particular depend on the kind of obligation they express (moral, hierarchical, political...), but also on the nature of the object of such obligation. Indeed, this subject may be a state of affair or a given action. The former would be represented by formulas of the form $F\varphi$, where φ is a formula of the language, which says that a situation in which φ is true is forbidden; the latter by some $F\alpha$ where α represents an action – and how to represent such notions? – and not a formula.

We present in Section 3.2.1 classical deontic logic, which is a modal logic with a similar language than epistemic logic's but with another interpretation. In this standard theory, the deontic modalities are applied to formulas. In section 3.2.2 we present the logics of permissions, for which obligation and prohibition are applied to actions, opening the way to what is our proposal, presented in Chapters 6 and 7.

3.2.1 Standard Deontic Logic

Like many logics, the classical deontic logic is rooted in Antiquity and the Middle Ages, e.g., in the *Obligatio* game/procedure (see [Boh, 1993]). This game, that can be seen as a logical game of counterfactual reasoning (see [Spade, 1982]) or some kind of training on thesis defence (see [Spade, 1992]), is based on arguments that have to follow some rules. *Obligationes* are not deontic thoughts, they are not *obligations*, but they led to thought on deontic concepts because they carried the idea of rules and thus of obligations in a formal frame.

Since then, deontic logics have been developed on a twofold way, both non-modal and modal, namely [Mally, 1926] and [von Wright, 1951]. Let us present briefly the Standard Deontic Logic based on von Wright work, and its limitations in our context.

Standard Deontic Logic (SDL) is a modal logic whose operators formalize the following basic notions:

- it is obligatory that φ : $O\varphi$
- it is permitted that φ: $P\varphi$
- it is forbidden that φ: $F\varphi$
- it is omissible that φ: $OM\varphi$.
- it is indifferent (or optional) that φ: $I\varphi$.

It is built upon propositional logic. In this standard framework, all these modalities can be expressed using only one of them, typically the first one, obligation. This framework is a normal KD logic.

Syntax, Semantics and Properties

Definition 3.12 (Syntax of $\mathcal{L}_{sdl}$) *The language of the standard deontic logic $\mathcal{L}_{sdl}$ over a countable set of propositional atoms $PROP$ is defined as follows, where $p \in PROP$:*

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid O\varphi.$$

We use the classical propositional abbreviations and define the following ones: $P\varphi := \neg O \neg\varphi$, $F\varphi := O\neg\varphi$, $OM\varphi := \neg O\varphi$, $I\varphi := \neg O\varphi \wedge \neg O \neg\varphi$.

We interpret this language on the class of serial Kripke models, using the possible world semantics. That implies the following properties

- $\models O(\varphi \longrightarrow \psi) \longrightarrow (O\varphi \longrightarrow O\psi)$

- $\models O\varphi \longrightarrow P\varphi$
- If $\models \varphi$ then $\models O\varphi$.

We call SDL the normal KD axiomatization of this logic summed up in Table 3.2.

instantiations of propositional tautologies	
$O(\varphi \longrightarrow \psi) \longrightarrow (O\varphi \longrightarrow O\psi)$	distribution
$P\top$ or $\neg O\varphi \vee \neg O\neg\varphi$	D
From φ and $\varphi \longrightarrow \psi$ infer ψ	modus ponens
From φ infer $O\varphi$	necessitation

Table 3.2: Axiomatization SDL

SDL is sound and complete with respect to the class of serial models (where we interpret the language) (see [Chellas, 1980]).

Moreover, it has been proved (see for example [Blackburn et al., 2001]) that:

- $\mathcal{L}_{sdl}$ is decidable, and the decision procedure is $PSPACE$-complete
- The model checking problem for $\mathcal{L}_{sdl}$ is in P

Classical Paradoxes

Standard deontic logic is famous for its 'paradoxes'. Here are some of the most known (cited from [Meyer et al., 1994]):

1.	Empty normative system	$\models O\top$
2.	Ross' paradox	$\models O\varphi \rightarrow O(\varphi \vee \psi)$
3.	No free choice permission	$\models P(\varphi \vee \psi) \leftrightarrow P\varphi \vee P\psi$
4.	Penitent's paradox	$\models F\varphi \rightarrow F(\varphi \wedge \psi)$
5.	Good Samaritan paradox	$\models \varphi \rightarrow \psi$ implies $\models O\varphi \rightarrow O\psi$
6.	Chrisholm paradox	$\models (O\varphi \wedge O(\varphi \rightarrow \psi) \wedge (\neg\varphi \rightarrow \neg O\psi) \wedge \neg\varphi) \rightarrow \bot$
7.	Forreseter's paradox of gentle murder	$\models \psi \rightarrow \varphi$ implies $\models (F\varphi \wedge (\varphi \rightarrow O\psi) \wedge \varphi) \rightarrow \bot$
8.	Conflicting obligations	$\models O\varphi \rightarrow \neg O\neg\varphi$
9.	Derived obligations	$\models O\varphi \rightarrow O(\psi \rightarrow \varphi)$
10.	Deontic detachment	$\models (O(\varphi) \wedge O(\varphi \rightarrow \psi)) \rightarrow O\psi$

The normal modal interpretation of deontic concepts imposes these validities, perceived by deontic logicians as paradoxes. But that comes in our mind from a erroneous interpretation of the subjects of the obligation, that are in this formalism propositions and not actions.

For example, the strangeness of the second validity is often illustrated by the following example: 'if you are obliged to read a letter, you are obliged to read it or to burn it'. It is clearly a counterintuitive sentence. But the example supposes that ψ and φ are actions ('to burn the letter') while they are actually propositions. A better example would thus be 'if the letter must be red, then it must be red or green'. It is still a strange sentence, but its strangeness does not come from the implication, but from the fact that the color of the letter is obligatory...

In the same way, here are some sentences that reveal the claimed nature of paradox of each validity:

3. If someone is allowed to hit his dog, he is allowed to hit his dog or his boss.
4. If someone is forbidden to commit a crime, he is forbidden to commit a crime and to repent.
5. The doctor operates her wounded patient implies that the patient is wounded, thus if it is obliged that the doctor operates her wounded patient then it is obliged that the patient is wounded
6. You are obliged to go to a party; it is obliged that, if you go, you tell you are coming; if you do not go, you are obliged not to tell you are coming. In this situation, we can affirm that you go to the party!
7. One is forbidden to murder; still, if one murders someone, one has to do it gently; moreover, a gentle murder implies a murder. But in this situation murders are impossible!
8. There are no conflicting obligations (sic)
9. If a child is obliged to brush her teeth, then it is obliged that if martians exist the child brushes her teeth.
10. Same as 5.

As we can see, all these paradoxes come essentially from the fact that it seems that deontic norms are applied to actions, while they actually applied to formulas. To be able to formalize concepts of obligation in a context of acting agents we need some kind of dynamic modality that represents actions, as in our proposals (Sections 6 and 7). The next section presents some of those already existing frameworks.

3.2.2 Dynamic Logics of Permission

As we have seen, to dynamic logicians (and in particular to dynamic epistemic ones), obligations and permissions clearly apply to actions. It seems thus strange that people associate those with static observations, and 'confuse' the non-deterministic choice between two actions with the disjunction of two propositions. For deontic logic this frame of mind was reset by John-Jules Meyer with his *different approach to deontic logic* (see [Meyer, 1988]), an approach that was later followed up by [van der Meyden, 1996], the starting point for some proposals in this thesis.

Both of these works are adaptations of propositional dynamic logic (PDL) presented in [Fischer and Ladner, 1979]: PDL allows to represent at the same time the truth value of propositions (*i.e.* objective facts) and the effect of actions on them. In other words, we can model in the same framework static situations and dynamic transitions.

In the language $\mathcal{L}_{pdl}$ we distinguish *assertions* (that describe states of facts) from *actions* (that describe transitions between states of facts). The set of actions is inductively constructed over a countable set of atomic actions Act in the following way:

$$\alpha ::= a \mid \alpha; \beta \mid \alpha \cup \beta$$

where a is an atomic action. Action $\alpha; \beta$ describes the succession of action α and action β, action $\alpha \cup \beta$ is the non-deterministic choice between those two actions. Figure 3.6 gives an example of model of PDL, that we define afterwards. In this example Brune initially receives a letter from Cha, and we look at Brune's possible actions.

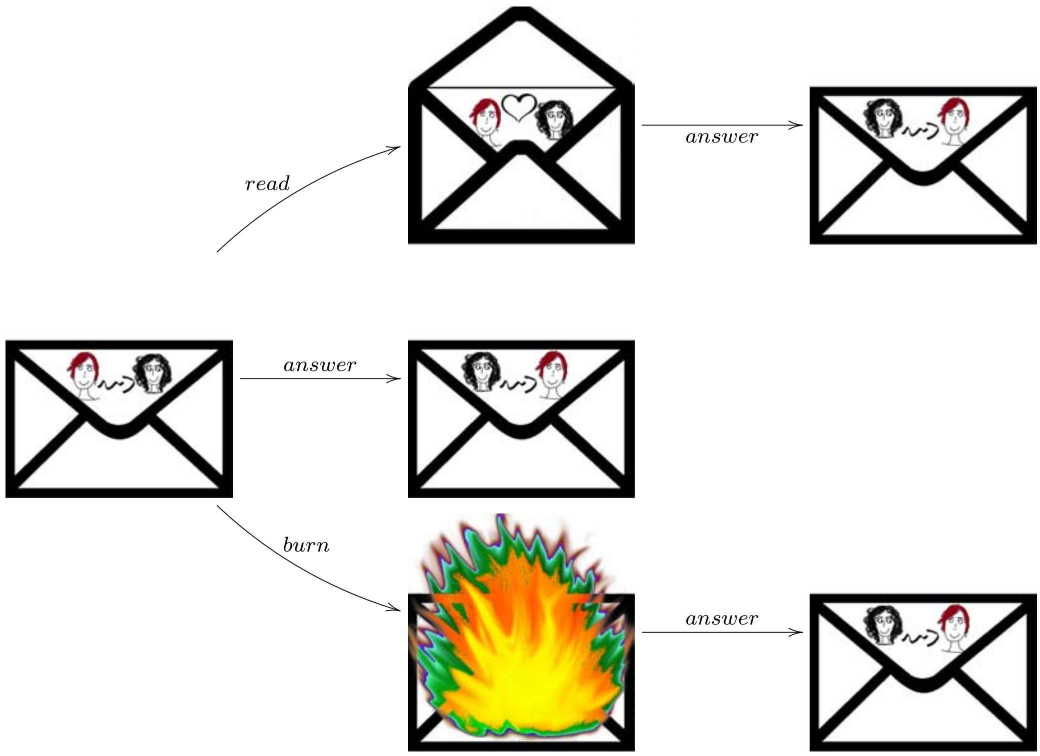


Figure 3.6: Brune actions while receiving a letter from Cha

Brune has three possible actions $read$, $answer$ and $burn$. Therefore '$read; answer$' corresponds to the fact that she reads the letter and then answers to her friend, '$read \cup burn$' the fact that she does one of the two actions.

Now that the notion of action is introduced, we define the language $\mathcal{L}_{pdl}$ as the following set of assertions over the set of actions and a countable set of propositional atoms $PROP$ by the following:

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi \vee \psi \mid [\alpha]\varphi$$

We read $[\alpha]\varphi$ as "It is necessary that after executing α, φ is true". Its dual, $\langle\alpha\rangle\varphi := \neg[\alpha]\neg\varphi$ can be read as "There is an execution of α after which φ is true".

Here are some examples of formulas we can express in this language:

- $[read]letter\_is\_open$: 'after Brune having read the letter, it is open'
- $\langle read \cup burn\rangle(letter\_is\_open \vee letter\_is\_burnt)$: there is an action between read and burn that Brune can do, and after having executed it the letter is open or burnt.

Actually, $\mathcal{L}_{pdl}$ admits two more constructions of action:

- the test of φ, noted $\varphi?$, that cannot be executed if φ is false and has no effect if φ is true.
- the iterated execution α noted α^*. The number of executions is chosen nondeterministically, *i.e.* $\alpha^* := \top? \cup (\alpha) \cup (\alpha\ ;\alpha) \cup (\alpha\ ;\alpha\ ;\alpha) \cup \ldots$

A model of PDL is a Kripke model $(S, V, \mathcal{R})$ where $\mathcal{R} = \{R_a\}_{a \in Act}$. Figure 3.6 gives an example of such a model, with $Act = \{read, burn, answer\}$. From the set $\mathcal{R}$ of binary relations we can construct a binary relation R_α for every action α by induction on the structure of α in the following way:

- for all $a \in Act$, R_a is already defined
- for all $\psi \in \mathcal{L}_{pdl}$, $R_{\psi?} = \{(s,s) \mid s \in S \text{ and } \mathcal{M}, s \models \psi\}$
- $R_{\alpha;\beta} = \{(s,t) \in S^2 \mid \exists u \in S \text{ s.t. } (s,u) \in R_\alpha \text{ and } (u,t) \in R_\beta\}$
- $R_{\alpha\cup\beta} = R_\alpha \cup R_\beta$
- $R_{\alpha^*} = \{(s,t) \in S^2 \mid \exists t_0, \ldots, t_n \in S \text{ s.t. } s = t_0,\ t = t_n, \text{ and for all } 0 \leqslant i \leqslant n-1,\ (t_i, t_{i+1}) \in R_\alpha\}$

The semantics of the dynamic operator $[\alpha]$ is thus defined in the following way:

- for all action α, $\mathcal{M}, s \models [\alpha]\varphi$ iff for all $t \in S$, $sR_\alpha t$ implies $\mathcal{M}, t \models \varphi$

The following is an equivalent definition of the semantics of this dynamic operator, by induction on the structure of the action considered:

- for all $a \in Act$, $\mathcal{M}, s \models [a]\varphi$ iff for all $t \in S$, $sR_a t$ implies $\mathcal{M}, t \models \varphi$
- for all $\psi \in \mathcal{L}_{pdl}$, $\mathcal{M}, s \models [\psi?]\varphi$ iff $\mathcal{M}, s \models \psi \longrightarrow \varphi$
- $\mathcal{M}, s \models [\alpha;\beta]\varphi$ iff $\mathcal{M}, s \models [\alpha][\beta]\varphi$

- $\mathcal{M}, s \models [\alpha \cup \beta]\varphi$ iff $\mathcal{M}, s \models [\alpha]\varphi \wedge [\beta]\varphi$

The definition for the iterated execution of α may thus be understood in the following way:

- $\mathcal{M}, s \models [\alpha^*]\varphi$ iff $\mathcal{M}, s \models \varphi \wedge [\alpha][\alpha^*]\varphi$

The basic idea of Meyer was to add to PDL a special atom *viol* (or its negation *perm*) to the set $PROP$ of propositional atoms. Thus, we would say that an action is permitted if its execution does not lead to a state of violation, *i.e.* a state where *viol* is true. Formally, $P\alpha := [\alpha]\neg viol$. We could have considered another kind of permission, which is weaker: $P'\alpha := \langle\alpha\rangle\neg viol$. In this case, we consider that action α is permitted if there is at least one execution of α that does not lead to a state of violation.

Figure 3.7 illustrates this framework in the same example of Brune acting after having received a (love) letter from Cha. In the figure violations states are marked by a big V. As depicted in the figure, we consider that both 'burning the letter' and 'answering it (without having read it)' lead to a violation state. We consider also that if Brune, after having read the letter, answers NO then the resulting state violates the rule. We consider in this situation the action $answer := answer\_YES \cup answer\_NO$

Here are some formulas that are true in the initial state using the notation P for the strong permission and P' for the weak one :

- $P(read)$: 'Brune is permitted to read Cha's letter'
- $\neg P(burn) \wedge \neg P(answer)$: 'Brune is neither permitted to burn Cha's letter nor to answer it (without having first read it)'
- $\langle read\rangle(P'(answer) \wedge \neg P(answer))$: 'After having read Cha's letter, Brune is weakly -but not strongly- permitted to answer it'. Indeed, there is an execution of $answer$ (namely $answer\_YES$) that is permitted, but not all ($answer\_NO$ is not).
- $\langle burn\rangle P(answer)$: 'After having burnt the letter, Brune is allowed to answer it'

Van der Meyden noted a limit in Meyer's work: an action is permitted or not depending exclusively on its resulting state of affairs. This may seam reasonable, but it brings about some counterintuitive implications, as highlighted by this last example. In fact, it would validate the following sentence: "If after having burnt the letter Brune is allowed to answer it, then Brune has the right to burn the letter and then answer it". Indeed, it would be translated by the following: $\langle burn\rangle Panwer \longrightarrow P(burn\ ; answer)$. This is clearly a validity of Meyer's models, inasmuch as the resulting state of affairs after executing $burn$ and then $answer$ is the same as the state of affairs resulting from the execution of $(burn; anwer)$. If the reader is not convinced that this result is counter-intuitive, the more famous following example should definitely convince her: "If after shooting the president one has the right to remain silent then one has the right to shoot the president and then remain silent".

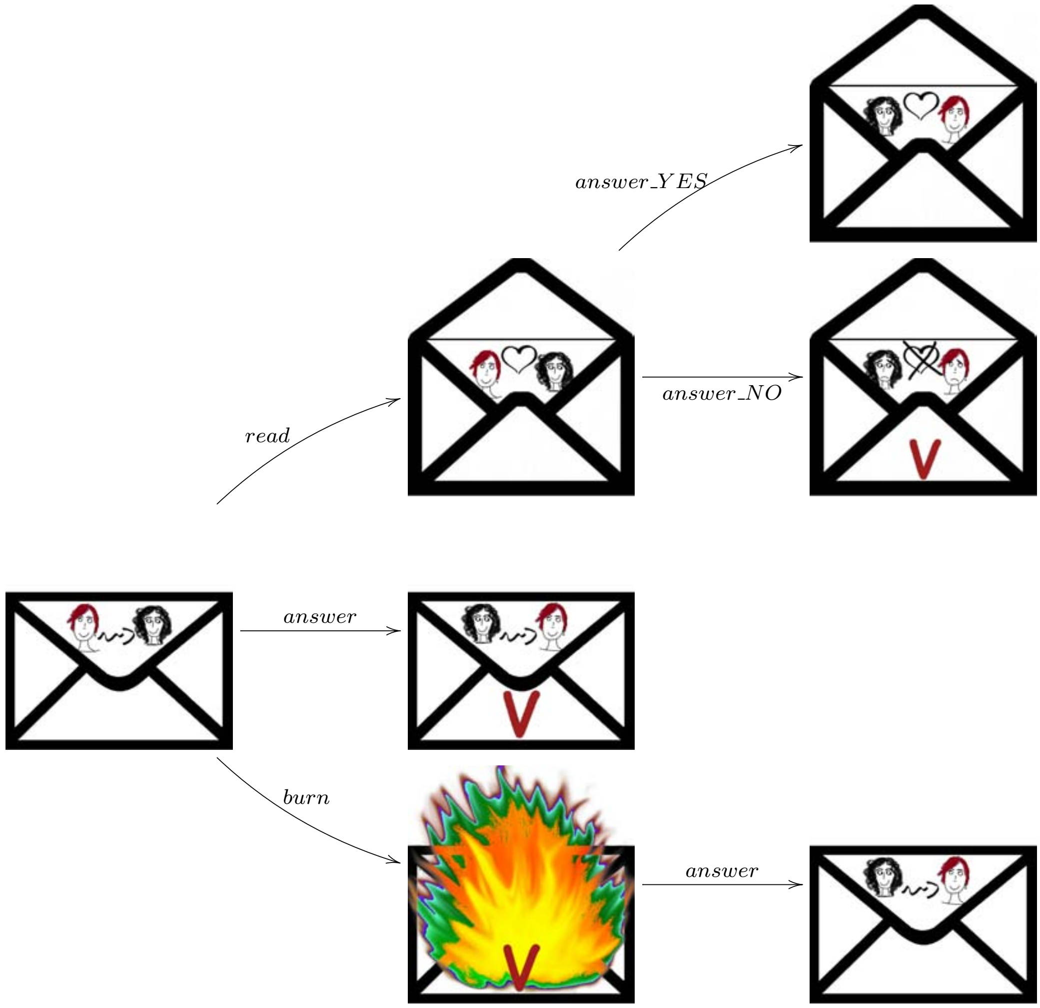


Figure 3.7: Considering states of violation for Brune's actions

To solve this problem, Van der Meyden proposed to label transitions (*i.e.* the atomic actions) instead of worlds (*i.e.* the resulting state of affairs). Formally, Van der Meyden's logic is an adaptation of PDL in which the models contain a special set $\mathcal{P} \subseteq S \times \mathbf{Act} \times S$. $\mathcal{P}$ is the set of permitted transitions: a triple (s, a, s') is in $\mathcal{P}$ iff the transition labelled by a from s to s' is permitted. The same example as before is represented in this framework by Figure 3.8. The plain (and green) transitions are the permitted ones, the dotted (and red) are the ones that are not in $\mathcal{P}$.

The syntax of this language contains the following construct $\Diamond(\alpha, \varphi)$ which means "*there is a way to execute action α which is permitted and after which φ is true*". In the example, in the initial situation, we have $\Diamond(read, \top)$ and $\Diamond((read; answer), \top)$. He also introduces a weak form of obligation $\mathcal{O}(\alpha, \varphi)$. The meaning of $\mathcal{O}(\alpha, \varphi)$ is "*after any permitted execution of α, φ is true*". This formalization allows to consider situations in which two different actions that end in the same state of affairs are not permitted in the same way. Section 7.3 presents

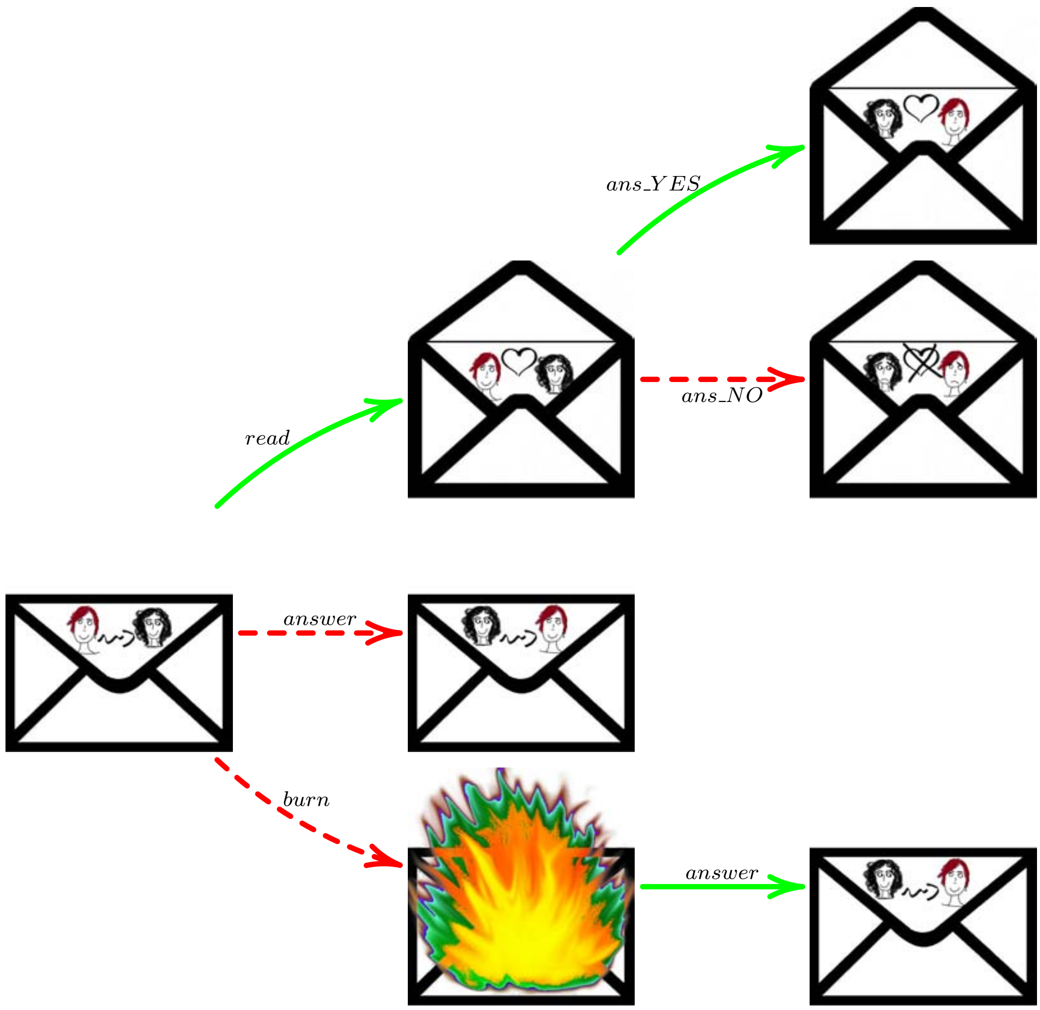


Figure 3.8: Considering labelled transition for Brune's actions

an example of such a situation.

3.3 Permission and Epistemic Actions

To formalize the notion of 'right to say' we are particularly interested in frameworks that consider the permission on epistemic events, as presented in Section 3.1. Such frameworks already exist.

[van Benthem et al., 2009] (see also [Hoshi, 2008]) propose a logic for protocols in dynamic epistemic logic that can be interpreted as a logic for permitted events – and in particular permitted announcements. A protocol is a set of sequences of events, and an announcement is an example of such an event; "*being in the protocol*" can therefore be understood as "*being permitted to be said*". One purpose of this publication was to merge epistemic temporal logic - [Parikh and Ramanujam, 2003] - with dynamic epistemic logic - [Baltag et al., 1998, van Ditmarsch et al., 2007]. The axiomatization of the language with added protocols is facilitated by the translation of the latter into the former.

We only present what [van Benthem et al., 2009] call the *forest* generated by a pointed epistemic model $(\mathcal{M}, s)$ and sequences of announcements. It corresponds to all the models we may obtain from the initial epistemic model by applying the considered announcements. In fact, the announcements considered are pointed event models (see Definition 3.6), and we face sequences of such epistemic events. We call *protocol* a prefix-closed set of such sequences. We call *history* a set consisting of such sequences preceded by a state in the epistemic model wherein they are executed. For example, given an initial state s, and say a sequence of first $\psi!$ and then $\varphi!$ as allowed according to protocol, we write $s\psi\varphi$ for that history: the announcements in sequence are simply written one after the other.

Relative to a protocol Π we can construct a temporal epistemic model $\mathcal{M}_\Pi$ that contains the initial model and all the models obtained from it by applying a sequence of epistemic events belonging to Π. We can then express the following: 'in the context $(\mathcal{M}, s)$, given the protocol Π and after having past the history h (that starts from s), it is permitted to say ψ after which φ is true'. It would be translated by $\mathcal{M}_\Pi, h \models \langle\psi\rangle\varphi$ and the semantics gives us the following: $\mathcal{M}_\Pi, h \models \langle\psi\rangle\varphi$ iff:

- $\mathcal{M}_\Pi, h \models \psi$
- $h' = h\psi \in \Pi$
- $\mathcal{M}_\Pi, h' \models \varphi$.

[Aucher et al., 2010] also propose to merge deontic concepts with epistemic ones. They start from the *Epistemic Deontic Logic* (EDL) which language contains an epistemic operator K - as in EL - and a deontic one O - as in SDL -. Note that it is a single-agent framework. We can thus express, for example $OK\varphi \wedge KO\neg K\psi$: 'It is obligatory that the agent knows φ and the agent knows it is obligatory that she does not know ψ'. In fact, following [Castaneda, 1981], they distinguish formulas depending on whether they are within or without the scope of deontic operators. In order to skip details, just assume that the deontic operators apply to the knowledge of the agent: if p is an atomic proposition, we can express OKp ('the agent is *obliged to know* p') but not Op ('it is obligatory that p'). They call P (for *permitted*) the dual operator of O: $P\varphi := \neg O\neg\varphi$. They interpret this language in models in which there is an $S5$-relation R_K for the epistemic operator, and a KD-relation R_O for the deontic operator.

Therefore, they call *epistemic norm* a formula of the form $\psi \longrightarrow O\varphi$ or $\psi \longrightarrow P\varphi$ and *privacy policy* a set of epistemic norms. A pointed EDL-model is said *compliant* with a given privacy policy if all the epistemic norms it contains are satisfied in the given state of the model.

From there, they develop a new logic called *Dynamic Epistemic Deontic Logic* ($DEDL$) adding to the language of EDL a dynamic operator $[send\ \varphi]$ (in fact there is also another operator $[prom\ \varphi]$ that we skip here). It is a dynamic epistemic operator that can be understood as the announcement (by the modeler) of φ. They can thus express for example the following: $[send\ p](Kp \wedge OKq)$: 'after the announcement of p, the agent knows p and it is obligatory that she knows q'. In this logic, they propose to formalize the notion of 'right to say' in the following way: it is permitted to announce something if the result of such an announcement is compliant with the given privacy policy. We denote as $P\varphi!$ such a permission *to announce* φ (and no more *to know* φ). Therefore, let PP be a privacy policy and $\bigwedge PP$ the conjunction of all the epistemic norms appearing in PP, then we have $\models_{EDL} P\varphi! \longleftrightarrow [send\ \varphi](\bigwedge PP)$.

Yet, these two frameworks have limitations that justify to develop our proposal (See Chapters 6 and 7). Comparisons between our work and theirs are proposed in Section 6.5.2. for [van Benthem et al., 2009] and in Section 8.1 for [Aucher et al., 2010].

Chapter 4

Logics of Objective Beliefs

4.1 Introduction

The Logic of Objective Beliefs (LOB) has been presented in section 3.1.4. We saw that 'objective belief' is a useful notion to formalize situations in which some information is observed by the agents. If an agent believes something about such observation, then it is true. As we saw, a typical situation is the *Texas Hold'em* Poker: some cards are known to be seen by the agents nd some others are revealed publicly on the table. In this section, we present new results regarding this framework, while extending it with some useful notions. Let us enter inside the game!

After the initial deal and the *flop*, the players have to bet three times; each bet is followed by the deal of an additional card. During these bets, even if she thinks her hand is losing, Brune could try to bluff. She would not do it if she thinks that Alex is certain to have the winning hand. In this framework, we would like to represent the fact that an agent believes something ("I have the winning hand") whatever she may observe later.

Let us suppose that, while cheating, Alex sees Brune's hand without her knowing. He would have a serious advantage on his opponent. In particular, he knows who has the winning hand, and he can deduce if his opponent is sure or not that her hand is the winning one. How could we represent an update function that allows to distinguish between the deal of a card and a cheating access to new information?

This section, which principal results have been published in french in [Balbiani and Seban, 2008], complete these requirements. We first present the Logic of Objective Beliefs with an update function expressible in this logic, with a slightly different semantics from [Hommersom et al., 2004]. We can express sentences like " After the *flop* deal, Brune considers possible that she has a losing hand" or "Alex believes that if the *river* is the ace of spades, he will have a winning hand". Secondly, we present the concept of arbitrary update of objective beliefs. We can thus express, for example, the following: "Brune can learn something that will let her know she has the winning hand, but whatever Alex learns about the deal, he would still consider possible that Brune considers possible she has the losing hand".

4.2 Some Properties of the Logic of Objective Beliefs

First we present the following axiomatization LOB of this logic, according to the intuition we have about objective beliefs, then we show that this axiomatization is sound and complete on a certain class of models.

(CPL) axioms of classical propositional logic,

(K) $B_i(\varphi \to \psi) \to (B_i\varphi \to B_i\psi)$,

(4) $B_i\varphi \to B_iB_i\varphi$,

(5) $\hat{B}_i\varphi \to B_i\hat{B}_i\varphi$,

(T_{bool}) if φ is boolean, then $B_i\varphi \to \varphi$.

As we exposed before, axioms (4) and (5) express the introspective character of belief: "if agent i believes that φ, then she believes that she believes that φ" and "if agent i considers possible that φ, then she believes that she considers possible that φ", for any formula φ. The axiom (T_{bool}) expresses the objective dimension of the belief: for any boolean formula φ (*i.e.* φ corresponds to an objective fact) "if agent i believes that φ, then φ". In other words, we consider that the beliefs of the agents regarding the real world are consistent with it. It corresponds to the property of *o*-seriality. The theorems of LOB are all the formulas deductible from the axioms using the following deduction rules:

(MP) if φ is a theorem and $\varphi \to \psi$ is a theorem then ψ is a theorem,

(GD) if φ is a theorem then $B_i\varphi$ is a theorem.

Following [Hommersom et al., 2004], we consider the class $\mathcal{C}_0$ of models $\mathcal{M} = (S, R, V)$ in which for every agent i, R_i is transitive, euclidian and *o-serial*. This notion of *o*-seriality has been presented in Definition 3.11.

Proposition 4.1 shows soundness and completeness of logic LOB with respect to the class of models $\mathcal{C}_0$, a result that does not appear in [Hommersom et al., 2004].

Proposition 4.1 *Let $\varphi \in \mathcal{L}_{lob}$. Then φ is a theorem of LOB iff φ is valid in any model of the class $\mathcal{C}_0$.*

Proof To show soundness (*i.e.* the direct implication), it is sufficient to show that axioms are valid and that the inference rules preserve validity. First, we show that o-serial models validate schema T_{bool}. Indeed, let $\mathcal{M}$ be an o-serial model and let s be a state of $\mathcal{M}$. Then if $\mathcal{M}, s \models B_i\varphi$, then there exists $t \in S$ such that (1) sR_it and (2) $V(s) = V(t)$. (1) implies that $\mathcal{M}, t \models \varphi$. Combined with (2), recording that φ is boolean, we obtain that $\mathcal{M}, s \models \varphi$. Second, as we saw in section 2.2, axioms (CPL), K, 4 and 5 are valid in the class of transitive and euclidian Kripke models. Third, (MP) and (GD) preserves validity by definition of the semantics. Indeed, let $\mathcal{M}$ be a model of $\mathcal{C}_0$. On one hand if $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \neg\varphi \vee \psi$ then $\mathcal{M} \models \psi$. On the other hand, if $\mathcal{M} \models \varphi$ then for all $s \in S$ and all $t \in S$ such that sR_it, we have $\mathcal{M}, t \models \varphi$. Therefore for all $s \in S$, $\mathcal{M}, s \models B_i\varphi$. Consequently $\mathcal{M} \models B_i\varphi$. We obtain the wanted result.

Let us now prove completeness (*i.e.* the indirect implication). We define the canonical model $\mathcal{M}^c_{lob} = (S^c, R^c_i, V^c)$ in the classical way (cf. Definition 2.19). Recall in particular that for all maximal consistent sets $x, y \in S^c$ and for all agent i, xR_iy iff $B_i(x) \subseteq y$, where $B_i(x) = \{\varphi \mid B_i\varphi \in x\}$. Let us see that the canonical model is in the class $\mathcal{C}_0$:

- it is transitive: Let x, y, z be such that xR_iy and yR_iz, *i.e.* $B_i(x) \subseteq y$ and $B_i(y) \subseteq z$. It is enough to show that $B_i(x) \subseteq B_i(y)$. Therefore, take $\varphi \in B_i(x)$ then $B_i\varphi \in x$ thus $B_iB_i\varphi \in x$ and $B_i\varphi \in B_i(x)$ which leads to $B_i\varphi \in y$ *i.e.* $\varphi \in B_i(y)$.

- it is euclidean: We use axiom 5, or rather its contrapositive $\hat{B}_iB_i\varphi \longrightarrow B_i\varphi$. Let x, y, z be such that xR_iy and xR_iz, *i.e.* $B_i(x) \subseteq y$ and $B_i(x) \subseteq z$. It is enough to show that $B_i(y) \subseteq z$. Therefore, let $\varphi \in B_i(y)$ then $B_i\varphi \in y$ thus $\hat{B}_iB_i\varphi \in x$. Indeed, if it were not the case, then $B_i\neg B_i\varphi \in x$ and thus $\neg B_i\varphi \in y$ which is false. Thus $\hat{B}_iB_i\varphi \in x$ and by axiom 5 $B_i\varphi \in x$, which means that $\varphi \in B_i(x)$. With $B_i(x) \subseteq z$ we obtain $\varphi \in z$ *Q.E.D.*

- it is o-serial: Let $x \in S^c$ and let $y^\circ = B_i(x) \cup \{p \in PROP \mid p \in x\}$. y° is a consistent set. Suppose the opposite, then there exist $\varphi_1, \ldots \varphi_n \in B_i(x)$ and $p_1, \ldots p_m \in x \cap PROP$ such that $\vdash (\varphi_1 \wedge \ldots \wedge \varphi_n \wedge p_1 \wedge \ldots \wedge p_m) \longrightarrow \bot$, thus $\vdash \neg\varphi_1 \vee \cdots \vee \neg\varphi_n \vee \neg p_1 \vee \cdots \vee \neg p_m$ and $\vdash B_i(\neg\varphi_1 \vee \cdots \vee \neg\varphi_n \vee \neg p_1 \vee \cdots \vee \neg p_m)$. Combining with the fact that $B_i(\varphi_1 \wedge \ldots \wedge \varphi_n) \in x$ we obtain that $B_i(\neg p_1 \vee \cdots \vee \neg p_m) \in x$ which implies that $\neg p_1 \vee \cdots \vee \neg p_m \in x$ by axiom T_{bool}, a contradiction with respect to the hypothesis. So y° is consistent, it can thus be extended to a maximal-consistent set of formulas y (using a classical Lindenbaum lemma proof), that includes $B_i(x)$ and $PROP \cap x$ as y° did. Therefore, xR_iy and for all $p \in PROP$, $y \in V(p)$ iff $x \in V(p)$. We obtain the wanted result.

Now it is sufficient to prove that for all formula $\varphi \in \mathcal{L}_{lob}$, $\mathcal{M}^c_{lob}, x \models \varphi$ iff $\varphi \in x$. Thus, as $\mathcal{M}^c_{lob}$ is a model, a formula that is valid is a validity of this model. It is then in any maximal consistent set of the theory, which proves that it is a theorem. We prove this truth lemma in a classical way, by induction on the structure of φ. This proof is exactly the same as the proof for the logic K (see [Fagin et al., 1995]). Let us analyse the specific induction case where $\varphi = B_i\psi$.

($\Rightarrow$) Suppose that (1) $\mathcal{M}^c_{lob}, x \models B_i\psi$. Thus $K_i(x) \cup \{\neg\psi\}$ is not consistent, and it has a finite subset $\{\varphi_1, \ldots, \varphi_n, \neg\psi\}$ which is not consistent. Therefore, by propositional reasoning, $\vdash \varphi_1 \rightarrow (\varphi_2 \rightarrow (\ldots \rightarrow (\varphi_n \rightarrow \psi) \ldots))$. Hence, using the necesitation rule, we obtain $\vdash K_i(\varphi_1 \rightarrow (\varphi_2 \rightarrow (\ldots \rightarrow (\varphi_n \rightarrow \psi) \ldots)))$. Using axiom K n times we get the following:

$\vdash K_i(\varphi_1 \to (K_i\varphi_2 \to (\ldots \to (K_i\varphi_n \to K_i\psi)\ldots)))$. Using that $\{\varphi_1, \ldots, \varphi_n\} \subseteq K_i(x)$ we have that $K_i\psi \in x$, *Q.E.D.*

($\Leftarrow$) If $B_i\psi \in x$ then $\psi \in B_i(x)$. Now, for all $y \in S^c$ s.t. $xR_i^c y$ we have $B_i(x) \subseteq y$. Thus for all $y \in S^c$ s.t. xR_iy, $\psi \in y$, *i.e.* $\mathcal{M}_{lob}^c, y \models \psi$ by IH. Therefore, $\mathcal{M}_{lob}^c, x \models B_i\psi$.

$\square$

Consider the formula $\hat{B}_i\top$. This formula is valid in any model $\mathcal{M} = (S, R, V)$ where R_i is serial for agent $i \in AG$. Therefore, LCO is an extension of the logic $KD45_{AG}$ obtained by replacing schema (T_{bool}) by axiom $\hat{B}_i\top$. The inclusion of $KD45_{AG}$ in LCO is a strict one if $PROP \neq \emptyset$. To show this, we associate to any subset X of $PROP$, the logic LCO_X obtained by replacing schema (T_{bool}) by (T_{bool}^X): "$B_i\varphi \to \varphi$ for all φ boolean and based on X". The function $X \longmapsto LCO_X$ is clearly strictly increasing: each added atom gives new theorems. Furthermore, $LCO = LCO_{PROP}$ and $KD45_{AG} = LCO_{\emptyset}$.

We use in the following sections the notion of bisimulation presented in Section 2.2. Indeed, in this logic as in K, the following proposition is true:

Proposition 4.2 *Let $\mathcal{M} = (S, R, V)$, $\mathcal{M}' = (S', R', V')$ two models of the class $\mathcal{C}_0$ and let $s_0 \in S$, $s_0' \in S'$ be two states. If $(\mathcal{M}, s_0) \leftrightarrows (\mathcal{M}', s_0')$ then $(\mathcal{M}, s_0) \models \varphi$ iff $(\mathcal{M}', s_0') \models \varphi$ for any formula $\varphi \in \mathcal{L}_{lob}$.*

The proof of this proposition can be found in [Hommersom et al., 2004].

4.2.1 Decidability and Complexity:

We prove in this paragraph that the logic of objective beliefs is $PSPACE$-complete, a notion presented in Section 2.2. We prove first that this logic has the finite-model property.

Proposition 4.3 *Let $\varphi \in \mathcal{L}_{lob}$ be a formula. If φ is satisfied in a model of $\mathcal{C}_0$ then φ is satisfied in a finite model of $\mathcal{C}_0$.*

We do not prove it here, but the proof is very similar to (in fact identical to a part of) the proof of Proposition 6.32. Indeed, we take a model that satisfies φ, and we take its filtration (see Definition 2.16) through the set of all subformulas of φ. See the proof of Proposition 6.32 for details.

Proposition 4.4 *If $|AG| \geqslant 2$, the problem of satisfiability of $\mathcal{L}_{lob}$ with respect to $\mathcal{C}_0$ is decidable and $PSPACE$-hard.*

Proof This proof is largely inspired from the equivalent proof for $KD45_n$ in [Halpern and Moses, 1992]. Recall that a *quantified Boolean formula* (QBF) can be written in the following form: $A = Q_1p_1Q_2p_2\ldots Q_mp_mA'$ where for all i, $Q_i \in \{\forall, \exists\}$ and A' is a Boolean formula whose primitive propositions are among $p_1, \ldots, p_m$. Recall that the problem of deciding whether a QBF is true or not is $PSPACE$-complete [Stockmeyer and Meyer, 1973].

We get our result by proving that, given a QBF A we can construct a $\mathcal{L}_{lob}$-formula ψ_A such that ψ_A is satisfiable in $\mathcal{C}_0$ iff A is true. Take a *QBF* A of the previous form. We call $A_1 := Q_2p_2 \dots Q_mp_mA'$ and more generally $A_k := Q_{k+1}p_{k+1} \dots Q_mp_mA'$. Therefore $A = Q_1p_1A_1 = Q_1p_1Q_2p_2 \dots Q_kp_kA_k$.

The formula ψ_A we construct enforces the existence of a *binary tree-like* model, in which each leaf represents a distinct truth assignment to the primitive propositions $p_1, \dots, p_m$. Our primitive propositions are the p_i's and additional $d_0, \dots, d_m$, where, intuitively, d_j denotes the fact that we already did at least j consecutive assignments of p_i's values. In other words, d_j is true in the nodes of depth at least j. For technical reasons that we present later, the link between a node of depth i and a node of depth $i+1$ is a succession of two arrows (one for agent 1 and the other for agent 2) with an intermediate node which is exactly identical to its antecedent. An example of such a model is given in figure 4.1. We then define the following $\mathcal{L}_{lob}$-formulas:

- *depth* captures the intended relation between the d_j's:

$$depth := \bigwedge_{i=1}^{m} (d_i \longrightarrow d_{i-1})$$

- *determined* says, intuitively, that the truth value of the proposition p_i is determined at depth i in the tree. If p_i is true (resp. false) in a node s of depth $j \geqslant i$, then it is true (resp. false) in all the nodes that are under s:

$$determined := \bigwedge_{i=1}^{m} (d_i \longrightarrow ((p_i \longrightarrow B_1B_2(d_i \longrightarrow p_i)) \wedge (\neg p_i \longrightarrow B_1B_2(d_i \longrightarrow \neg p_i))))$$

- $branching_A$ says that for any node of depth i, if the truth value of p_{i+i} is quantified universally (resp. existentially) in A, it is possible to find two successor nodes (resp. one successor node) at depth $i+1$ such that p_{i+1} is true at one and false at the other (resp. p_{i+1} has the expected truth value):

$$\bigwedge_{\{i:Q_{i+1}=\forall\}} ((d_i \wedge \neg d_{i+1} \longrightarrow (\hat{B}_1\hat{B}_2(d_{i+1} \wedge \neg d_{i+2} \wedge p_{i+1}) \wedge \hat{B}_1\hat{B}_2(d_{i+1} \wedge \neg d_{i+2} \wedge \neg p_{i+1}))\wedge$$

$$\bigwedge_{\{i:Q_{i+1}=\exists\}} ((d_i \wedge \neg d_{i+1} \longrightarrow (\hat{B}_1\hat{B}_2(d_{i+1} \wedge \neg d_{i+2} \wedge p_{i+1}) \vee \hat{B}_1\hat{B}_2(d_{i+1} \wedge \neg d_{i+2} \wedge \neg p_{i+1})).$$

Finally,

$$\psi_A := d_0 \wedge \neg d_1 \wedge (B_1B_2)^m(depth \wedge determined \wedge branching_A \wedge (d_m \longrightarrow A')).$$

We can now prove that ψ_A is satisfiable in $\mathcal{C}_0$ iff A is true. First, we show that if A is true then ψ_A is satisfied by the model $\mathcal{M}_{\psi_A}$ corresponding to the requirements explained before. We describe now such a model starting from an initial state s_0 satisfying only d_0. In this description, we call "node of depth i" a node satisfying all the d_j with $j \leqslant i$ and no other. Furthermore, if a node s has depth i, we call "successor" of s a node s' of depth $i+1$ such that $s(R_1 \circ R_2)s'$, with an intermediate node that is identical to s.

Now, if $Q_1 = \forall$, s_0 has two successors $s_1^{p_1}$ and $s_1^{\neg p_1}$, p_1 being satisfied in the first and not being satisfied in the second. If $Q_i = \exists$, we construct a unique successor s_1, p_1 being satisfied in it iff $A_1(\top/p_1)$ is true. Then we reproduce this process for the new created states, considering as A the actualized formula $A_1(\top/p_1)$ if p_1 is satisfied and $A_1(\bot/p_1)$ if p_1 is not satisfied, and maintaining in every further state the valuation of p_1.

In other words, for all i, if $Q_i = \forall$, each node of depth i has two successors, p_i being satisfied in the first and not being satisfied in the second. If $Q_i = \exists$, each node of depth i node has a unique successor, p_i being satisfied iff $A_i(*_1/p_1, \ldots *_{i-1}/p_{i-1}, p_i := \top)$ is true (where $*_k$ corresponds to the actual valuation assigned to p_k).

We then end the model taking for each relation R_i its reflexive-symmetric closure. An example of such a model is given in Figure 4.1. Note that such a model is in $\mathcal{C}_0$ as every relation is an equivalence relation.

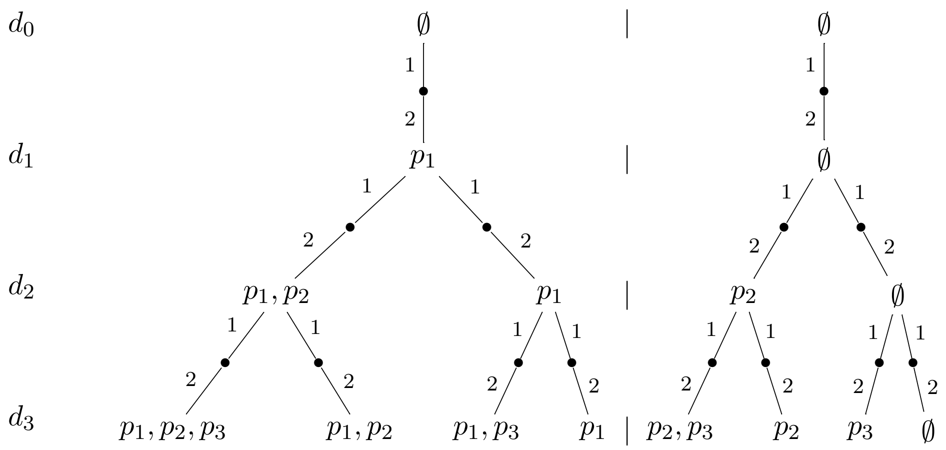


Figure 4.1: Possible models $\mathcal{M}_{\psi_A}$ for $A = \exists p_1 \forall p_2 \forall p_3 A'$
One of these models, in its upper state, satisfies ψ_A if $A = \exists p_1 \forall p_2 \forall p_3 A'$ is true. The left one if $A_1(\top/p_1)$ is true, the right one if $A_1(\bot/p_1)$ is true.

Now suppose that A is satisfiable, we get $\mathcal{M}_{\psi_A}, s_0 \models \psi_A$ by construction. Indeed, $d_0 \wedge \neg d_1$ is satisfied in s_0, and *depth* $\wedge$ *branching*$_A$ is clearly satisfied in all the model. To see that *determined* is a validity of the model, note that by a unique arrow $R_1 \circ R_2$ from a node s, we cannot reach a node of depth higher or equal to the depth of s that is not a successor of s. Finally, we have to see that $d_m \longrightarrow A'$ is a validity of $\mathcal{M}_{\psi_A}$, which means exactly that if the value of each p_i is fixed, arbitrarily if $Q_i = \forall$ and choosing the right one if $Q_i = \exists$, the

A' corresponding to this choice is a true boolean proposition. That exactly means that A is true.

Conversely, suppose that a model $\mathcal{M} = (S, V, \{R_1, R_2\})$ in $\mathcal{C}_0$ such that $\mathcal{M}, s \models \psi_A$ exists. Given a state $t \in S$, let A_j^t be the QBF that results by starting with $Q_{j+1}p_{j+1} \ldots Q_m p_m A'$ and replacing all occurrence of p_i in A', with $i < j$, by $\top$ if $t \in V(p)$ and by $\bot$ otherwise. Note that $A_0^t = A$ and that A_m^t corresponds to A' where each atomic proposition has been replaced by its valuation in t. Now we have $\mathcal{M}, s \models (B_1 B_2)^m (d_m \longrightarrow A')$. Thus if $(s,t) \in (R_1 \circ R_2)^m$ and $\mathcal{M}, t \models d_m$ then A_m^t is true. With the fact that $\mathcal{M} \models B_1 B_2(branching_A)$ and an easy induction on j, we can prove that for all $t \in S$, if $(s,t) \in (R_1 \circ R_2)^{m-j}$ and $\mathcal{M}, t \models d_{m-j} \wedge \neg d_{m-j+1}$ then the QBF A_{m-j}^t is true. In particular, since $\mathcal{M}, s \models d_0 \wedge \neg d_1$, $A_0^s = A$ is true.

□

Proposition 4.5 *The problem of satisfiability of $\mathcal{L}_{lob}$ with respect to $\mathcal{C}_0$ is in $PSPACE$.*

This proof is identical to the similar proof of the $PSPACE$ complexity of the problem of satisfiability of the logic K [Halpern and Moses, 1992]. It uses a notion of *tableau* that generalises the notion of propositional tableau. Such a tableau method is proposed for another logic in Section 6.6. The idea of the proof is the following:

- First, we show that this logic has the tree-model property. That means that if a formula is satisfiable, it is satisfied in a tree-like model. More precisely, we prove that a formula is satisfiable iff the tableau method terminates and allow to construct such a tree-like model. It would take at most a time exponential in the size of the formula φ that is satisfied.

- Second, we show that such a model has a depth that is polynomial in the size φ.

- Third, we construct this tableau depth first, in other words we construct the tree-like tableau branch by branch. Once a branch constructed, we examine if it can satisfy the formula. If it does not, we just forget it a pass to the following branch.

In this way we obtain an algorithm that solves the problem in exponential time and polynomial space.

Proposition 4.6 *The problem of the model checking of $\mathcal{L}_{lob}$ with respect to $\mathcal{C}_0$ is in P.*

Proof To prove it, just recall that the problem of the model checking of the language of modal logic with respect to all Kripke models is in P (see [Gradel and Otto, 1999]). As the language is the same, it remains true if we take the particular class of models $\mathcal{C}_0$.

□

4.2.2 Adding Update

Let us now tackle the question of updating objective beliefs. What happens in a given model if a group of agents learns that a boolean formula is satisfied? To answer this question, we first present the framework proposed by [Hommersom et al., 2004], showing some new properties.

Definition 4.7 *Let $\mathcal{M} = (S, R, V)$ be a model, φ a boolean formula and $G \subseteq AG$ a finite group of agents. The update of $\mathcal{M}$ by φ and G is the model $\mathcal{U}_{\varphi,G}(\mathcal{M}) = (S', R', V')$ defined in the following way:*

- $S' = S \times \{0, 1\}$,
- $(x, a)R'_i(y, b)$ *iff one of the following conditions is satisfied :*
  - $a = 0$, $b = 0$ *and* xR_iy,
  - $a = 1$, $b = 1$, xR_iy, $(\mathcal{M}, y) \models \varphi$ *and* $i \in G$,
  - $a = 1$, $b = 0$, xR_iy *and* $i \notin G$.
- $V'(x, a) = V(x)$.

In the case where $G = \emptyset$ we obtain that $\mathcal{U}_{\varphi,\emptyset}(\mathcal{M})$ satisfies $(x, a)R'_i(y, b)$ iff (xR_iy and $b = 0$).

Clearly this update is an informative event, *i.e.* an event that does not change the valuation of the propositional atoms but only the knowledge that agents may have about the situation. It is thus equivalent to an action model, as presented in Section 3.1.3. Figure 4.2 presents this action model:

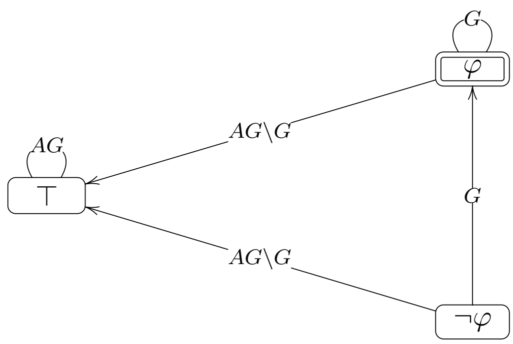


Figure 4.2: Action model equivalent to the update $\mathcal{U}_{\varphi,G}(\mathcal{M})$ for objective beliefs

Here are some properties of the result of the operation $\mathcal{U}$ on a model. These properties justify the choice of this operation $\mathcal{U}$ to model the update of an objective belief by a group of agents.

Proposition 4.8 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ be a boolean formula and $G \subseteq AG$ be a group of agents. For all $s \in S$, if $(\mathcal{M}, s) \models \varphi$ then the submodel of $\mathcal{U}_{\varphi,G}(\mathcal{M})$ generated from $(s, 1)$ is a model of the class $\mathcal{C}_0$.*

In other words, the class $\mathcal{C}_0$ is stable by the application of operation $\mathcal{U}$.

Proposition 4.9 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ be a boolean formula and $G \subseteq AG$ be a group of agents. For all $s \in S$ and all $i \in AG$,*

- *for $i \in G$, $(\mathcal{U}_{\varphi,G}(\mathcal{M}), (s, 1)) \models B_i\varphi$*
- *for $i \notin G$, $(\mathcal{U}_{\varphi,G}(\mathcal{M}), (s, 1)) \models B_i\varphi$ iff $(\mathcal{M}, s) \models B_i\varphi$.*

In other words, the agents of group G believe the formula appearing in the update is true; the other agents believe it only if they already believed it before the update was made. The proof of the two previous propositions is a simple verification and can be found in [Hommersom et al., 2004]. The two following propositions say that updating by the boolean constant $\top$ or updating by any formula for an empty set of agents change nothing to any agent beliefs. This is quite intuitive! More precisely, they assert that such update gives a model bisimilar to the initial one. The notion of bisimulation, denoted $\rightleftarrows$, is introduced in Definition 2.13.

Proposition 4.10 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ be a boolean formula, $s_0 \in S$ be a state of the model and $G \subseteq AG$ be a group of agents. Then $(\mathcal{U}_{\top,G}(\mathcal{M}), (s_0, 1)) \rightleftarrows (\mathcal{M}, s_0)$.*

Proof Let $\mathfrak{R}$ be the binary relation between $S \times \{0, 1\}$ and S defined in the following way: $(s, a)\mathfrak{R}s'$ iff $s = s'$.

We show that $\mathfrak{R}$ is a bisimulation between the sub-model of $\mathcal{U}_{\top,G}(\mathcal{M})$ generated from $(s_0, 1)$ and the sub-model of $\mathcal{M}$ generated from s_0. First, clearly $(s_0, 1)\mathfrak{R}s_0$. Now for all $a \in \{0, 1\}$, all $s \in S$ (*i.e.* for all situation such that $(s, a)\mathfrak{R}s$)

atoms for all $p \in \Theta$: $(s, a) \in V(p)$ iff $s \in V'(p)$ (by Definition 4.7);

forth for all $i \in AG$ and all $(t, b) \in S \times \{0, 1\}$: if $(s, a)R_i(t, b)$, then sR_it by Definition 4.7, with $(t, b)\mathfrak{R}t$;

back for all $i \in AG$ and all $t \in S$: if sR_it, then we distinguish two cases:

- if $a = 0$ then $(s, 0)R_i(t, 0)$, with $(t, 0)\mathfrak{R}t$
- if $a = 1$ then two cases again:
  - if $i \notin G$ then $(s, 1)R_i(t, 0)$
  - if $i \in G$ then $(s, 1)R_i(t, 1)$, because $\mathcal{M}, t \models \top$.

In all cases we obtained $b \in \{0, 1\}$ such that $(s, a)R_i(t, b)$, with again $(t, b)\mathfrak{R}t$.

□

Proposition 4.11 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ be a boolean formula and $s_0 \in S$ be a state of the model. Then $(\mathcal{U}_{\varphi,\emptyset}(\mathcal{M}), (s_0, 1)) \rightleftarrows (\mathcal{M}, s_0)$.*

Proof It is easy to prove it in the same way as for Proposition 4.10, with the same relation $\mathfrak{R}$.

□

Proposition 4.12 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ, ψ be boolean formulas, $s_0 \in S$ be a state of the model and $G \subseteq AG$ be a group of agents. If $(\mathcal{M}, s_0) \models \varphi$ and $(\mathcal{M}, s_0) \models \psi$ then $(\mathcal{U}_{\psi,G}(\mathcal{U}_{\varphi,G}(\mathcal{M})), ((s_0, 1), 1)) \rightleftarrows (\mathcal{U}_{\varphi\wedge\psi,G}(\mathcal{M}), (s_0, 1))$.*

Two successive updates are thus equivalent to a unique one.
Proof It is easy to prove it in the same way as for Proposition 4.10, with the following binary relation $\mathfrak{R}$ between $(S \times \{0,1\}) \times \{0,1\}$ and $S \times \{0,1\}$:

- $((s, a), b)\mathfrak{R}(t, c)$ iff $s = t$, $a = c$ and $b = c$.

□

Proposition 4.13 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, φ and ψ be boolean formulas, $s_0 \in S$ be a state of the model and $G, H \subseteq AG$ be groups of agents. If $(\mathcal{M}, s_0) \models \varphi$ and $(\mathcal{M}, s_0) \models \psi$ then $(\mathcal{U}_{\psi,H}(\mathcal{U}_{\varphi,G}(\mathcal{M})), ((s_0, 1), 1)) \rightleftarrows (\mathcal{U}_{\varphi,G}(\mathcal{U}_{\psi,H}(\mathcal{M})), ((s_0, 1), 1))$.*

Therefore, if we consider two successive updates, the order is not important.
Proof It is easy to prove it in the same way as for Proposition 4.10, with the following relation $\mathfrak{R}$ over $(S \times \{0,1\}) \times \{0,1\}$:

- $((s, a), b)\mathfrak{R}((t, c), d)$ iff $s = t$, $a = d$ and $b = c$.

□

Proposition 4.14 *Let $\mathcal{M} = (S, R, V)$ and $\mathcal{M}' = (S', R', V')$ be models of the class $\mathcal{C}_0$, φ be a boolean formula, $s_0 \in S$ and $s'_0 \in S'$ be states of these models and $G \subseteq AG$ be a group of agents. If $(\mathcal{M}, s_0) \models \varphi$ and $(\mathcal{M}, s_0) \rightleftarrows (\mathcal{M}', s'_0)$ then $(\mathcal{U}_{\varphi,G}(\mathcal{M}), (s_0, 1)) \rightleftarrows (\mathcal{U}_{\varphi,G}(\mathcal{M}'), (s'_0, 1))$.*

In other words, the bisimilarity between two models remains after the update of these models by a same formula for a same group of agents.
Proof If we call $\mathfrak{R}$ the bisimulation between the submodel of $\mathcal{M}$ generated from s_0 and the submodel of $\mathcal{M}'$ generated from s'_0, then it is easy to verify that the binary relation $\mathfrak{R}^{\mathcal{U}}$ between $S \times \{0,1\}$ and $S' \times \{0,1\}$ defined as follow is a bisimulation:

- $(s, a)\mathfrak{R}^{\mathcal{U}}(s', a')$ iff $u\mathfrak{R}u'$ and $a = a'$.

□

The last but not the least:

Proposition 4.15 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, ψ be a boolean formulas, $s_0 \in S$ be a state of the model and $G \subseteq AG$ be a group of agents. If $\mathcal{M}, s_0 \models \psi$ then $(\mathcal{U}_{\psi,G}(\mathcal{M}), (s_0, 0)) \leftrightarrows (\mathcal{M}, s_0)$.*

Recall that after the update, the agents that are not in G believe that $(s_0, 0)$ is the actual state. Therefore this proposition claims that the beliefs of agents that are not in G do not change after the update.

Proof It is easy to prove it in the same way as for Proposition 4.10, with the following relation $\mathfrak{R}$ between $S \times \{0, 1\}$ and S

- $(s, a)Z(t)$ iff $s = t$ and $a = 0$.

□

4.3 Update of Objective Beliefs

4.3.1 Syntax and Semantics

LOB only uses the modal operators B_i. Therefore, it cannot express update of beliefs. If we look again to our *Texas Hold'em* example, with only the notion of belief in our language, how can we express the fact that after a card is dealt, on the table all the agents update their beliefs? In this section, we propose new operators that, added to *LOB*, give us the possibility to analyse dynamics of the update of objective beliefs. We adapt the framework proposed by [Hommersom et al., 2004] to propose the Logic of Update of Objective Beliefs (*LUOB*) which language contains, besides the operators B_i, operators of the form $[\psi, G]$ for any boolean formula ψ and any group of agents $G \subseteq AG$. We then present new results of expressivity and decidability/complexity.

As we just said, the language of *LUOB* ($\mathcal{L}_{luob}$) over a countable set of propositional atoms *PROP* and a countable set of agents *AG* is defined inductively as follow:

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid B_i\varphi \mid [\psi, G]\varphi$$

where $p \in PROP$, $i \in AG$, $G \subseteq AG$ and $\psi \in \mathcal{L}_{pl}$.

We understand $[\psi, G]\varphi$ as "after the agents of the group G learn ψ, φ is true", so these operators introduce an idea of update. Therefore, we generalize the satisfiability relation presented in Section 3.1.4 in the following way:

- $\mathcal{M}, s \models [\psi, G]\varphi$ iff ($\mathcal{M}, s \models \psi$ implies $\mathcal{U}_{\psi,G}(\mathcal{M}), (s, 1) \models \varphi$).

Once again, the notion of bisimilarity is useful in this framework, as two bisimilar models satisfy the same formula. More precisely:

Proposition 4.16 *Let $\mathcal{M} = (S, R, V)$, $\mathcal{M}' = (S', R', V')$ be two models and $s_0 \in S$, $s'_0 \in S'$ be two states of these models. If $(\mathcal{M}, s_0) \rightleftarrows (\mathcal{M}', s'_0)$ then $(\mathcal{M}, s_0) \models \varphi$ iff $(\mathcal{M}', s'_0) \models \varphi$ for every formula $\varphi \in \mathcal{L}_{luob}$.*

Proof It can be proved by an easy induction on the structure on φ.

□

Therefore, we obtain the following:

Proposition 4.17 *Let $\mathcal{M}$ be a model, s be a state of $\mathcal{M}$, $\psi_1, \psi_2 \in \mathcal{L}_{pl}$ and $G_1, G_2 \subseteq AG$.*

1. $\mathcal{M}, s \models [\psi_1, G_1]\psi_2$ *iff* $\mathcal{M}, s \models \psi_1 \longrightarrow \psi_2$
2. $\mathcal{M}, s \models [\psi_1, G_1][\psi_2, G_2]\varphi \longleftrightarrow [\psi_2, G_2][\psi_1, G_1]\varphi$.

This proposition asserts that two successive updates are interchangeable.
Proof

1. $\mathcal{M}, s \models [\psi_1, G_1]\psi_2$ iff ($\mathcal{M}, s \models \psi_1$ implies $\mathcal{U}_{\psi_1,G_1}(\mathcal{M}), (s, 1) \models \psi_2$). But by definition of the update, $(s, 1)$ has the same valuation of s. Therefore, as ψ_2 is a propositional formula, $\mathcal{M}, s \models [\psi_1, G_1]\psi_2$ iff ($\mathcal{M}, s \models \psi_1$ implies $\mathcal{M}, s \models \psi_2$) *Q.E.D.*

2. $\mathcal{M}, s \models [\psi_1, G_1][\psi_2, G_2]\varphi$
   iff $\mathcal{M}, s \models \psi_1$ implies $\mathcal{U}_{\psi_1,G_1}(\mathcal{M}), (s, 1) \models [\psi_2, G_2]\varphi$
   iff $\mathcal{M}, s \models \psi_1$ implies ($\mathcal{M}, s \models \psi_2$ implies $\mathcal{U}_{\psi_2,G_2}(\mathcal{U}_{\psi_1,G_1}(\mathcal{M})), ((s, 1), 1) \models \varphi$) (by 1.)
   iff $\mathcal{M}, s \models \psi_1 \wedge \psi_2$ implies $\mathcal{U}_{\psi_2,G_2}(\mathcal{U}_{\psi_1,G_1}(\mathcal{M})), ((s, 1), 1) \models \varphi$
   iff $\mathcal{M}, s \models \psi_1 \wedge \psi_2$ implies $\mathcal{U}_{\psi_1,G_1}(\mathcal{U}_{\psi_2,G_2}(\mathcal{M})), ((s, 1), 1) \models \varphi$ (Prop. 4.13 and 4.16)
   iff $\mathcal{M}, s \models [\psi_2, G_2][\psi_1, G_1]\varphi$ (conversely)

□

Definition 4.18 *Let $\psi_1, \ldots, \psi_n \in \mathcal{L}_{pl}, G_1, \ldots, G_n \subseteq AG$ be given. For every model $\mathcal{M} = (S, \mathcal{R}, V)$ and every state of the model $s \in S$, we write $\mathcal{UP}_n(\mathcal{M}) = \mathcal{U}_{\psi_n,G_n}(\ldots(\mathcal{U}_{\psi_1,G_1}(\mathcal{M}))\ldots)$ and $\mathcal{UP}_n(\mathcal{M}, s) = \mathcal{UP}_n(\mathcal{M}), ((s, 1), \ldots, 1)$.*

Note that the formulas $\psi_1, \ldots, \psi_n$ and the groups $G_1, \ldots, G_n$ are supposed to be clear in this definition. This definition comes from the interchangeability of successive updates that allows to omit the order in a multiple update. More precisely, by propositions 4.13 and 4.16, the order in such a succession of updates is not important when considering the satisfaction of a formula. Therefore using the semantics and Proposition 4.17.2 we obtain the following:

Proposition 4.19 *Let $\mathcal{M} = (S, \mathcal{R}, V)$ be a model, let $s \in S$, let $\psi_1, \ldots, \psi_n \in \mathcal{L}_{pl}$ and let $G_1, \ldots, G_n \subseteq AG$. Then for every $\mathcal{L}_{luob}$-formula φ we have:*

$$\mathcal{M}, s \models [\psi_1, G_1]\ldots[\psi_n, G_n]\varphi \textit{ iff } (\mathcal{M}, s \models \psi_1 \wedge \ldots \wedge \psi_n \textit{ implies } \mathcal{UP}_n(\mathcal{M}, s) \models \varphi)$$

Another interesting result of Corollary of Proposition 4.16 is the following:

Proposition 4.20 *Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, ψ be a boolean formulas, $s_0 \in S$ be a state of the model and $G \subseteq AG$ be a group of agents. If $\mathcal{M}, s_0 \models \psi$ then for all $\varphi \in \mathcal{L}_{luob}$ we have $\mathcal{U}_{\psi,G}(\mathcal{M}), (s_0, 0) \models \varphi$ iff $\mathcal{M}, s_0 \models \varphi$.*

Proof It is a translation of Proposition 4.15 through proposition 4.16.

$\square$

This proposition asserts that, if the update concerns a formula satisfied in s_0, then a formula of $\mathcal{L}_{luob}$ is satisfied in $(s_0, 0)$ after the update iff it was satisfied before. We get the following result concerning the belief of the agents after un update.

Proposition 4.21 *Let $\mathcal{M}$ be a model, s be a state of $\mathcal{M}$, $G \subseteq AG$ a group of agents, $i \in AG \backslash G$, and $\psi \in \mathcal{L}_{pl}$. Therefore, if $\mathcal{M}, s \models \psi$ then $\mathcal{U}_{\psi,G}(\mathcal{M}), (s, a) \models B_i\varphi$ iff $\mathcal{M}, s \models B_i\varphi$.*

Proof We pose $\mathcal{U}_{\psi,G}(\mathcal{M}) = (S', \mathcal{R}', V')$.

$\mathcal{U}_{\psi,G}(\mathcal{M}), (s, a) \models B_i\varphi$

iff for all $(t, b) \in S'$ s.t. $(s, a)R'_i(t, b)$ we have $\mathcal{U}_{\psi,G}(\mathcal{M}), (t, b) \models \varphi$

iff for all $t \in S$ s.t. sR_it we have $\mathcal{U}_{\psi,G}(\mathcal{M}), (t, 0) \models \varphi$
(because if $i \notin G$ then $(s, a)R'_i(t, b)$ iff sR_it and $b = 0$).

iff for all $t \in S$ s.t. sR_it we have $\mathcal{M}, t \models \varphi$ (by Proposition 4.20)

iff $\mathcal{M}, s \models B_i\varphi$

$\square$

4.3.2 Axiomatization and Completeness

Here are the axioms of $LUOB$.

(LOB) the axioms of the logic of objective beliefs,

$(R1)$ $[\psi, G]p \longleftrightarrow (\psi \to p)$,

$(R2)$ $[\psi, G]\bot \longleftrightarrow (\psi \to \bot)$,

$(R3)$ $[\psi, G]\neg\varphi \longleftrightarrow (\psi \to \neg[\varphi, G]\varphi)$,

$(R4)$ $[\psi, G](\varphi \vee \chi) \longleftrightarrow [\psi, G]\varphi \vee [\psi, G]\chi$,

$(R5)$ $[\psi, G]B_i\varphi \longleftrightarrow (\psi \to B_i[\psi, G]\varphi)$ when $i \in G$,

$(R6)$ $[\psi, G]B_i\varphi \longleftrightarrow (\psi \to B_i\varphi)$ when $i \notin G$.

Axioms $(R1)$, $(R2)$, $(R3)$, $(R4)$, $(R5)$ and $(R6)$ have an easy interpretation, similar to the one for PAL. Just as $\mathcal{L}_{pal}$ is not more expressive than $\mathcal{L}_{el}$, the fact that these axioms are *reduction* axioms implies that the language $\mathcal{L}_{luob}$ is not more expressive than $\mathcal{L}_{lob}$. $LUOB$ theorems are all the formulas deducible from these axioms, using the inference rules (MP) and (GD) and the following inference rule:

$(G\mathcal{U})$ if ψ is a theorem then $[\varphi, G]\psi$ is a theorem.

As Proposition 4.27 states, these axioms and inference rules give a sound and complete axiomatization for the language $\mathcal{L}_{luob}$ with respect to the class $\mathcal{C}_0$. But we first introduce a translation from $\mathcal{L}_{luob}$ to $\mathcal{L}_{lob}$:

Definition 4.22 *For all $\varphi \in \mathcal{L}_{luob}$, we define $\tau([\psi_1, G_1] \dots [\psi_n, G_n], \varphi) \in \mathcal{L}_{lob}$ for all $n \in \mathbb{N}$, all $\psi_1, \dots, \psi_n \in \mathcal{L}_{pl}$ and all $G_1, \dots, G_n \subseteq AG$ inductively on the structure of the formula φ in the following way:*

- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], p) = \psi_1 \wedge \dots \wedge \psi_n \to p$,
- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], \bot) = \psi_1 \wedge \dots \wedge \psi_n \to \bot$,
- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], \neg\varphi) = \psi_1 \wedge \dots \wedge \psi_n \to \neg\tau([\psi_1, G_1] \dots [\psi_n, G_n]\varphi)$,
- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], \varphi \vee \varphi') = \tau([\psi_1, G_1] \dots [\psi_n, G_n]\varphi) \vee \tau([\psi_1, G_1] \dots [\psi_n, G_n]\varphi')$,
- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], B_i\varphi) = \psi_1 \wedge \dots \wedge \psi_n \to B_i\tau(\mu\ \varphi)$ *where μ is the sequence obtained from $[\psi_1, G_1] \dots [\psi_n, G_n]$ eliminating all the $[\psi, G]$ such that $i \notin G$,*
- $\tau([\psi_1, G_1] \dots [\psi_n, G_n], [\psi, G]\varphi) = \tau([\psi_1, G_1] \dots [\psi_n, G_n][\psi, G], \varphi)$.

We show that in Proposition 4.26 that for all $\varphi \in \mathcal{L}_{luob}$, $\models \varphi \longleftrightarrow \tau(\emptyset, \varphi)$. To do so, we need the two following lemmas:

Lemma 4.23 *Let $n \in \mathbb{N}$, $\psi_1, \dots, \psi_n \in \mathcal{L}_{pl}$ and $G_1, \dots, G_n \subseteq AG$. Let μ be defined as in Definition 4.22 and for every model $\mathcal{M}$ and every state s, let $\mathcal{UP}_n(\mathcal{M}, s)$ be defined as in Definition 4.19. Then for all $\theta \in \mathcal{L}_{luob}$, all $i \in AG$,*

$$\text{if } \mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n \text{ then } (\mathcal{M}, s \models B_i(\mu\ \theta) \text{ iff } \mathcal{UP}_n(\mathcal{M}, s) \models B_i\theta)$$

Proof Let m be the number of groups G_k containing i. We have $0 \leqslant m \leqslant n$. By Proposition 4.13 we can suppose, without loss of generality, that $i \in G_k$ iff $k \leqslant m$. Now, we call $\mathcal{UP}_m(\mathcal{M}, s) = \mathcal{U}_{\psi_1,G_1}(\dots(\mathcal{U}_{\psi_m,G_m}(\mathcal{M}))\dots,((s,1),\dots,1)$. We pose $\mathcal{UP}_n(\mathcal{M}) = (S^n, \mathcal{R}^n, V^n)$ and $\mathcal{UP}_m(\mathcal{M}) = (S^m, \mathcal{R}^m, V^m)$.

Now $\mathcal{UP}_n(\mathcal{M}, s) \models B_i\theta$
iff $\mathcal{UP}_m(\mathcal{M}, s) \models B_i\theta$ (by Proposition 4.21 applied $m - n$ times)
iff for all $t \in S^m$ such that $sR_i^m t$, $\mathcal{UP}_m(\mathcal{M}), t \models \theta$
iff for all $t \in S$ s.t. $sR_i t$ and $\mathcal{M}, t \models \psi_1 \wedge \dots \wedge \psi_m$, $\mathcal{UP}_m(\mathcal{M}, t) \models \theta$
iff for all t such that $sR_i t$, $\mathcal{M}, t \models \mu\ \theta$ (by Proposition 4.19)
iff $\mathcal{M}, s \models B_i(\mu\ \theta)$

□

Lemma 4.24 *Let $\varphi \in \mathcal{L}_{luob}$. Then for all $n \in \mathbb{N}$, all $G_1, \dots, G_n \subseteq AG$, all $\psi_1, \dots \psi_n \in \mathcal{L}_{pl}$:*

1. $\mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n], \varphi) \mid \leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \varphi \mid) \times \mid \varphi \mid$,

2. $\tau([\psi_1, G_1] \ldots [\psi_n, G_n], \varphi) \leftrightarrow [\psi_1, G_1] \ldots [\psi_n, G_n]\varphi$ *is valid in all models of the class* $\mathcal{C}_0$.

Remark 4.25 *First we remark that* $\psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_n \rightarrow \psi$ *is an abuse of notation. In fact, we should write* $(\neg(\underbrace{\neg(\ldots\neg(}_{n-2}\neg(\neg\psi_1 \vee \neg\psi_2)\underbrace{\vee\neg\psi_3)\ldots\vee\neg\psi_n)}_{n-2} \vee \psi)$. *Therefore we have :*

$$\begin{aligned} \mid \psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_n \rightarrow \psi \mid &= \mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6(n-1) + \mid \psi \mid +4 \\ &\leqslant \mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \psi \mid. \end{aligned}$$

Proof

1. We prove it by induction on the structure of φ, noting $\chi = ([\psi_1, G_1] \ldots [\psi_n, G_n], \varphi)$. We note by a '$\Box$' the use of Remark 4.25 in this proof.

   base case $\varphi = p$ or $\bot$: It is a direct application of Remark 4.25.

   inductive cases: let us suppose that it is true for θ and θ'.

   - $\varphi = \neg\theta : \mid \tau(\chi) \mid \leqslant^{\Box} (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n) + \mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n], \theta) \mid + 1$
     $\leqslant_{(IH)} (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n) + (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times \mid \theta \mid + 1$
     $= (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times (\mid \theta \mid + 1) - \mid \theta \mid + 1$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times (\mid \theta \mid + 1)$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \varphi \mid) \times (\mid \varphi \mid)$
   - $\varphi = \theta \vee \theta' : \mid \tau(\chi) \mid = \mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n], \theta) \mid + \mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n], \theta') \mid + 3$
     $\leqslant_{(IH)} (\mid \psi_1 \mid + \ldots + 6n + \mid \theta \mid) \times \mid \theta \mid + (\mid \psi_1 \mid + \ldots + 6n + \mid \theta' \mid) \times \mid \theta' \mid + 3$
     $\leqslant (\mid \psi_1 \mid + \ldots + 6n + \mid \theta \mid + \mid \theta' \mid + 3) \times (\mid \theta \mid + \mid \theta' \mid + 3)$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \varphi \mid) \times (\mid \varphi \mid)$
   - $\varphi = B_i\theta$: $\mid \tau(\chi) \mid \leqslant^{\Box} (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n) + \mid \tau(\mu, \theta) \mid + 1$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n) + \mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n], \theta) \mid + 1$
     $\leqslant_{(IH)} (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n) + (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times \mid \theta \mid + 1$
     $= (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times (\mid \theta \mid + 1) - \mid \theta \mid + 1$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times (\mid \theta \mid + 1)$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \varphi \mid) \times (\mid \varphi \mid)$
   - $\varphi = [\psi, G]\theta$: $\mid \tau(\chi) \mid = \mid \tau([\psi_1, G_1] \ldots [\psi_n, G_n][\psi, G], \theta) \mid$
     $\leqslant_{(IH)} (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + \mid \psi \mid + 6(n+1) + \mid \theta \mid) \times \mid \theta \mid$
     $= (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times \mid \theta \mid + (\mid \psi \mid + 6) \times \mid \theta \mid$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \theta \mid) \times (\mid \theta \mid + \mid \psi \mid + 6)$
     $\leqslant (\mid \psi_1 \mid + \ldots + \mid \psi_n \mid + 6n + \mid \varphi \mid) \times (\mid \varphi \mid)$

2. We prove it by induction on the structure of φ, noting $\chi := [\psi_1, G_1] \ldots [\psi_n, G_n]\varphi$. As in Proposition 4.19, we write $\mathcal{UP}(\mathcal{M}, s)$ for $\mathcal{U}_{\psi_n, G_n}(\ldots(\mathcal{U}_{\psi_1, G_1}(\mathcal{M}))\ldots, ((s, 1), \ldots, 1)$. We note by a '$(*)$' the use of Proposition 4.19 in this proof.

- $\varphi = p$:

  $\mathcal{M}, s \models \tau([\psi_1, G_1] \dots [\psi_n, G_n], p)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n \rightarrow p$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{M}, s \models p$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{UP}(\mathcal{M}, s) \models p$ (the valuation is unchanged)

  iff $\mathcal{M}, s \models [\psi_1, G_1] \dots [\psi_n, G_n]p$ by $(*)$

- $\varphi = \bot$: identical

- $\varphi = \neg\theta$:

  $\mathcal{M}, s \models \tau([\psi_1, G_1] \dots [\psi_n, G_n], \neg\theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{M}, s \models \neg\tau([\psi_1, G_1] \dots [\psi_n, G_n], \theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{M}, s \not\models [\psi_1, G_1] \dots [\psi_n, G_n]\theta$ (by IH)

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{UP}(\mathcal{M}, s) \models \neg\theta$ by $(*)$

  iff $\mathcal{M}, s \models [\psi_1, G_1] \dots [\psi_n, G_n]\neg\theta$ by $(*)$ again

- $\varphi = \theta \vee \theta'$: similar

- $\varphi = B_i\theta$:

  $\mathcal{M}, s \models \tau([\psi_1, G_1] \dots [\psi_n, G_n], B_i\theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{M}, s \models B_i\tau(\mu\ , \theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies for all $t \in S$ s.t. sR_it, $\mathcal{M}, t \models \tau(\mu\ , \theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies for all $t \in S$ s.t. sR_it, $\mathcal{M}, t \models \mu\ \theta$ (by IH)

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{M}, s \models B_i(\mu\ \theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n$ implies $\mathcal{UP}(\mathcal{M}, s) \models B_i\theta$ (by Lemma 4.23)

  iff $\mathcal{M}, s \models [\psi_1, G_1] \dots [\psi_n, G_n]B_i\theta$

- $\varphi = [\psi, G]\theta$:

  $\mathcal{M}, s \models \tau([\psi_1, G_1] \dots [\psi_n, G_n], [\psi, G]\theta)$

  iff $\mathcal{M}, s \models \tau([\psi_1, G_1] \dots [\psi_n, G_n][\psi, G], \theta)$

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n \wedge \psi$ implies $\mathcal{M}, s \models [\psi_1, G_1] \dots [\psi_n, G_n][\psi, G]\theta$ (by IH)

  iff $\mathcal{M}, s \models \psi_1 \wedge \dots \wedge \psi_n \wedge \psi$ implies $\mathcal{UP}(\mathcal{M}, s) \models \theta$ by $(*)$

  iff $\mathcal{M}, s \models [\psi_1, G_1] \dots [\psi_n, G_n]\theta$ by $(*)$ again

□

Hence the following:

Proposition 4.26

- $\mid \tau(\emptyset, \varphi) \mid \leq \mid \varphi \mid^2$,
- *$\tau(\emptyset, \varphi) \leftrightarrow \varphi$ is valid in all models of the class $\mathcal{C}_0$.*

In particular, we showed that $\mathcal{L}_{luob}$ is not more expressive that $\mathcal{L}_{lob}$. The translation above sets that every formula of $\mathcal{L}_{luob}$ is equivalent to a formula of $\mathcal{L}_{lob}$. We can now prove the completeness of the given axiomatization.

Proposition 4.27 *Let $\varphi \in \mathcal{L}_{luob}$. φ is a LUOB theorem iff φ is valid in every model of the class $\mathcal{C}_0$.*

Proof For the soundness, by Proposition 4.1, it is sufficient to show soundness of the additive axioms and inference rules. The soundness of the rule $G\mathcal{U}$ is evident and the soundness of the axioms is given by Proposition 4.26.

Let us now prove completeness. We define the canonical model $\mathcal{M}^c_{luob} = (S^c, R^c_i, V^c)$ in the classical way (cf. Definition 2.19). Its membership to the class $\mathcal{C}_0$ can be proved as in the proof of Proposition 4.1.

Now it is sufficient to prove that for all formula $\varphi \in \mathcal{L}_{luob}$, $\mathcal{M}^c_{luob}, x \models \varphi$ iff $\varphi \in x$. We prove this truth lemma by induction on the structure of φ. This proof is identical to the proof of Proposition 4.1 for the base case and the first inductive cases. In particular, we have already shown that for all $\varphi \in \mathcal{L}_{lob}, \mathcal{M}^c_{luob}, x \models \varphi$ iff $\varphi \in x$. Let us analyse the specific induction case where $\varphi = [\psi, G]\chi$:

- $\mathcal{M}^c, x \models [\psi, G]\chi$
  iff $\mathcal{M}^c, x \models \tau([\psi, G]\chi$ by Proposition 4.26
  iff $\tau([\psi, G]\chi) \in x$ because $\tau([\psi, G]\chi) \in \mathcal{L}_{lob}$
  iff $[\psi, G]\chi \in x$ by axioms $R1, \ldots, R6$.

□

4.3.3 Decidability and Complexity

In this section, we give some technical results on *LUOB*. First, *LUOB* has the finite model property. In other words:

Proposition 4.28 *Let $\varphi \in \mathcal{L}_{luob}$. If φ is satisfied in a model of the class $\mathcal{C}_0$ then φ is satisfied in a finite model of the class $\mathcal{C}_0$.*

Proof By propositions 4.3 and 4.26. More precisely, if φ is satisfied in $\mathcal{C}_0$ then $\tau(\varphi)$ is satisfied in $\mathcal{C}_0$ (Proposition 4.26). Thus $\tau(\varphi)$ is satisfied in a finite model of $\mathcal{C}_0$ (Proposition 4.3) which implies that φ is satisfied in a finite model of $\mathcal{C}_0$ (Proposition 4.26 again).

□

This proposition implies that the problem of satisfiability of $\mathcal{L}_{luob}$ is decidable. We also have the decidability of the problem of the model checking. More precisely:

Proposition 4.29 *The problem of satisfiability of $\mathcal{L}_{luob}$ is PSPACE-complete.*

Proposition 4.30 *The problem of model checking for $\mathcal{L}_{luob}$ with respect to $\mathcal{C}_0$ is in P.*

Proof The first is a corollary of propositions 4.4 and 4.5 and the second a corollary of Proposition 4.6 through the translation τ, using the first property of Proposition 4.26.

□

4.4 Arbitrary Update of Objective Beliefs

4.4.1 Syntax and Semantics

With $LUOB$ we could express and understand situations in which agents update their objective beliefs about the world, as in poker. However, if we limit our language to the operators we have (B_i and $[\varphi, G]$) it is impossible to express the notion of arbitrary update. For example, we would like to say "Alex believes that whatever the agents learn in the future of the game, he will still believe his hand is winning". This notion of arbitrary update should be added to our language, with a new modal operator. We propose in this section such operators inspired from [Balbiani et al., 2007]. In addition to $LUOB$ gives us the Logic of Arbitrary Update of Objective Beliefs ($LAUOB$). These additional operators are of the form $[?, G]$ for any group of agents $G \subseteq AG$, with the following reading for $[?, G]\psi$: "whatever group G agents learn, ψ is true". These operators then introduce the wanted notion of arbitrary update. Let us define more precisely the language of $LAUOB$ $\mathcal{L}_{lauob}$ over $PROP$ and AG as follow:

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid B_i\varphi \mid [\psi, G]\varphi \mid [?, G]\varphi$$

where $p \in PROP$, $i \in AG$, $G \subseteq AG$ and $\psi \in \mathcal{L}_{pl}$.

We then generalize the satisfiability relation, for the same models of the class $\mathcal{C}_0$, in the following way:

- $(\mathcal{M}, s) \models [?, G]\varphi$ iff for every boolean formula ψ, $(\mathcal{M}, s) \models [\psi, G]\varphi$.

Here again, bisimulation is a useful notion, as two bisimilar models satisfy the same formulas:

Proposition 4.31 *Let $\mathcal{M} = (S, R, V)$, $\mathcal{M}' = (S', R', V')$ be two models and $s_0 \in S$, $s'_0 \in S'$ be two states of these models. If $(\mathcal{M}, s_0) \rightleftarrows (\mathcal{M}', s'_0)$ then $(\mathcal{M}, s_0) \models \varphi$ iff $(\mathcal{M}', s'_0) \models \varphi$ for any formula $\varphi \in \mathcal{L}_{lauob}$.*

This proposition can easily be showed by induction on the structure of φ.

Proposition 4.32 *The following formulas are valid in all models of the class $\mathcal{C}_0$:*

(T) $[?, G]\varphi \rightarrow \varphi$,

(4) $[?, G]\varphi \rightarrow [?, G][?, G]\varphi$,

(CR) $\langle ?, G\rangle[?, H]\varphi \rightarrow [?, H]\langle ?, G\rangle\varphi$.

First, remark that the formulas of the form $[?, G \cup H]\varphi \rightarrow [?, G][?, H]\varphi$ are not all valid in any model of the class $\mathcal{C}_0$. For example, the formula $[?, \{i, j\}](B_ip \rightarrow B_iB_jp) \rightarrow [?, \{i\}][?, \{j\}](B_ip \rightarrow B_iB_jp)$ is not valid in the model presented in Figure 4.3 (that belongs to the class $\mathcal{C}_0$).

$$\begin{tikzcd} {}_1\bullet^{p} \arrow[loop left, "{i,j}"] \arrow[rr, dash, "{i,j}"] & & {}_0\bullet^{\neg p} \arrow[loop right, "{i,j}"] \end{tikzcd}$$

Figure 4.3: Two counter-examples in one model

In fact, if i and j learn p together (*i.e.* if the update is public), as described by the first part of the formula, they will believe (correctly) that the other believes p. However, if they learn p privately, one after the other, as described by the second part of the formula, they will not update anything about the other agent beliefs.

Also remark that the formulas of the form $[?, G]\langle ?, H\rangle\varphi \rightarrow \langle ?, H\rangle[?, G]\varphi$ are not valid in any model of the class $\mathcal{C}_0$ either. For example, the formula $[?, \{i\}]\langle ?, \{j\}\rangle(B_i p \oplus B_j p) \rightarrow \langle ?, \{j\}\rangle[?, \{i\}](B_i p \oplus B_j p)$ — where the operator $\oplus$ denotes the exclusive disjunction— is not valid in this very model. Indeed, whatever i learns about the value of p, j can learn something so that one and only one of the two agents believes that p. But it is not true that one of the agents can learn something so that whatever the other learns one and only one of the two agents will believe that p.

Proof (of Proposition 4.32) Let $\mathcal{M} = (S, R, V)$ be a model of the class $\mathcal{C}_0$, $s \in S$ be a state of the model and $\varphi \in \mathcal{L}_{lauob}$ be a formula.
(T): Suppose that $(\mathcal{M}, s) \models [?, G]\varphi$ and $(\mathcal{M}, s) \not\models \varphi$. Then, $(\mathcal{M}, s) \models [\top, G]\varphi$. Therefore, $(\mathcal{U}_{\top,G}(\mathcal{M}), (s, 1)) \models \varphi$. By Proposition 4.10, the submodel of $\mathcal{U}_{\top,G}(\mathcal{M})$ generated from $(s, 1)$ and the submodel of $\mathcal{M}$ generated from s are bisimilar. This is in contradiction with Proposition 4.31.
(4): Suppose that $(\mathcal{M}, s) \models [?, G]\varphi$ and $(\mathcal{M}, s) \not\models [?, G][?, G]\varphi$. Then, there exists a boolean formula ψ_1 such that $(\mathcal{M}, s) \not\models [\psi_1, G][?, G]\varphi$. Therefore, $(\mathcal{M}, s) \models \psi_1$ and $(\mathcal{U}_{\psi_1,G}(\mathcal{M}), (s, 1)) \not\models [?, G]\varphi$. Then, there exists a boolean formula ψ_2 such that $(\mathcal{U}_{\psi_1,G}(\mathcal{M}), (s, 1)) \not\models [\psi_2, G]\varphi$. Therefore, $(\mathcal{U}_{\psi_1,G}(\mathcal{M}), (s, 1)) \models \psi_2$ and $(\mathcal{U}_{\psi_2,G}(\mathcal{U}_{\psi_1,G}(\mathcal{M})), ((s, 1), 1)) \not\models \varphi$. By Proposition 4.12, the submodel of $\mathcal{U}_{\psi_1,G}(\mathcal{U}_{\psi_2,G}(\mathcal{M}))$ generated from $((s, 1), 1)$ and the submodel of $\mathcal{U}_{\psi_1\wedge\psi_2,G}(\mathcal{M})$ generated from $(s, 1)$ are bisimilar. This is in contradiction with Proposition 4.31.
(CR): Suppose that $(\mathcal{M}, s) \models \langle ?, G\rangle[?, H]\varphi$ and $(\mathcal{M}, s) \not\models [?, H]\langle ?, G\rangle\varphi$. Then there exists a boolean formula ψ_1 such that $(\mathcal{M}, s) \models \langle \psi_1, G\rangle[?, H]\varphi$ and a boolean formula ψ_2 such that $(\mathcal{M}, s) \not\models [\psi_2, H]\langle ?, G\rangle\varphi$. Therefore, $(\mathcal{M}, s) \models \psi_1$, $(\mathcal{U}_{\psi_1,G}(\mathcal{M}), (s, 1)) \models [H]\varphi$, $(\mathcal{M}, s) \models \psi_2$, $(\mathcal{U}_{\psi_2,G}(\mathcal{M}), (s, 1)) \not\models \langle ?, G\rangle\varphi$. Thus, $(\mathcal{U}_{\psi_1,G}(\mathcal{M}), (s, 1)) \models \psi_2$, $(\mathcal{U}_{\psi_2,H}(\mathcal{U}_{\psi_1,G}(\mathcal{M})), ((s, 1), 1)) \models \varphi$, $(\mathcal{U}_{\psi_2,H}(\mathcal{M}), (s, 1)) \models \psi_1$ and $(\mathcal{U}_{\psi_1,G}(\mathcal{U}_{\psi_2,H}(\mathcal{M})), ((s, 1), 1)) \not\models \varphi$. By Proposition 4.13, the submodel of $\mathcal{U}_{\psi_2,H}(\mathcal{U}_{\psi_1,G}(\mathcal{M}))$ generated from $((s, 1), 1)$ and the submodel of $\mathcal{U}_{\psi_1,G}(\mathcal{U}_{\psi_2,H}(\mathcal{M}))$ generated from $((s, 1), 1)$ are bisimilar. This is in contradiction with Proposition 4.31.

$\square$

4.4.2 Axiomatization and Completeness

Here are $LAUOB$ axioms:

$(LUOB)$ $LUOB$ axioms

$(S0)$ $[?, G](\psi \rightarrow \chi) \rightarrow ([?, G]\psi \rightarrow [?, G]\chi)$,

$(S1)$ $[?, G]\varphi \rightarrow [\psi, G]\varphi$, where $\psi \in \mathcal{L}_{pl}$.

The axiom $S1$ has a simple interpretation. If after whatever group G of agents learn, φ becomes true, then for any boolean formula ψ, after group G of agents learn ψ, φ becomes true . However, it is not sufficient to ensure completeness of $LAUOB$ for the class $\mathcal{C}_0$. In order to obtain this result, we have to add the following inference rules:

(GAU) if φ is a theorem then $[?, G]\varphi$ is a theorem,

(X) if $\boldsymbol{\theta}([\psi, G]\varphi)$ is a theorem for all $\psi \in \mathcal{L}_{pl}$ then $\boldsymbol{\theta}([?, G]\varphi)$ is a theorem.

In the inference rule X, $\boldsymbol{\theta}$ represents a *necessity form.* Necessity forms were introduced by [Goldblatt, 1982] and are similar to the notion of admissible form. More precisely:

Definition 4.33 (Necessity form for $\mathcal{L}_{lauob}$) *A necessity form is an element of the set defined inductively as follows:*

- *$\sharp$ is a necessity form,*
- *if $\boldsymbol{\theta}$ is a necessity form then for every formula $\varphi \in \mathcal{L}_{lauob}$, $(\varphi \longrightarrow \boldsymbol{\theta})$ is a necessity form,*
- *if $\boldsymbol{\theta}$ is a necessity form then for every agent $i \in AG$, $B_i\boldsymbol{\theta}$ is a necessity form.*
- *if $\boldsymbol{\theta}$ is a necessity form then for every $\psi \in \mathcal{L}_{pl}$ and every group $G \subseteq AG$, $[\psi, G]\boldsymbol{\theta}$ is a necessity form.*

Note that $\sharp$ appears exactly once in every necessity form. Now for every necessity form $\boldsymbol{\theta}$ and for every formula $\varphi \in \mathcal{L}_{lauob}$, $\boldsymbol{\theta}(\varphi)$ denotes the formula obtained from $\boldsymbol{\theta}$ by replacing the unique occurrence of $\sharp$ in $\boldsymbol{\theta}$ by φ.

Proposition 4.34 and 4.39 establish the soundness and the completeness of $LAUOB$ with respect to the class of models $\mathcal{C}_0$

Proposition 4.34 *Let $\varphi \in \mathcal{L}_{lauob}$ be a formula. Then φ is a theorem of LAUOB only if φ is valid in every model of the class $\mathcal{C}_0$.*

Proof It is sufficient to show soundness of the additive axioms and inference rules. It is evident for $S0$, $S1$ and GAU. We show soundness of the rule X. Suppose that $\boldsymbol{\theta}([?, G]\varphi)$ is not valid, *i.e.* there exists a pointed model $\mathcal{M}, s$ such that $\mathcal{M}, s \models \neg\boldsymbol{\theta}([?, G]\varphi)$. Therefore there exists $\psi \in \mathcal{L}_{pl}$ such that $\mathcal{M}, s \models \neg\boldsymbol{\theta}([\psi, G]\varphi)$. Hence, $\boldsymbol{\theta}([\psi, G]\varphi)$ is not valid for this particular ψ, which implies that it is not true that for all $\psi \in \mathcal{L}_{pl}$, $\boldsymbol{\theta}([\psi, G]\varphi)$ is valid.

□

We define now the canonical model for this $LAUOB$. It is a different notion than the notion of canonical model we saw until now. This difference comes from the infinitary nature of the inference rule (X). Let us see it in details:

A set x of formulas is called a theory if it satisfies the following conditions:

- x contains the set of all theorems;
- x is closed under the rule of modus ponens and the rule (X).

Obviously, the least theory is the set of all theorems whereas the greatest theory is the set of all formulas. The latter theory is called the trivial theory. A theory x is said to be consistent if $\bot \notin x$. Let us remark that the only inconsistent theory is the set of all formulas. We shall say that a theory x is maximal if for all formulas φ, $\varphi \in x$ or $\neg\varphi \in x$. We abbreviate *mct* for maximal consistent theory.

Let x be a set of formulas. For every $\varphi \in \mathcal{L}_{lauob}$, every $\psi \in \mathcal{L}_{pl}$, every $i \in AG$ and every $G \subseteq AG$ we define:

- $x + \varphi = \{\chi \in \mathcal{L}_{lauob} \mid \varphi \to \chi \in x\}$
- $B_i x = \{\chi \in \mathcal{L}_{lauob} \mid B_i\chi \in x\}$
- $[\psi, G]x = \{\chi \in \mathcal{L}_{lauob} \mid [\psi, G]\chi \in x\}$
- $[?, G]x = \{\chi \in \mathcal{L}_{lauob} \mid [?, G]\chi \in x\}$

Lemma 4.35 *Let x be a theory, $\varphi \in \mathcal{L}_{lauob}$, $\psi \in \mathcal{L}_{pl}$, $i \in AG$ and $G \subseteq AG$. Then $x + \varphi$, $B_i x$, $[\psi, G]x$ and $[?, G]x$ are theories. Moreover $x + \varphi$ is consistent iff $\neg\varphi \notin x$.*

Proof

- $x + \varphi$ is a theory.

  First, let us prove that $x + \varphi$ contains the set of all theorems, by proving a useful property: $x \subseteq x + \varphi$. Let $\chi \in x$, we know that $\chi \to (\varphi \to \chi)$ is a theorem. By modus ponens we then have that $\chi \in x + \varphi$.

  Now let us prove that $x + \varphi$ is closed under modus ponens. Let χ_1, χ_2 be formulas such that $\chi_1 \in x + \varphi$ and $\chi_1 \to \chi_2 \in x + \varphi$. Thus $\varphi \to \chi_1 \in x$ and $\varphi \to (\chi_1 \to \chi_2) \in x$. But then $\varphi \to \chi_2 \in x$.

  Third, let us prove that $x + \varphi$ is closed under (X). Let $\boldsymbol{\theta}$ be a possibility form and ψ be a formula such that $\boldsymbol{\theta}([\psi, G]\chi) \in x + \varphi$ for all $\psi \in \mathcal{L}_{pl}$. It follows that $\varphi \to \boldsymbol{\theta}([\psi, G]\chi) \in x$ for all $\psi \in \mathcal{L}_{pl}$. Since x is a theory, then $\varphi \to \boldsymbol{\theta}([?, G]\chi) \in x$. Consequently, $\boldsymbol{\theta}([?, G]\chi) \in x + \varphi$. It follows that $x + \varphi$ is closed under (X).

- $B_i x$ is a theory.
  First, let us prove that $B_i x$ contains the set of all theorems. Let ψ be a theorem. By the necessitation of knowledge, $B_i\psi$ is also a theorem. Since x is a theory, then $B_i\psi \in x$. Therefore, $\psi \in B_i x$. It follows that $B_i x$ contains the set of all theorems. Second, let us prove that $B_i x$ is closed under modus ponens. Let ψ, χ be formulas such that $\psi \in B_i x$ and $\psi \to \chi \in K_a x$. Thus, $B_i\psi \in x$ and $B_i(\psi \to \chi) \in x$. Since $B_i\psi \to (B_i(\psi \to \chi) \to B_i\chi)$ is a theorem and x is a theory, then $B_i\psi \to (B_i(\psi \to \chi) \to B_i\chi) \in x$. Since x is closed under modus ponens, then $B_i\chi \in x$. Hence, $\chi \in B_i x$. It follows that $B_i x$ is closed under modus ponens. Third, let us prove that $B_i x$ is closed under (X). Let $\boldsymbol{\theta}$ be a necessity form, $G \subseteq AG$ be a group of agents and φ be a formula such that $\boldsymbol{\theta}([\psi, G]\varphi) \in B_i x$ for all $\psi \in \mathcal{L}_{pl}$. It follows that $B_i(\boldsymbol{\theta}([\psi, G]\varphi)) \in x$ for all $\psi \in \mathcal{L}_{pl}$. Since x is a theory, then $B_i\boldsymbol{\theta}([\psi, G]\varphi)) \in x$. Consequently, $\boldsymbol{\theta}([\psi, G]\varphi) \in B_i x$. It follows that $B_i x$ is closed under (X).

- $[\psi, G]x$ and $[?, G]x$ are theories.
  We obtain this result with the same proof than the previous one, considering that this two modal operators satisfy the axiom (K) and the necessitation rule.

- Finally, $x+\varphi$ is consistent only if $\neg\varphi \notin x$ (because $\varphi \in x+\varphi$). Reciprocally, $\perp \in (x+\varphi)$ implies that $(\varphi \to \perp) \in x$ and this implies that $\neg\varphi \in x$.

$\square$

Lemma 4.36 (Lindenbaum lemma) *Let x be a consistent theory. There exists a maximal consistent theory y such that $x \subseteq y$.*

Proof Let $\varphi_0, \varphi_1, \ldots$ be a list of the set of all formulas. We define a sequence $y_0, y_1, \ldots$ of consistent theories as follows. First, let $y_0 = x$. Second, suppose that, for some $n \geq 0$, y_n is a consistent theory containing x that has been already defined. If $y_n + \varphi_n$ is inconsistent and $y_n + \neg\varphi_n$ is inconsistent then, by lemma 4.35, $\neg\varphi_n \in y_n$ and $\neg\neg\varphi_n \in y_n$. Since $\neg\varphi_n \to (\neg\neg\varphi_n \to \perp)$ is a theorem, then $\neg\varphi_n \to (\neg\neg\varphi_n \to \perp) \in y_n$. Since y_n is closed under modus ponens, then $\perp \in y_n$: a contradiction. Hence, either $y_n + \varphi_n$ is consistent or $y_n + \neg\varphi_n$ is consistent. If $y_n + \varphi_n$ is consistent then we define $y_{n+1} = y_n + \varphi_n$. Otherwise, $\neg\varphi_n \in y_n$ and we consider two cases.

Either φ_n is not a conclusion of (X). Then, we define $y_{n+1} = y_n$.

Or φ_n is a conclusion of (X). In this case, let $\boldsymbol{\theta}([?, G_1]\chi_1)$, ..., $\boldsymbol{\theta}([?, G_k]\chi_k)$ be all the representations of φ_n as a conclusion of (X). Such representations are necessarily finitely many because there is a finite number of modal operators of the form $[?, G]$ in φ_n. We define the sequence $y_n^0, \ldots, y_n^k$ of consistent theories as follows. First, let $y_n^0 = y_n$. Second, suppose that, for some $i < k$, y_n^i is a consistent theory containing y_n that has been already defined. Then it contains $\neg\boldsymbol{\theta}([?, G_1]\chi_1) = \varphi_n$. Since y_n^i is closed under (X), then there exists a formula

$\psi \in \mathcal{L}_{pl}$ such that $\boldsymbol{\theta}([\psi, G_1]\chi_1)$ is not in y_n^i. Then, we define $y_n^{i+1} = y_n^i + \neg\boldsymbol{\theta}([\psi, G_1]\chi_1)$. Now, we put $y_{n+1} = y_n^k$. Finally, we define $y = y_0 \cup y_1 \cup \ldots$. Clearly if y is a theory then it is a maximal consistent theory such that $x \subseteq y$. Let us then prove that it is a theory.

1. It contains the set of all theorems because $x \subseteq y$

2. It is closed under modus ponens. Indeed, if $\{\chi, (\chi \longrightarrow \varphi)\} \subseteq y$ then there exists $n \in \mathbb{N}$ such that $\{\chi, (\chi \longrightarrow \varphi)\} \subseteq y_n$. Thus, y_n being a theory, we obtain $\varphi \in y_n$ and then $\varphi \in y$.

3. It is closed under (X) by construction: suppose that $\boldsymbol{\theta}([\psi, G]\chi) \in y$ for all $\psi \in \mathcal{L}_{pl}$. Let us call φ_n the formula it represented by $\boldsymbol{\theta}([?, G]\chi)$. We want to show that $\varphi_n \in y$. Suppose the opposite, $\neg\varphi_n \in y$. This means that $\neg\varphi_n \in y_n$ by construction, *i.e.* $\neg\boldsymbol{\theta}([?, G]\chi) \in y_n$. Therefore there exists a $\psi \in \mathcal{L}_{pl}$ such that $\neg\boldsymbol{\theta}([\psi, G]\chi) \in y_{n+1}$ by construction again. This is a contradiction, considering that $y_{n+1} \subseteq y$.

□

The canonical model of $LAUOB$ is the structure $\mathcal{M}_c = (W^c, \mathcal{R}^c, V^c)$ defined as follows:

- W^c is the set of all maximal consistent theories ;
- For all agents i, R_i is the binary relation on W^c defined by xR_iy iff $B_ix \subseteq y$;
- For all atoms p, $V^c(p) = \{x \in W^c \mid p \in x\}$.

Proposition 4.37 *The canonical model of $LAUOB$ is a model of the class $\mathcal{C}_0$.*

Proof Identical to the proof of Proposition 4.1.

□

Proposition 4.38 (Truth lemma) *Let φ be a formula in $\mathcal{L}_{lauob}$. Then for all maximal consistent theories x, for all $n \in \mathbb{N}$, for all $\psi_1, \ldots, \psi_n \in \mathcal{L}_{pl}$ and all $G_1, \ldots, G_n \subseteq AG$ such that $\psi_1 \wedge \ \ldots\ \wedge \psi_n \in x$, for all $\varphi \in \mathcal{L}_{lauob}$:*

$$\mathcal{UP}_n(\mathcal{M}^c, x) \models \varphi \textit{ iff } [\psi_1, G_1] \ldots [\psi_n, G_n]\varphi \in x.$$

Proof The proof is by induction on the structure of φ. The base case follows from the definition of V. The Boolean cases are trivial. It remains to deal with the modalities.

- $\varphi = B_i\chi$: Let μ be the sequence obtained from $[\psi_1, G_1] \ldots [\psi_n, G_n]$ eliminating all the $[\psi, G]$ such that $i \notin G$. Without loss of generality, we consider that $\mu = [\psi_1, G_1] \ldots [\psi_m, G_m]$ with $0 \leqslant m \leqslant n$.

  $\mathcal{UP}_n(\mathcal{M}^c, x) \not\models B_i\chi$

  iff $\mathcal{UP}_m(\mathcal{M}^c, x) \not\models B_i\chi$ (by Proposition 4.21)

  iff there exists a *mct* y such that xR_iy, $\psi_1 \wedge \ \ldots\ \wedge \psi_m \in y$ and $\mathcal{UP}_m(\mathcal{M}^c, y) \not\models \chi$

  iff (1) there exists a *mct* y s.t. $B_ix \subseteq y$ and $[\psi_1, G_1] \ldots [\psi_m, G_m]\chi \notin y$ (by IH).

Now, (1) is equivalent to (2) : $[\psi_1, G_1] \ldots [\psi_m, G_m] B_i \chi \notin x$.

Indeed, if $[\psi_1, G_1] \ldots [\psi_m, G_m] B_i \chi \in x$ then $B_i[\psi_1, G_1] \ldots [\psi_m, G_m]\chi \in x$ by Axiom $R5$, and thus $[\psi_1, G_1] \ldots [\psi_m, G_m]\chi \in y$ for any *mct* y such that $B_i x \subseteq y$.

Conversely, if $[\psi_1, G_1] \ldots [\psi_m, G_m] B_i \chi \notin x$ then $B_i[\psi_1, G_1] \ldots [\psi_m, G_m]\chi \notin x$ by Axiom $R5$. Let $y = B_i x + \neg[\psi_1, G_1] \ldots [\psi_m, G_m]\chi$. The reader may easily verify that y is a consistent theory. By Lemma 4.36, there is a maximal consistent theory z such that $y \subseteq z$. Hence, $B_i x \subseteq z$ and $[\psi_1, G_1] \ldots [\psi_m, G_m]\chi \notin z$ *Q.E.D.*

We end this case by noting that $[\psi_1, G_1] \ldots [\psi_m, G_m] B_i \chi \notin x$ is equivalent to $[\psi_1, G_1] \ldots [\psi_n, G_n] B_i \chi \notin x$ by axiom $R6$.

- $\varphi = [\psi, G]\chi : \mathcal{UP}_n(\mathcal{M}^c, x) \models [\psi, G]\chi$
  iff $\mathcal{U}_{\psi,G}(\mathcal{UP}_n(\mathcal{M}^c), (\ldots(x, 1), \ldots, 1) \models \chi$
  iff $\mathcal{UP}_{n+1}(\mathcal{M}^c, x) \models \chi$ (with an evident notation)
  iff $[\psi_1, G_1] \ldots [\psi_n, G_n][\psi, G]\varphi \in x$ by IH

- $\varphi = [?, G]\chi : \mathcal{UP}_n(\mathcal{M}^c, x) \models [?, G]\chi$
  iff for all $\psi \in \mathcal{L}_{pl}, \mathcal{UP}_n(\mathcal{M}^c, x) \models [\psi, G]\chi$
  iff for all $\psi \in \mathcal{L}_{pl}, \mathcal{UP}_n(\mathcal{M}^c, x) \models \psi$ implies $\mathcal{U}_{\psi,G}(\mathcal{UP}_n(\mathcal{M}^c, x)) \models \chi$
  iff for all $\psi \in \mathcal{L}_{pl}, [\psi_1, G_1] \ldots [\psi_n, G_n]\psi \in x$ implies $\mathcal{U}_{\psi,G}(\mathcal{UP}_n(\mathcal{M}^c, x)) \models \chi$ by IH
  iff for all $\psi \in \mathcal{L}_{pl}, \psi_1 \wedge \ldots \psi_n \wedge \psi \in x$ (recall $\psi \in \mathcal{L}_{pl}$) implies $\mathcal{UP}_{n+1}(\mathcal{M}^c, x) \models \chi$
  iff for all $\psi \in \mathcal{L}_{pl}, \psi_1 \wedge \ldots \psi_n \wedge \psi \in x$ implies $[\psi_1, G_1] \ldots [\psi_n, G_n][\psi, G]\chi \in x$ by IH
  iff for all $\psi \in \mathcal{L}_{pl}, \psi_1 \wedge \ldots \psi_n \in x$ implies $\begin{cases} \psi \in x \text{ and } [\psi_1, G_1] \ldots [\psi, G]\chi \in x \\ \psi \notin x \text{ and } [\psi_1, G_1] \ldots [\psi, G]\chi \in x \end{cases}$
  iff $\psi_1 \wedge \ldots \psi_n \in x$ implies for all $\psi \in \mathcal{L}_{pl}$, $[\psi_1, G_1] \ldots [\psi_n, G_n][\psi, G]\chi \in x$
  iff $\psi_1 \wedge \ldots \psi_n \in x$ implies for all $\psi \in \mathcal{L}_{pl}$, $[\psi_1, G_1] \ldots [\psi_n, G_n][?, G]\chi \in x$ (by (X))

□

Theorem 4.39 *The axiomatization LAUOB is sound and complete with respect to the class of models $\mathcal{C}_0$.*

Proof Soundness has been proved in Proposition 4.34.

Let $\varphi \in \mathcal{L}_{lauob}$ be a a valid formula, then it is valid in the canonical model. Therefore by Proposition 4.38 it is in every maximal consistent theory. Hence, it is a theorem of $LAUOB$. Indeed, if it were not the case, then there would exists a consistent theory x such that $\neg\varphi \in x$. Therefore, by Lemma 4.36 there exists a *mct* y such that $x \subseteq y$. Therefore $\neg\varphi \in y$. Contradiction.

□

Clearly, $\mathcal{L}_{lauob}$ is at least as expressive as $\mathcal{L}_{luob}$. However is $\mathcal{L}_{lauob}$ more expressive than $\mathcal{L}_{luob}$? To answer this question, we consider the formula $\chi := \langle ?, \{i, j\}\rangle(B_i p \wedge \neg B_j B_i p)$ and the models $\mathcal{M}$ and $\mathcal{M}_q$ (members of $\mathcal{C}_0$) presented in Figure 4.4.

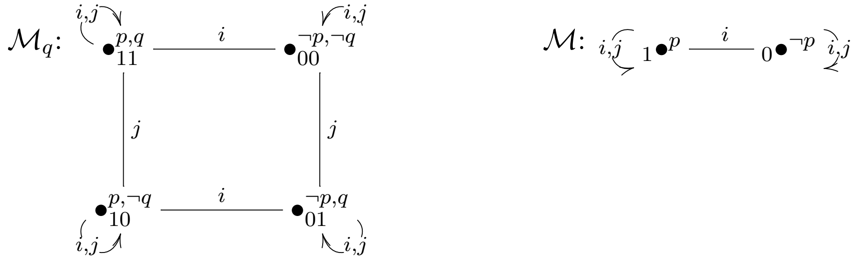


Figure 4.4: Distinguishing $\mathcal{L}_{lauob}$ from $\mathcal{L}_{luob}$

We let the reader see that $(\mathcal{M}_q, 11) \models \langle ?, \{i,j\}\rangle(B_i p \wedge \neg B_j B_i p)$ and $(\mathcal{M}, 1) \not\models \langle ?, \{i,j\}\rangle(B_i p \wedge \neg B_j B_i p)$. Let us suppose that χ is equivalent to a formula $\chi' \in \mathcal{L}_{luob}$. Then χ' would be satisfied in the same way in every model composing any couple of bisimilar models with respect to the language restricted to the atoms appearing in χ'. Let us take an atom q that does not appear in χ'. Thus $(\mathcal{M}_q, 11)$ and $(\mathcal{M}, 1)$ are bisimilar with respect to the language restricted to the atoms appearing in χ'. Then the formula χ is not equivalent to any formula of $\mathcal{L}_{luob}$. That means that $\mathcal{L}_{lauob}$ is more expressive than $\mathcal{L}_{luob}$ (it contains at least χ in extra).

Neither we have results on the finite model property of this logic, nor on the complexity of the problem of satisfiability. But we get the following:

Proposition 4.40 *The problem of the model checking of $\mathcal{L}_{lauob}$ with respect to $\mathcal{C}_0$ is PSPACE-complete.*

Proof This proof is analogous to the proof of the complexity of the model checking of $\mathcal{L}_{gal}$ with respect to the models of $\mathcal{M}_{S5}$ (Theorem 5.33). See Section 5 to get more details.

□

4.5 Case Study

We examine the situation presented in Figure 4.5. Brune and Alex are in a Poker final face to face. Brune gets a pair of Kings in her hand. She knows it is a very good game. Alex has the ace of spades, and another spades. When Brune proposes her bet, Alex checks to see what will happen. The game becomes particularly interesting when, as for this example, several players have a good hand and imagine easily to have the winning one.

Let us specify our language in this case. We pose $AG = \{a, b\}$ for Alex and Brune, and $PROP = \{VC_i | V \in \{1, 2, 3, 4, \ldots, J, Q, K\}, C \in \{\clubsuit, \heartsuit, \diamondsuit, \spadesuit\}, i \in AG \cup \{t\}\} \cup \{\Phi_a\}$. We read VC_i by 'the present deal gives the card of value V and color C to i". When i is t (for table), or to be more readable when i is missing, it means that the card is (or will be) dealt on the table. Φ_a means that Alex has the winning hand (following poker rules). We abbreviate $\Phi_b := \neg\Phi_a$.

Figure 4.5: Brune wins with a "poker"

We consider the initial poker model $\mathcal{M}_{\mathcal{P}}$. All possible deals are considered (as states of the model). Any agent can distinguish two states if and only her own cards are different. More precisely, we pose $\mathcal{M}_{\mathcal{P}} = (S, V, \{R_i\}_{i \in AG})$. Thus, we have: $S = \{(c_a, c_b, \{p^5, p^6, p^7\}, p^8, p^9) \mid c_i = \{p_i^1, p_i^2\}$, the propositions appearing are all different and the p_a^j are of VC_a form $\}$. Also, for every proposition p, $V(p)$ is the set of deals in which p appears. Finally, R_i links two states if and only if the c_i is identical in both deals. Note that we have $\binom{52}{2} \times \binom{50}{2} \times \binom{48}{3} \times 45 \times 44 = 5.56 \times 10^{13}$ different possible deals. Therefore this is a gigantic model of fifty million of millions of states.

This representation clearly has some limits. A poker player probably does not represent herself all the possible deals; she would rather think in terms of probability to win. Nevertheless, this probability corresponds to a simplified representation of this huge model. Furthermore, a poker game includes other kinds of communicative acts besides announcements, that give information on the players' intentions or feelings. These communicative acts are probably the heart of the game, and we cannot formalize them here. Yet, our position is not to propose a formalism that contains all the characteristics of poker, but only the aspects linked to the notion of objective belief that are clearly part of this game.

We examine two different situations, presented in Figure 4.6, situations in which the initial cards received by Brune (c_b) are different, but the other cards (dealt to Alex and dealt on the table) are identical. Alex has the ace and the 9 of spades. In the first case (d_1) Brune has a pair of Kings (diamonds and clubs), in the second (d_2) she has a pair of 2 (diamonds and hearts). In both cases, the cards dealt on the table are first the 7, the 2 and the King of spades, then the King of hearts and eventually the 9 of diamonds.

In both contexts, we use the following abbreviations:

- $FLOP := 7\spadesuit \wedge 2\spadesuit \wedge K\spadesuit$

Figure 4.6: Two deals: d_1 and d_2

- $TURN := FLOP \wedge K\heartsuit$
- $RIVER := TURN \wedge 9\diamondsuit$

Here are some formulas that are true in this model, in the first context (Brune having a pair of Kings)

- $\mathcal{M}_{\mathcal{P}}, d_1 \models \langle FLOP, \{a, b\}\rangle(B_a 7\spadesuit \wedge B_a B_b 7\spadesuit)$: After the flop, Alex believes that the 7 of spades is dealt on the table, and that Brune believes the same fact.
- $\mathcal{M}_{\mathcal{P}}, d_1 \models \langle FLOP, \{a, b\}\rangle B_b [R\heartsuit, \{a, b\}]\Phi_b$: After the flop, Brune believes that if the King of hearts is dealt, she will have the winning hand.
- $\mathcal{M}_{\mathcal{P}}, d_1 \models \langle TURN, \{a, b\}\rangle B_b \Phi_b$: After the turn, Brune believes she has the winning hand

What is the difference between the two situations? In both cases, both players have a very good hand (Alex has a flush and Brune either a poker or a full house). In both cases also, Brune has the winning hand. The difference is that at some moment of the game (here after the turn), Brune will be sure that she has the winning hand. In other (formal) words:

- $\mathcal{M}_{\mathcal{P}}, d_1 \models \langle ?, \{a,b\}\rangle(B_b\Phi_b \wedge \hat{B}_a\neg B_b\Phi_b)$
- $\mathcal{M}_{\mathcal{P}}, d_2 \models [?, \{a,b\}](\Phi_b \wedge \neg B_b\Phi_b)$

Now let us consider the possibility that a player cheats by looking at the other player's hand. With that possibility, for the deal d_1, Alex could learn that Brune believes that she has the winning hand. While for the deal d_2 he could learn instead that she does not have the information that she has the winning hand:

- $\mathcal{M}_{\mathcal{P}}, d_1 \models \langle ?, a\rangle(B_aB_b\Phi_b)$
- $\mathcal{M}_{\mathcal{P}}, d_2 \models \langle ?, a\rangle B_a(\Phi_b \wedge \neg B_b\Phi_b)$

In this second case he could be tempted to try a bluff. But Brune may cheat as well. In that case she would know that she has a winning hand, but Alex may never get this information, whatever he could learn afterwards (even by cheating):

- $\mathcal{M}_{\mathcal{P}}, d_2 \models \langle ?, b\rangle(B_b\Phi_b \wedge [?, a]\hat{B}_a\neg B_b\Phi_B)$

Chapter 5

Group Announcement Logic

The previous sections worked on the kind of information that is given during a communication. The following one deals with the results a group of agents can achieve by announcing something: agents of the group G can obtain φ if they each of them can do a (simultaneous) announcement such that after such announcements φ becomes true.

[van Benthem, 2004] and [Balbiani et al., 2007] suggested an interpretation of the standard modal diamond where $\Diamond\varphi$ means "there is an announcement after which φ" (see Section 3.1). This was in a setting going back to the Fitch-paradox (see [Brogaard and Salerno, 2004]). The new interpretation of the diamond $\Diamond$ in the Fitch setting firstly interprets $\Diamond\varphi$ as 'sometime later, φ', and secondly specifies this temporal specification as what may result of a specific event, namely a public announcement: 'after some announcement, φ'. In other words, the semantics is: $\Diamond\varphi$ is true if and only if $\langle\psi\rangle\varphi$ is true for some ψ; the expression $\langle\psi\rangle\varphi$ stands for 'ψ is true and after ψ is announced, φ is true.' There are some restrictions on ψ. The resulting arbitrary announcement logic is axiomatisable and has various pleasing properties (see [Balbiani et al., 2007], and for more detail the extended journal version [Balbiani et al., 2008]). Arbitrary announcement logic makes no assumption on the interpretation of $\Diamond\varphi$ about *who* makes the announcement, or indeed whether or not the announcement *can* be truthfully made by anyone. In the current chapter we investigate a variant of arbitrary announcement logic. Instead of $\Diamond\varphi$ we use a more specific operator, namely $\langle G\rangle\varphi$. Here G is a subgroup of all agents *that simultaneously make truthful public announcements*, i.e., announcements of formulas they know. In other words, let $G = \{1, \ldots, k\}$, then: $\langle G\rangle\varphi$ is true if and only if there exist formulas $\psi_1, \ldots, \psi_k$ such that $\langle K_1\psi_1 \wedge \ldots K_k\psi_k\rangle\varphi$ is true; now, the expression $\langle K_1\psi_1 \wedge \ldots K_k\psi_k\rangle\varphi$ stands for $K_1\psi_1 \wedge \ldots K_k\psi_k$ is true and after agents $1, \ldots, k$, simultaneously announce $\psi_1, \ldots, \psi_k$, then φ is true'. Note that the remaining agents, not included in the set G of k agents, are not involved in making the announcement, although they are aware of that action happening. The resulting logic is called *Group Announcement Logic (GAL)*. Informally speaking, $\langle G\rangle\varphi$ expresses the fact that coalition G has the ability to make φ come about. Logics modelling the coalitional abilities of agents have been an active area of research in multi-agent systems in recent years, the most prominent frameworks being Pauly's Coalition Logic ([Pauly, 2002]) and Alur, Henzinger and Kupferman's Alternating-time Temporal Logic ([Alur et al., 2002]). The main constructs of these logics are indeed of the form $\langle G\rangle\varphi$ with the intuitive meaning that coalition G can achieve φ. In this chapter we investigate these notions when the actions that can be performed are truthful public announcements.

Section 5.1 defines group announcement logic, presents various interaction axioms between the different modalities that express intuitive properties of such joint announcements, and the axiomatization. Section 5.2 is entirely devoted to expressivity matters, and Section 5.3 to model checking. The relation between group announcement logic and various notions of group ability, including knowledge 'de re' and knowledge 'de dicto', is discussed in detail in Section 5.4, which is followed by a more applied Section 5.5 that embeds these observations into security protocols for two agents (sender and receiver) in the presence of a finite number of eavesdroppers intercepting all communications between them. Most of this chapter has been published in [Ågotnes et al., 2010].

5.1 Group Announcement Logic

The main construct of the language of *Group Announcement Logic (GAL)* is $\langle G\rangle\varphi$, intuitively meaning that there is some announcement the group G can truthfully make after which φ will be true. Such a simultaneous announcement may sound like a lot of unintelligible noise. But in fact it merely means a *joint public action*—not necessarily involving talking. We later find ways to model subsequent announcements as sequences of simultaneous actions, making the basic semantic idea even less appear as shouting in groups.

5.1.1 Language

The language $\mathcal{L}_{gal}$ of GAL over a set of propositions $PROP$ and a set of agents AG is defined by extending the language $\mathcal{L}_{pal}$ of PAL (introduced in Section 3.1.1) with a new operator $[G]$ for each coalition G:

Definition 5.1 (Language)

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid K_i\varphi \mid [\varphi_1]\varphi_2 \mid [G]\varphi$$

where $i \in AG$ is an agent, $G \subseteq AG$ is a set of agents and $p \in PROP$. We write $\langle G\rangle\varphi$ for the dual $\neg[G]\neg\varphi$ and $\langle i\rangle\varphi$ for $\langle\{i\}\rangle\varphi$. For the subset of atoms occurring in a formula φ we, again, write Θ_φ.

We adopt the standard definition for the notion of subformula.

5.1.2 Semantics

The interpretation of formulas in a pointed Kripke structure is defined by extending the definition for PAL (see Definition 3.2) with a clause for the new operator:

Definition 5.2 (Semantics of GAL)

$\mathcal{M}, s \models [G]\varphi$ *iff for every set* $\{\psi_i : i \in G\} \subseteq \mathcal{L}_{el}$, $\mathcal{M}, s \models [\bigwedge_{i\in G} K_i\psi_i]\varphi$

We get the following meaning for the dual $\langle G\rangle\varphi := \neg[G]\neg\varphi$:

$$\mathcal{M}, s \models \langle G\rangle\varphi \text{ iff there exists a set } \{\psi_i : i \in G\} \subseteq \mathcal{L}_{el} \text{ such that } \mathcal{M}, s \models \langle \bigwedge_{i\in G} K_i\psi_i\rangle\varphi$$

If we write this out in detail, we get: $\mathcal{M}, s \models \langle G\rangle\varphi$ iff there exists a set $\{\psi_i : i \in G\} \subseteq \mathcal{L}_{el}$ such that $\mathcal{M}, s \models \bigwedge_{i\in G} K_i\psi_i$ and $\mathcal{M}|\bigwedge_{i\in G} K_i\psi_i, s \models \varphi$.

Observe that $\langle G\rangle$ quantifies only over purely epistemic formulas. The reason for this is as follows. First, in the semantics of $\langle G\rangle\varphi$ the formulas ψ_i in $\bigwedge_{i\in G} K_i\psi_i$ *cannot* be unrestricted $\mathcal{L}_{gal}$ formulas, as that would make the definition circular: such a ψ_i could then be the formula $\langle G\rangle\varphi$ itself that we are trying to interpret. We therefore avoid quantifying over formulas containing $\langle G\rangle$ operators. However, as public announcement logic is equally expressive as epistemic logic within the class of all models ([Plaza, 1989]), the semantics obtained by quantifying over the fragment of the language without $\langle G\rangle$ operators is the same as the semantics obtained by quantifying only over epistemic formulas.

As usual, a formula φ is *valid on* $\mathcal{M}$, notation $\mathcal{M} \models \varphi$, iff $\mathcal{M}, s \models \varphi$ for all s in the domain of $\mathcal{M}$; and a formula φ is *valid*, $\models \varphi$, iff $\mathcal{M} \models \varphi$ for all $\mathcal{M}$. The denotation of φ on $\mathcal{M}$, notation $[\![\varphi]\!]_{\mathcal{M}}$ is defined as $\{s \in S \mid \mathcal{M}, s \models \varphi\}$. The set of validities of the logic is called GAL (group announcement logic).

Proposition 5.3 *Let two models $\mathcal{M} = (S, \mathcal{R}, V)$ and $\mathcal{M}' = (S', \mathcal{R}', V')$ be given. Let $\varphi \in \mathcal{L}_{gal}$ be a formula. For all $s \in S$ and for all $s' \in S'$, if $(\mathcal{M}, s) \underline{\leftrightarrow} (\mathcal{M}', s')$ then $\mathcal{M}, s \models \varphi$ iff $\mathcal{M}', s' \models \varphi$.*

Proof The proof is by induction on the number n of group announcement modality that appear in the formula. If $n = 0$, it is the epistemic case, already underlined in Proposition 2.14. Now let us suppose that it is true for all formula with at most $n-1$ group announcement operators and let us prove for any formula with at most n by induction on the structure of φ. The base case is by the the main IH, the boolean cases are trivial. For the epistemic modality, as in [Fagin et al., 1995], we use the **back** and **forth** conditions in the definition of bisimulation. It remains to deal with the group announcement modality:

$$\begin{array}{ll} & \mathcal{M}, s \models \langle G\rangle\varphi \\ \text{iff} & \text{there exists } \psi_1, \ldots, \psi_n \in \mathcal{L}_{el} \text{ such that } \mathcal{M}, s \models \langle \bigwedge_{i\in AG} K_i\psi_i\rangle\varphi \\ \text{iff} & \text{there exists } \psi_1, \ldots, \psi_n \in \mathcal{L}_{el} \text{ such that } \mathcal{M}', s' \models \langle \bigwedge_{i\in AG} K_i\psi_i\rangle\varphi \text{ by the main IH} \\ \text{iff} & \mathcal{M}', s' \models \langle G\rangle\varphi \end{array}$$

$\square$

5.1.3 Logical Properties

To sharpen the intuition about the logic we mention some relevant validities, with particular attention to interaction between group announcement and epistemic modal operators. Examples are $\models [G]\varphi \longrightarrow [G][G]\varphi$ (Corollary 5.6), $\models \langle G\rangle[G]\varphi \rightarrow [G]\langle G\rangle\varphi$ (Corollary 5.12), and $K_i[i]\varphi \longleftrightarrow [i]K_i\varphi$ (Proposition 5.13).

Elementary validities

Proposition 5.4

1. $\langle G\rangle p \to p$ *and* $\langle G\rangle \neg p \to \neg p$. *(atomic propositions do not change value)*
2. $\langle \emptyset\rangle \varphi \leftrightarrow \varphi$ *and* $[\emptyset]\varphi \leftrightarrow \varphi$ *(the empty group is powerless)*
3. $\langle K_{j_1}\psi_{j_1} \wedge \cdots \wedge K_{j_k}\psi_{j_k}\rangle \varphi \to \langle \{j_1, \ldots, j_k\}\rangle \varphi$
4. $\varphi \longrightarrow \langle G\rangle \varphi$ *(truth axiom)*

The easy proofs are ommited. They use the following ideas:

1. In public announcement logic, and its 'derivatives', factual truths never change value.
2. The conjunction of an empty set of formulas is, as usual, taken to be a tautology.
3. Obvious (note that $\psi_{j_1}, \ldots, \psi_{j_k}$ are purely epistemic formulas).
4. If all agents announce 'true', nothing changes to the system.

An announcement by the empty group (the second property above) corresponds to a "clock tick", a dynamic transition without informative effect. We could also see this as "nobody says a thing" (and this now happens...). In fact you could even see this as 'everybody says true', an announcement by the public (as in the fourth property): in other words, the group of all agents have the option *not* to exercise their power.

Sequences of group announcements Intuitively, $\langle G\rangle\varphi$ means that G can achieve a situation where φ is true in "one step", by making a joint announcement. One can easily imagine situations where it could be interesting to reason about what a group can achieve by making *repeated* announcements, i.e., by a sequence of announcements, one after the other, or a communication protocol. A general example is a conversation over an open channel. We want to express that "there is some sequence, of arbitrary length, of announcements by G which will ensure that φ becomes true".

For arbitrary public announcement logic ($APAL$), the validity of the principle $\Box\varphi \longrightarrow \Box\Box\varphi$ follows from the simple observation that a sequence of two announcements ψ and χ is equivalent to the single announcement of $\psi \wedge [\psi]\chi$ (see [Plaza, 1989]). Less obvious is that $[G]\varphi \longrightarrow [G][G]\varphi$ is also valid, because now we have to show that two conjunctions of *known* formulas are again such a conjunction.

Proposition 5.5 $\models [G \cup H]\varphi \longrightarrow [G][H]\varphi$

Proof The diamond version $\langle G\rangle\langle H\rangle\varphi \longrightarrow \langle G \cup H\rangle\varphi$ of this validity makes clear that the requirement is that two successive announcements respectively by the agents in G simultaneously and in H simultaneously can also be seen as a single announcement by the agents in $G \cup H$ simultaneously. Let us prove how it can be done. Consider two successive announcements $\bigwedge_{i \in G} K_i \varphi_i$ and $\bigwedge_{j \in H} K_j \psi_j$. Let a Kripke structure $\mathcal{M}$ and a state s in $\mathcal{M}$ be given such that $\mathcal{M}, s \models \bigwedge_{i \in G} K_i \varphi_i$, and similarly $\bigwedge_{j \in H} K_j \psi_j$ is true in state s in the restriction of $\mathcal{M}$ to the $\bigwedge_{i \in G} K_i \varphi_i$-states: $\mathcal{M} | \bigwedge_{i \in G} K_i \varphi_i, s \models \bigwedge_{j \in H} K_j \psi_j$.

Then we have:

$\mathcal{M}, s \models \langle \bigwedge_{i \in G} K_i \varphi_i \rangle \langle \bigwedge_{j \in H} K_j \psi_j \rangle \theta$

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G} K_i \varphi_i \wedge [\bigwedge_{g \in G} K_g \varphi_g] \bigwedge_{j \in H} K_j \psi_j \rangle \theta$

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G} K_i \varphi_i \wedge \bigwedge_{i \in H \backslash G} K_i \top \wedge [\bigwedge_{g \in G} K_g \varphi_g] (\bigwedge_{j \in H} K_j \psi_j \wedge \bigwedge_{j \in G \backslash H} K_j \top) \rangle \theta$
because for any agent i, $K_i \top$ is a valid formula

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G \cup H} (K_i \varphi_i \wedge [\bigwedge_{g \in G} K_g \varphi_g] K_i \psi_i) \rangle \theta$
with $\forall i \in H \backslash G$, $\varphi_i = \top$ and $\forall j \in G \backslash H$, $\psi_j = \top$.

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G \cup H} (K_i \varphi_i \wedge ((\bigwedge_{g \in G} K_g \varphi_g) \longrightarrow K_i [\bigwedge_{g \in G} K_g \varphi_g] \psi_i)) \rangle \theta$
by a reduction axiom of PAL

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G \cup H} K_i \varphi_i \wedge \bigwedge_{i \in G \cup H} ((\bigwedge_{j \in G} K_j \varphi_j) \longrightarrow K_i [\bigwedge_{j \in G} K_j \varphi_j] \psi_i) \rangle \theta$
by distributing the $\wedge$

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G \cup H} K_i \varphi_i \wedge \bigwedge_{i \in G \cup H} K_i [\bigwedge_{j \in G} K_j \varphi_j] \psi_i \rangle \theta$
because $\bigwedge_{j \in G} K_j \varphi_j$ is assumed true in the left conjunct of the announcement.

only if $\mathcal{M}, s \models \langle \bigwedge_{i \in G \cup H} K_i (\varphi_i \wedge [\bigwedge_{j \in G} K_j \varphi_j] \psi_i) \rangle \theta$.

□

Corollary 5.6 $\models [G]\varphi \longrightarrow [G][G]\varphi$

We thus get exactly the property alluded to above:

Corollary 5.7 $\mathcal{M}, s \models \langle G \rangle \varphi$ *iff there is a finite sequence of announcements by agents in G after which* φ *is true.*

In Section 5.5 we discuss a security protocol example involving sequences of announcements. Note that our result does *not* mean that sequences of announcements can simply be replaced by a single announcement: whether agents are willing to do an announcement may depend on the postconditions of such announcements. These may be *known to be satisfied* after each announcement in the sequence, but not *known to be satisfied* initially after the entire sequence. These matters will be discussed in great detail later.

Church-Rosser We prove that for all groups G and H of agents, for every formula $\varphi \in \mathcal{L}_{gal}$, $\langle G\rangle[H]\varphi \rightarrow [H]\langle G\rangle\varphi$ is a valid formula. The principle is fairly intuitive: it says that when in a given epistemic state group G or group H make a group announcement, there are additional announcements by group H (after G's announcement) and group G (after H's announcement), in order to reach a new common state of information. Unfortunately, its proof is rather involved. This is because group announcements implicitly quantify over *all* propositional variables in the language. Towards the proof, we first define the *group-announcement depth* $d(\varphi)$ of a formula φ:

Let $p \in PROP$, $\psi, \psi_1, \psi_2 \in \mathcal{L}_{gal}$, $i \in N$, and $G \subseteq N$ be given; then $d(p) = 0$; $d(\neg\psi) = d(K_i\psi) = d(\psi)$; $d(\psi_1 \wedge \psi_2) = d([\psi_1]\psi_2) = max(d(\psi_1), d(\psi_2))$; and $d([G]\psi) = d(\psi) + 1$. The following lemma holds for any number k, but we will only use it for $k \leq |AG|$.

Lemma 5.8 *Let $Q = \{q_i\}_{i\in\mathbb{N}^\star} \subseteq PROP$ be pairwise distinct primitive propositions, for some $k \in \mathbb{N}$ let $\theta_1, \ldots, \theta_k$ be epistemic formulas such that for $i = 1$ to $i = k$, $\Theta_{\theta_i} \cap Q = \emptyset$ and let $\varphi \in \mathcal{L}_{gal}$ be such that $\Theta_\varphi \cap Q = \emptyset$.*

For all $\psi \in \mathcal{L}_{gal}$, define $\begin{cases} \psi^\alpha = \psi(\theta_1/q_1, .., \theta_k/q_k, q_1/q_{k+1}, q_2/q_{k+2}, ..) \\ \psi^{-\alpha} = \psi(q_{k+1}/q_1, q_{k+2}/q_2, ..) \end{cases}$

Then, for all structures $\mathcal{M} = (S, \sim_1, \ldots, \sim_n, V)$ there is a valuation function $V' : PROP \rightarrow 2^S$ such that

1. $[\![\varphi]\!]_\mathcal{M} = [\![\varphi]\!]_{\mathcal{M}'}$
2. *for all $\psi \in \mathcal{L}_{el}$,*
   - $[\![\psi]\!]_{\mathcal{M}'} = [\![\psi^\alpha]\!]_\mathcal{M}$
   - $[\![\psi]\!]_\mathcal{M} = [\![\psi^{-\alpha}]\!]_{\mathcal{M}'}$
3. *for all $i \leq k, [\![q_i]\!]_{\mathcal{M}'} = [\![\theta_i]\!]_{\mathcal{M}'} = [\![\theta_i]\!]_\mathcal{M}$*

where $\mathcal{M}' = (S, \sim_1, \ldots, \sim_n, V')$.

Proof We define V' as: $\begin{cases} V'(p) = V(p), \text{ for all } p \notin Q \\ V'(q_i) = [\![\theta_i]\!]_\mathcal{M}, \text{ for all } i \leq k \\ V'(q_{k+i}) = V(q_i), \text{ for all } i \geq 1 \end{cases}$

Items 2 and 3 follow directly from the definition of V'. We prove item 1 by induction on the structure of φ, by showing the somewhat stronger following property $P(\varphi)$:

for all submodels $\mathcal{M}_*$ of $\mathcal{M}$, and for all states $s \in \mathcal{M}_*$: $\mathcal{M}_*, s \models \varphi$ iff $\mathcal{M}'_*, s \models \varphi$.

Base case: $\varphi = p \in PROP$. $\mathcal{M}_*, s \models p$ iff $\mathcal{M}'_*, s \models p$ follows directly from the definition of V'. We also have that $\mathcal{M}_*, s \models \bot$ iff $\mathcal{M}'_*, s \models \bot$

Inductive cases: Let us suppose that the property P is true for all eventual ψ, ψ_1 and ψ_2, and let us prove it for formulas φ of the form $\neg\psi, \psi_1 \wedge \psi_2, \hat{K}_i\psi$ and $[\psi_1]\psi_2$. In fact, let $\mathcal{M}_*$ be a submodel of $\mathcal{M}$:

- $\neg\psi$: $\mathcal{M}_*, s \models \neg\psi$
  iff $\mathcal{M}_*, s \not\models \psi$
  iff $\mathcal{M}'_*, s \not\models \psi$ (by IH)
  iff $\mathcal{M}'_*, s \models \neg\psi$.

- $\psi_1 \wedge \psi_2$: $\mathcal{M}_*, s \models \psi_1 \wedge \psi_2$
  iff $(\mathcal{M}_*, s \models \psi_1$ and $\mathcal{M}_*, s \models \psi_2)$
  iff $(\mathcal{M}'_*, s \models \psi_1$ and $\mathcal{M}'_*, s \models \psi_2)$ (by IH)
  iff $\mathcal{M}'_*, s \models \psi_1 \wedge \psi_2$.

- $K_i\psi$: $\mathcal{M}_*, s \models K_i\psi$
  iff for all $t \sim_i s$, $\mathcal{M}_*, t \models \psi$
  iff for all $t \sim'_i s$, $\mathcal{M}_*', t \models \psi$ (by IH and as $\sim_i = \sim'_i$)
  iff $\mathcal{M}'_*, s \models K_i\psi$.

- $[\psi_1]\psi_2$: $\mathcal{M}_*, s \models [\psi_1]\psi_2$
  iff $(\mathcal{M}_*, s \models \psi_1$ implies $\mathcal{M}_*|\psi_1, s \models \psi_2)$
  iff $(\mathcal{M}'_*, s \models \psi_1$ implies $(\mathcal{M}_*|\psi_1)', s \models \psi_2)$(using IH twice)
  iff $(\mathcal{M}'_*, s \models \psi_1$ implies $\mathcal{M}'_*|\psi_1, s \models \psi_2)$
  (using IH again for ψ_1, and that V' on the restriction is the restriction of V')
  iff $\mathcal{M}'_*, s \models [\psi_1]\psi_2$.

- $\langle G\rangle\psi$:
  $\mathcal{M}_*, s \models \langle G\rangle\psi$
  only if there are $\chi_1, \ldots, \chi_{|G|}$ in $\mathcal{L}_{el}$ such that $\mathcal{M}_*, s \models \langle \bigwedge K_i\chi_i\rangle\psi$
  only if there exists $\{\chi_i\} \subseteq \mathcal{L}_{el}$ s.t. $\mathcal{M}_*, s \models \bigwedge K_i\chi_i$ and $\mathcal{M}_*|(\bigwedge K_i\chi_i), s \models \psi$
  only if there exists $\{\chi_i\}$ s.t $\mathcal{M}'_*, s \models (\bigwedge K_i\chi_i)^{-\alpha}$ and $(\mathcal{M}_*|(\bigwedge K_i\chi_i))', s \models \psi$ (by IH)
  only if there exists $\{\chi_i\}$ s.t $\mathcal{M}'_*|(\bigwedge K_i\chi_i)^{-\alpha}, s \models \psi$ (**)
  only if there are $\chi_1, \ldots, \chi_{|G|}$ in $\mathcal{L}_{el}$ such that $\mathcal{M}'_*, s \models \langle \bigwedge K_i\chi_i^{-\alpha}\rangle\psi$
  only if $\mathcal{M}'_*, s \models \langle G\rangle\psi$.
  $\mathcal{M}'_*, s \models \langle G\rangle\psi$
  only if there are $\chi_1, \ldots, \chi_{|G|}$ in $\mathcal{L}_{el}$ such that $\mathcal{M}'_*, s \models \langle \bigwedge K_i\chi_i\rangle\psi$
  only if there exists $\{\chi_i\} \subseteq \mathcal{L}_{el}$ s.t. $\mathcal{M}'_*, s \models \bigwedge K_i\chi_i$ and $\mathcal{M}'_*|(\bigwedge K_i\chi_i), s \models \psi$
  only if there exists $\{\chi_i\}$ s.t. $\mathcal{M}_*, s \models (\bigwedge K_i\chi_i)^{\alpha}$ and $(\mathcal{M}_*|(\bigwedge K_i\chi_i)^{\alpha})', s \models \psi$ (**)
  only if there are $\chi_1, \ldots, \chi_{|G|}$ in $\mathcal{L}_{el}$ such that $\mathcal{M}_*|(\bigwedge K_i\chi_i)^{\alpha}, s \models \psi$ (by IH)
  only if there are $\chi_1, \ldots, \chi_{|G|}$ in $\mathcal{L}_{el}$ such that $\mathcal{M}_*, s \models \langle \bigwedge K_i\chi_i^{\alpha}\rangle\psi$
  only if $\mathcal{M}_*, s \models \langle G\rangle\psi$.

  In (**) we have used (the already shown) property 2 for epistemic formulas.

$\square$

Proposition 5.9 *Let $k \geq 0$, ψ and χ be $\mathcal{L}_{gal}$-formulas and $G = \{i_1, \ldots, i_k\}$ be a set of agents. If $\mathcal{M}, s \models \langle G\rangle\psi \wedge \chi$ and $p_1, ..., p_k \in PROP\backslash(\Theta_\psi \cup \Theta_\chi)$, then there is a $\mathcal{M}''$ different* *from $\mathcal{M}$ only on the valuation of the atoms of $PROP \backslash (\Theta_\psi \cup \Theta_\chi)$ such that $\mathcal{M}'', s \models \langle K_{i_1} p_1 \wedge \cdots \wedge K_{i_k} p_k \rangle \psi \wedge \chi$.*

Proof We use the previous lemma twice:

1. Let Q be $PROP \backslash (\Theta_\psi \cup \Theta_\chi)$ and q_i be p_i for all $i \leqslant k$, θ_i be $\top$ for all $i \leqslant k$ and φ be $\langle G \rangle \psi$. By Lemma 5.8, there is V' such that $V'(p_i) = S$ and $[\![\langle G \rangle \psi \wedge \chi]\!]_{\mathcal{M}'} = [\![\langle G \rangle \psi \wedge \chi]\!]_{\mathcal{M}}$. As $\mathcal{M}, s \models \langle G \rangle \psi \wedge \chi$, we have that $\mathcal{M}', s \models \langle G \rangle \psi \wedge \chi$. Therefore there are $\tau_1, \ldots, \tau_k$ in $\mathcal{L}_{el}$ such that $\mathcal{M}', s \models \langle \bigwedge_{i \in G} K_i \tau_i \rangle \psi \wedge \chi$. Without loss of generality, we can assume that for all i, j, $p_i \notin \Theta_{\tau_j}$. Indeed, for all i, p_i is equivalent to $\top$ in $\mathcal{M}'$, therefore we can replace τ_j by $\tau_j(\top / p_i)$.

2. Let Q be $PROP \backslash (\Theta_\psi \cup \Theta_\chi \cup \bigcup_{i \in G} \Theta_{\tau_i})$, $k = |G|$, q_i be p_i for all $i \leqslant k$, θ_i be τ_i for all $i \leqslant k$ and φ be $\langle \bigwedge_{i \in G} K_i \tau_i \rangle \psi \wedge \chi$. By Lemma 5.8, there is V'' such that (with $\mathcal{M}''$ as $\mathcal{M}$ except for valuation V'') $[\![\langle \bigwedge K_i \tau_i \rangle \psi \wedge \chi]\!]_{\mathcal{M}''} = [\![\langle \bigwedge K_i \tau_i \rangle \psi \wedge \chi]\!]_{\mathcal{M}'}$ and for all $i \leq k$, $[\![p_i]\!]_{\mathcal{M}''} = [\![\tau_i]\!]_{\mathcal{M}''} = [\![\tau_i]\!]_{\mathcal{M}'}$. As $\mathcal{M}', s \models \langle \bigwedge_{i \in G} K_i \tau_i \rangle \psi \wedge \chi$ (by the first item), the first property implies that $\mathcal{M}'', s \models \langle \bigwedge_{i \in G} K_i \tau_i \rangle \psi \wedge \chi$. The second one implies that $\mathcal{M}'' \models \bigwedge K_i p_i \leftrightarrow \bigwedge K_i \tau_i$. We now have that $\mathcal{M}'', s \models \langle \bigwedge_{i \in G} K_i p_i \rangle \psi \wedge \chi$.

□

A generalization of Proposition 5.9 indirectly proves the soundness of a derivation rule in the axiomatization of GAL. Here, we need Proposition 5.9 to prove the validity of the generalized Church-Rosser schema. But first here is a useful remark:

Remark 5.10 *Let $\mathcal{M}, s$ be a pointed model, $p \in PROP$ and $i \in AG$. If $\mathcal{M}, s \models K_i p$, then for all $\varphi \in \mathcal{L}_{el}$ such that $\mathcal{M}, s \models \varphi$ we have $\mathcal{M}|\varphi, s \models K_i p$.*

Proposition 5.11 (Church-Rosser Generalized) *For any $G, H \subseteq AG$: $\models \langle G \rangle [H] \varphi \to [H] \langle G \rangle \varphi$.*

Proof Suppose the contrary: Let $\mathcal{M}$ be a model, s a state of $\mathcal{M}$, $\varphi \in \mathcal{L}_{gal}$ and $G, H \subseteq AG$ two groups of agents such that $\mathcal{M}, s \models \langle G \rangle [H] \varphi \wedge \langle H \rangle [G] \neg \varphi$. Then, using Proposition 5.9 twice, for $|G| = k$ and $|H| = k'$, we know that there are $\{p_i\}_{i \in G}$ and $\{q_i\}_{i \in H}$ subsets of Θ and $\mathcal{M}'$ differing from $\mathcal{M}$ only on the valuation of the p_i, q_i such that

$$\mathcal{M}', s \models \langle \bigwedge_{i \in G} K_i p_i \rangle [H] \varphi \wedge \langle \bigwedge_{i \in H} K_i q_i \rangle [G] \neg \varphi \ .$$

In particular,

$$\mathcal{M}', s \models \langle \bigwedge_{i \in G} K_i p_i \rangle [\bigwedge_{i \in H} K_i q_i] \varphi \wedge \langle \bigwedge_{i \in H} K_i q_i \rangle [\bigwedge_{i \in G} K_i p_i] \neg \varphi \ .$$

Note that $\langle\bigwedge_{i\in G} K_i p_i\rangle$ and $\langle\bigwedge_{i\in H} K_i q_i\rangle$ are conjunctions of known facts. Using Remark 5.10 we know that they remain true after further announcements. So we have

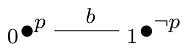
$$\mathcal{M}', s \models \langle \bigwedge_{i\in G} K_i p_i \wedge \bigwedge_{i\in H} K_i q_i\rangle\varphi \wedge \langle \bigwedge_{i\in G} K_i p_i \wedge \bigwedge_{i\in H} K_i q_i\rangle\neg\varphi$$

from which directly follows a contradiction.

$\square$

Corollary 5.12 (Church-Rosser) $\models \langle G\rangle[G]\varphi \rightarrow [G]\langle G\rangle\varphi$

We cannot in general reverse the order of G and H in Proposition 5.11. A simple counterexample is the following model, where b cannot distinguish between two states but a can.



We now have that $\mathcal{M}, 0 \models \langle a\rangle[b]K_b p \wedge \langle a\rangle[b]\neg K_b p$ because $\mathcal{M}, 0 \models \langle K_a p\rangle[b]K_b p \wedge \langle K_a \top\rangle[b]\neg K_b p$. Therefore $\langle G\rangle[H]\varphi \rightarrow [G]\langle H\rangle\varphi$ is *not* valid if $G = \{a\}$ and $H = \{b\}$.

More validities Just as for Church-Rosser, one would like to know whether the *APAL* validity $\square\lozenge\varphi \longrightarrow \lozenge\square\varphi$ has a *GAL* generalization. We know that there exists $G, H \subseteq N$ such that the schema $[G]\langle H\rangle\varphi \rightarrow \langle H\rangle[G]\varphi$ is *not* valid. A counterexample is the following model $\mathcal{M}$ (Figure 5.1) , i.e. for $G = \{a\}$ and $H = \{b\}$, with $\varphi = (K_a K_b p \vee K_b K_a q) \wedge \neg(K_a K_b p \wedge K_b K_a q)$.

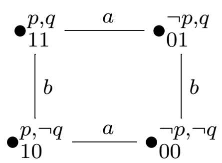


Figure 5.1: Counter-example of the validity of MacKinsey formula in GAL

φ asserts that a knows b knows p or b knows a knows q, but not both facts at the same time. Here we have $\mathcal{M}, 11 \models [a]\langle b\rangle\varphi \wedge [b]\langle a\rangle\neg\varphi$. Indeed, in this state, a can teach q to b and b can teach p to a. Thus, depending on what one agent does (to teach or not the corresponding fact), the other can decide whether teaching or not her knowledge. If both or none of them decide to teach her knowledge then φ will be true, if only one does it then φ will be false. The second agent speaking is the one who decides!

We do not know whether $[G]\langle G\rangle\varphi \rightarrow \langle G\rangle[G]\varphi$ is valid.

For arbitrary announcement logic we have that $K_i\square\varphi \longrightarrow \square K_i\varphi$, but not the other way round. Now, we can do more.

Proposition 5.13 *For arbitrary $i \in AG$ and $G \subseteq AG$:*

1. $\models K_i[i]\varphi \longleftrightarrow [i]K_i\varphi$

2. $\models K_i[G]\varphi \longrightarrow [G]K_i\varphi$ *(but not the other way round)*

Proof

1. For every model $\mathcal{M}$ and every state s, we have

$$\begin{array}{ll} & \mathcal{M}, s \models K_i[i]\varphi \\ \text{iff} & \text{for all } t \in S \text{ such that } sR_it,\ \mathcal{M}, t \models [i]\varphi \\ \text{iff} & \text{for all } \psi \in \mathcal{L}_{el}, \text{ for all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models [K_i\psi]\varphi \\ \text{iff} & \text{for all } \psi \in \mathcal{L}_{el}, \text{ for all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models K_i\psi \text{ implies } \mathcal{M}|K_i\psi, t \models \varphi \\ \text{iff} & \text{for all } \psi \in \mathcal{L}_{el}, \text{ for all } t \text{ s.t. } sR_it,\ \mathcal{M}, s \models K_i\psi \text{ implies } \mathcal{M}|K_i\psi, t \models \varphi \\ \text{iff} & \text{for all } \psi \in \mathcal{L}_{el},\ \mathcal{M}, s \models K_i\psi \text{ implies that for all } t \text{ s.t. } sR_it,\ \mathcal{M}|K_i\psi, t \models \varphi \\ \text{iff} & \mathcal{M}, s \models [i]K_i\varphi \end{array}$$

2. For every model $\mathcal{M}$ and every state s, we have

$$\begin{array}{ll} & \mathcal{M}, s \models K_i[G]\varphi \\ \text{only if} & \text{for all } t \in S \text{ such that } sR_it,\ \mathcal{M}, t \models [G]\varphi \\ \text{only if} & \text{for all } \psi_1, \ldots, \psi_n \in \mathcal{L}_{el}, \text{ for all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models [\bigwedge_{j\in\{1,\ldots,n\}} K_j\psi_j]\varphi \\ \text{only if} & \text{for all } \{\psi_j\} \subset \mathcal{L}_{el}, \text{ all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models \bigwedge K_j\psi_j \text{ implies } \mathcal{M}|\bigwedge K_j\psi_j, t \models \varphi \\ \text{only if} & \text{for all } \{\psi_j\} \subset \mathcal{L}_{el}, \text{ all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models \bigwedge K_j\psi_j \text{ implies } \mathcal{M}|\bigwedge K_j\psi_j, t \models \varphi \\ \text{only if} & \text{for all } \{\psi_j\} \subset \mathcal{L}_{el},\ \mathcal{M}, s \models \bigwedge K_j\psi_j \text{ implies that} \\ & \text{for all } t \text{ s.t. } sR_it,\ \mathcal{M}, t \models \bigwedge K_j\psi_j \text{ implies } \mathcal{M}|\bigwedge K_j\psi_j, t \models \varphi \\ \text{only if} & \mathcal{M}, s \models [G]K_i\varphi \end{array}$$

□

Finally, a rather puzzling property on the interaction between the announcements and knowledge by two agents. The intuition behind it is that announcements wherein you can make another agent learn facts even in the face of your own uncertainty, are rather rare.

Proposition 5.14 *For any atomic proposition* $p \in PROP$*:* $\models \langle a\rangle K_b p \leftrightarrow \langle b\rangle K_a p$.

Proof Assume $\mathcal{M}, s \models \langle a\rangle K_b p$. Then there is a $\psi_a \in \mathcal{L}_{el}$ such that $\mathcal{M}, s \models \langle K_a\psi_a\rangle K_b p$. This formula is equivalent to $K_a\psi_a \wedge (K_a\psi_a \rightarrow K_b[K_a\psi_a]p)$ and thus to $K_a\psi_a \wedge K_b(K_a\psi_a \rightarrow p)$—as p is an atom. Let us note $\mathcal{M}' = \mathcal{M}|K_b(K_a\psi_a \rightarrow p)$ and let us proof that $\mathcal{M}', s \models K_a p$. Indeed, let $t \in \mathcal{M}'$ s.t. $t \in \mathcal{R}'_a(x)$ (in the restricted model) and let us prove that $\mathcal{M}', t \models p$. But we have (1) $\mathcal{M}, t \models K_b(K_a\psi_a \rightarrow p)$ and (2) $t \in \mathcal{R}_a(x)$ (in the non-restricted model). But (1) implies that $\mathcal{M}, t \models K_a\psi_a \rightarrow p$ and (2) implies that $\mathcal{M}, t \models K_a\psi_a$ (because $\mathcal{M}, s \models K_a\psi_a$). Then $\mathcal{M}, t \models p$, and thus $\mathcal{M}', t \models p$.

□

We now proceed to a more systematic treatment of validities.

5.1.4 Axiomatization

The following is a sound and complete axiomatization of group announcement logic.

Definition 5.15 (*GAL* axioms and rules)

instantiations of propositional tautologies	
$K_i(\varphi \longrightarrow \psi) \longrightarrow (K_i\varphi \longrightarrow K_i\psi)$	*distribution (of knowl. over impl.)*
$K_i\varphi \longrightarrow \varphi$	*truth (T)*
$K_i\varphi \longrightarrow K_iK_i\varphi$	*positive introspection (4)*
$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi$	*negative introspection (5)*
$[\varphi]p \longleftrightarrow (\varphi \longrightarrow p)$	*atomic permanence*
$[\varphi]\neg\psi \longleftrightarrow (\varphi \longrightarrow \neg[\varphi]\psi)$	*announcement and negation*
$[\varphi](\psi \vee \chi) \longleftrightarrow ([\varphi]\psi \vee [\varphi]\chi)$	*announcement and disjunction*
$[\varphi]K_i\psi \longleftrightarrow (\varphi \longrightarrow K_i[\varphi]\psi)$	*announcement and knowledge*
$[\varphi][\psi]\chi \longleftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	*announcement composition*
$[G]\varphi \longrightarrow [\bigwedge_{i\in G} K_i\psi_i]\varphi$ *where* $\psi_i \in \mathcal{L}_{el}$	*group announcement*
From φ *and* $\varphi \longrightarrow \psi$, *infer* ψ	*modus ponens*
From φ, *infer* $K_i\varphi$	*necessitation of knowledge*
From φ, *infer* $[\psi]\varphi$	*necessitation of announcement*
From φ, *infer* $[G]\varphi$	*necessitation of group announcement*
From $\boldsymbol{\eta}([\wedge_{i\in G}K_i\psi_i]\chi)$ *for all* $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, *infer* $\boldsymbol{\eta}([G]\chi)$	*deriving group announcement /* $R^w([G])$

Our axiomatization of *GAL* is based on the standard $S5$ axioms for the epistemic operators K_i, the standard reduction axioms for the public announcement operators $[\varphi]$, and some additional axioms and derivation rules involving group announcement operators. These are the axiom *group and specific announcement*, and the derivation rules *necessitation of group announcement* and *deriving group announcement.* A formula $\varphi \in \mathcal{L}_{gal}$ is derivable, notation $\vdash \varphi$, iff φ belongs to the least set of formulas containing *GAL* axioms and closed with respect to the derivation rules.

The axiom $[G]\varphi \longrightarrow [\bigwedge_{i\in G} K_i\psi_i]\varphi$, where $\psi_i \in \mathcal{L}_{el}$, is obviously valid in all structures. Also the validity of "from φ, infer $[G]\varphi$" will be obvious. The derivation rule *deriving group announcement* $R^w([G])$ is used to introduce group announcement operators in derivations. In this rule, $\boldsymbol{\eta}$ is a necessity form for $\mathcal{L}_{gal}$, definable in the same way as in Definition 4.33 (but with a different language): $\sharp$ is a necessity form; if $\boldsymbol{\eta}$ is a necessity form and and φ is in $\mathcal{L}_{gal}$ then $(\varphi \rightarrow \boldsymbol{\eta})$ is a necessity form; if $\boldsymbol{\eta}$ is a necessity form and φ is in $\mathcal{L}_{gal}$ then $[\varphi]\boldsymbol{\eta}$ is a necessity form; if $\boldsymbol{\eta}$ is a necessity form then $K_i\boldsymbol{\eta}$ is a necessity form.

In this section we show completeness of *GAL* with respect to the class of epistemic models. But first the following:

Proposition 5.16 (Soundness) *Let* $\varphi \in \mathcal{L}_{gal}$. *Then* φ *is a theorem of GAL only if* φ *is valid in every epistemic model.*

Proof The only difficult result is the soundness of the rule $R^w(G)$ in the class of epistemic models. We thus show it by induction on the structure of the necessity form $\boldsymbol{\eta}$. The base case comes from the definition of the semantics. Let us look at the inductive cases:

- $\boldsymbol{\eta} = K_i\boldsymbol{\eta}'$: for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, $\mathcal{M}, s \models K_i\boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, and all sR_it, $\mathcal{M}, t \models \boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff for all sR_it, $\mathcal{M}, t \models \boldsymbol{\eta}'([G]\chi)$ by IH
  iff $\mathcal{M}, s \models K_i\boldsymbol{\eta}'([G]\chi)$
- $\boldsymbol{\eta} = \varphi \longrightarrow \boldsymbol{\eta}'$: for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, $\mathcal{M}, s \models \varphi \longrightarrow \boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff if $\mathcal{M}, s \models \varphi$ then for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, $\mathcal{M}, s \models \boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff if $\mathcal{M}, s \models \varphi$ then $\mathcal{M}, s \models \boldsymbol{\eta}'([G]\chi)$ by IH
  iff $\mathcal{M}, s \models \varphi \longrightarrow \boldsymbol{\eta}'([G]\chi)$
- $\boldsymbol{\eta} = [\varphi]\boldsymbol{\eta}'$: for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, $\mathcal{M}, s \models [\varphi]\boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff if $\mathcal{M}, s \models \varphi$ then for all $\{\psi_i\}_{i\in G} \subset \mathcal{L}_{el}$, $\mathcal{M}|\varphi, s \models \boldsymbol{\eta}'([\wedge_{i\in G}K_i\psi_i]\chi)$
  iff if $\mathcal{M}, s \models \varphi$ then $\mathcal{M}|\varphi, s \models \boldsymbol{\eta}'([G]\chi)$ by IH
  iff $\mathcal{M}, s \models [\varphi]\boldsymbol{\eta}'([G]\chi)$

□

The proof of completeness of the axiomatization is very similar to the proof of completeness of the axiomatization $LAUOB$ presented in Section 4.4.2. We follow its progress.

A set x of formulas is called a theory if it satisfies the following conditions:

- x contains the set of all theorems;
- x is closed under the rule of modus ponens and the rule $R^\omega([G])$.

Again, a theory x is said to be consistent if $\bot \notin x$, maximal if for all formulas φ, $\varphi \in x$ or $\neg\varphi \in x$. For all formulas φ we note $x + \varphi = \{\psi\colon \varphi \to \psi \in x\}$. For all agents i, let $K_ix = \{\varphi\colon K_i\varphi \in x\}$. For all formulas φ, let $[\varphi]x = \{\psi\colon [\varphi]\psi \in x\}$.

Lemma 5.17 *Let x be a theory, φ be a formula, and a be an agent. Then $x + \varphi$, K_ix and $[\varphi]x$ are theories. Moreover $x + \varphi$ is consistent iff $\neg\varphi \notin x$.*

Proof Identical to proof of Proposition 4.35, by substituting B_i by K_i and '$[\psi, G]$ with $\psi \in \mathcal{L}_{pl}$' by '$[\psi]$ with $\psi \in \mathcal{L}_{el}$'.

□

Lemma 5.18 (Lindenbaum lemma) *Let x be a consistent theory. There exists a maximal consistent theory y such that $x \subseteq y$.*

Proof Identical to the proof of Lemma 4.36 by substituting

- B_i by K_i

- '$[\psi, G]$ with $\psi \in \mathcal{L}_{pl}$' by '$[\psi]$ with $\psi \in \mathcal{L}_{el}$'
- $[?, G]$ by $[G]$ and
- 'there exists $\psi \in \mathcal{L}_{pl}$ such that $[\psi, G]\varphi$' by 'there exist $\psi_1, \ldots, \psi_{|G|} \in \mathcal{L}_{el}$ such that $[\bigwedge K_i\psi_i]\varphi$'

$\square$

The canonical model of $\mathcal{L}_{gal}$ is the structure $\mathcal{M}_c = (W^c, \sim^c, V^c)$ defined as follows:

- W^c is the set of all maximal consistent theories;
- For all agents i, $\sim_i$ is the binary (equivalence) relation on W defined by $x \sim_i y$ iff $K_i x = K_i y$;
- For all atoms p, $V^c(p)$ is the subset of W^c defined by $x \in V^c(p)$ iff $p \in x$.

Clearly, '=' is an equivalence relation, therefore $\sim_i$ also is. It then has the same properties than the relation R_i of an epistemic model, and that fact ensures that the canonical model of $\mathcal{L}_{gal}$ is an epistemic model. Note that, because of Axioms T, 4 and 5, $K_i x = K_i y$ iff $K_i x \subseteq y$.

We now prove a truth lemma for $\mathcal{L}_{gal}$ using a very special induction. Note that this property (and its proof) is quite different from the truth lemma appearing in [Ågotnes et al., 2010]. The property was not correctly proved in this paper indeed. To obtain a correct one, here are three definitions we shall use:

Definition 5.19 *The degree of a $\mathcal{L}_{gal}$-formula is defined inductively in the following way: for all $p \in PROP$, all $i \in AG$, all $G \subseteq AG$ and all $\varphi, \varphi_1, \varphi_2 \in \mathcal{L}_{gal}$: $deg(p) = deg(\bot) = 0$, $deg(\neg\varphi) = deg(K_i\varphi) = deg(\varphi)$, $deg\varphi_1 \vee \varphi_2 = max(deg(\varphi_1), deg(\varphi_2))$, $deg([\varphi_1]\varphi_2) = deg(\varphi_1) + deg(\varphi_2) + 2$, $deg([G]\varphi) = deg(\varphi) + 2$.*

Definition 5.20 *Let $\ll$ be the following binary relation on $\mathbb{N} \times \mathbb{N} \times \mathcal{L}_{gal}$:*

$$(k', n', \varphi') \ll (k, n, \varphi) \textit{ iff } \begin{cases} k' < k, \\ \textit{or } (k' = k \textit{ and } n' < n) \\ \textit{or } (k' = k \textit{ and } n' = n \textit{ and } \varphi' \in Sub(\varphi)) \end{cases}$$

*where $Sub(\varphi)$ is the set of strict subformulas of φ (*i.e.* subformulas of φ different from φ itself).*

Note that $\ll$ is a well-founded partial order on $\mathbb{N} \times \mathbb{N} \times \mathcal{L}_{gal}$. In fact, it is the lexicographical order based on the orders $\leqslant$ and 'being a subformula'. Here are some examples:

- $(0, 10^{18}, \varphi) \ll (1, 0, \varphi)$ because $0 < 1$
- $(0, 0, (p \wedge \neg q) \longrightarrow \langle G\rangle[p]q) \ll (0, 1, p)$ because $0 < 1$
- $(10, 10, p) \ll (10, 10, p \wedge q)$ because p is a subformula of $p \wedge q$.

Now the following:

Definition 5.21 ($\boldsymbol{\Pi(k,n,\varphi)}$) *For all $(k,n,\varphi) \in \mathbb{N} \times \mathbb{N} \times \mathcal{L}_{gal}$ we define the property $\Pi(k,n,\varphi)$: for all $\psi_1,\dots,\psi_n \in \mathcal{L}_{gal}$, if $n + deg(\psi_1) + \ldots + deg(\psi_n) + deg(\varphi) \leqslant k$ then for all $x \in W^c$, $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\varphi$ iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\varphi \in x$.*

This complex definition is useful to prove the following truth lemma by induction on (k,n,φ) following the order we just defined. We are now able to establish the truth lemma for $\mathcal{L}_{gal}$:

Proposition 5.22 (Truth lemma) *For all $(k,n,\varphi) \in \mathbb{N} \times \mathbb{N} \times \mathcal{L}_{gal}$, $\Pi(k,n,\varphi)$.*

To prove it, we first consider the following lemma:

Lemma 5.23 *For all $(k,n,\varphi) \in \mathbb{N} \times \mathbb{N} \times \mathcal{L}_{gal}$, if for all $(k',n',\varphi') \ll (k,n,\varphi)$ we have $\Pi(k',n',\varphi')$, then $\Pi(k,n,\top)$.*

Proof Let $\psi_1,\dots,\psi_n \in \mathcal{L}_{gal}$, be such that $n + \Sigma deg(\psi_i) + deg(\varphi) \leqslant k$ and $x \in W^c$, we want to show that $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\top$ iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\top \in x$. If $n = 0$, $\mathcal{M}_c, x \models \top$ is always true, and so is $\top \in x$. Suppose then that $n \geqslant 1$.

Hence we have: $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\top$

iff $\begin{cases} \mathcal{M}_c, x \models \psi_1 \\ \mathcal{M}_c, x \models \langle\psi_1\rangle\psi_2 \\ \ldots \\ \mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_{n-1}\rangle\psi_n \end{cases}$ iff $\begin{cases} \psi_1 \in x \text{ by } \Pi(k,0,\psi_1) \\ \langle\psi_1\rangle\psi_2 \in x \text{ by } \Pi(k,1,\psi_2) \\ \ldots \\ \langle\psi_1\rangle\ldots\langle\psi_{n-1}\rangle\psi_n \in x \text{ by } \Pi(k,n-1,\psi_n) \end{cases}$ if and only if $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\top \in x$.

We used that for all $i \leqslant n$, we have $n + deg(\psi_1) + \ldots + deg(\psi_i) \leqslant n + deg(\psi_1) + \ldots + deg(\psi_n) + deg(\varphi) \leqslant k$

$\square$

Proof (of Proposition 5.22) Let us prove it by induction on (k,n,φ). Suppose that for all $(k',n',\varphi') \ll (k,n,\varphi)$ we have $\Pi(k',n',\varphi')$. Let us prove $\Pi(k,n,\varphi)$ by reasoning on the form of φ. Note that by Lemma 5.23 we can already use that $\Pi(k,n,\top)$.

Let $\psi_1,\dots,\psi_n \in \mathcal{L}_{gal}$ be such that $n + \Sigma deg(\psi_i) + deg(\varphi) \leqslant k$ and $x \in W^c$

- $\varphi = p$: $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle p$
  iff $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\top$ and $\mathcal{M}_c, x \models p$ by the semantics
  iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\top \in x$ by $\Pi(k,n,\top)$ and $p \in x$ by definition of the valuation V^c
  iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle p \in x$

- $\varphi = \neg\chi$: $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\neg\chi$
  iff $\mathcal{M}_c, x \models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\top$ and $\mathcal{M}_c, x \not\models \langle\psi_1\rangle\ldots\langle\psi_n\rangle\chi$ by the semantics
  iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\top \in x$ by $\Pi(k,n,\top)$ and $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\chi \notin x$ by $\Pi(k,n,\chi)$
  (note that $n + \Sigma deg(\psi_i) + deg(\chi) = n + \Sigma deg(\psi_i) + deg(\neg\chi) \leqslant k$)
  iff $\langle\psi_1\rangle\ldots\langle\psi_n\rangle\neg\chi \in x$

- $\varphi = \chi_1 \vee \chi_2$: $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle(\chi_1 \vee \chi_2)$
  iff $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\chi_1$ or $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\chi_2$ by the semantics
  iff $\langle\psi_1\rangle \dots \langle\psi_n\rangle\chi_1 \in x$ by $\Pi(k, n, \chi_1)$ or $\langle\psi_1\rangle \dots \langle\psi_n\rangle\chi \in x$ by $\Pi(k, n, \chi_2)$
  (note that $n + \Sigma deg(\psi_i) + deg(\chi_i) \leqslant n + \Sigma deg(\psi_i) + deg(\varphi) \leqslant k$)
  iff $\langle\psi_1\rangle \dots \langle\psi_n\rangle(\chi_1 \vee \chi_2) \in x$

- $\varphi = K_i\chi$: First remark that we can prove $\Pi(k, n, \neg\chi)$ as in the first second case, considering that $Sub(K_i\chi) = Sub(\neg\chi)$.

  Now suppose that $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle K_i\chi$, therefore we have the two following properties: $\begin{cases} \mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \\ \text{for all } y \in W^c, \text{ if } xR_i^c y \text{ then } \mathcal{M}_c, y \models [\psi_1] \dots [\psi_n]\chi \end{cases}$
  The first implies that $\langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x$ by $\Pi(k, n, \top)$. Now suppose, towards a contradiction, that $\langle\psi_1\rangle \dots \langle\psi_n\rangle K_i\chi \notin x$. Then $[\psi_1] \dots [\psi_n]\neg K_i\chi \in x$ and using $\langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x$ we obtain $\langle\psi_1\rangle \dots \langle\psi_n\rangle\neg K_i\chi \in x$ and thus $\neg K_i[\psi_1] \dots [\psi_n]\chi \in x$. Let $y_0 = K_i x + \neg[\psi_1] \dots [\psi_n]\chi$, y_0 is thus a consistent theory, that can be extended, by Lemma 5.18, to a maximal consistent theory y. Therefore, $xR_i y$ and we obtain $\mathcal{M}_c, y \models [\psi_1] \dots [\psi_n]\chi$ or equivalently $\mathcal{M}_c, y \not\models \langle\psi_1\rangle \dots \langle\psi_n\rangle\neg\chi$. By $\Pi(k, n, \neg\chi)$, this implies that $\langle\psi_1\rangle \dots \langle\psi_n\rangle\neg\chi \notin y$. Contradiction.

  Conversely, suppose that $\langle\psi_1\rangle \dots \langle\psi_n\rangle K_i\chi \in x$. In particular $\langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x$ and $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\top$ by $\Pi(k, n, \top)$. We also have $K_i[\psi_1] \dots [\psi_n]\chi \in x$. Suppose towards a contradiction that $\mathcal{M}_c, x \not\models \langle\psi_1\rangle \dots \langle\psi_n\rangle K_i\chi$. Then $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\neg K_i\chi$. Therefore, there exists $y \in W^c$ such that $xR_i^c y$ and $\mathcal{M}_c, y \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\neg\chi$. By $\Pi(k, n, \neg\chi)$ we obtain $\langle\psi_1\rangle \dots \langle\psi_n\rangle\neg\chi \in y$. Now, $\mathcal{M}_c, x \models K_i[\psi_1] \dots [\psi_n]\chi$ and $xR_i y$ implies $\mathcal{M}_c, y \models [\psi_1] \dots [\psi_n]\chi$. With $\mathcal{M}_c, y \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\top$ we get $\mathcal{M}_c, y \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\chi$. By $\Pi(k, n, \chi)$ this implies that $\langle\psi_1\rangle \dots \langle\psi_n\rangle\chi \in y$. Contradiction.

- $\varphi = [\psi]\chi$: Here we have $n + \Sigma deg(\psi_i) + deg(\psi) + deg(\chi) + 2 \leqslant k$ $(*)$.

  We first prove (I): $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\psi\rangle\neg\chi$ iff $\langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\psi\rangle\neg\chi \in x$. Indeed, as $(k-1, n+1, \neg\chi) \ll (k, n, \varphi)$ we have $\Pi(k-1, n+1, \neg\chi)$. Moreover, $(n+1) + (\Sigma deg(\psi_i) + deg(\psi)) + deg(\neg\chi) \leqslant k - 1$ by $(*)$.
  Now $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle[\psi]\chi$
  iff $\begin{cases} \mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \\ \mathcal{M}_c, x \not\models \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\psi\rangle\neg\chi \end{cases}$
  iff $\begin{cases} \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x \text{ by } \Pi(k, n, \top) \\ \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\psi\rangle\neg\chi \notin x \text{ by } (I) \end{cases}$
  iff $\langle\psi_1\rangle \dots \langle\psi_n\rangle[\psi]\chi \in x$

- $\varphi = [G]\chi$: Here we have $n + \Sigma deg(\psi_i) + deg(\chi) + 2 \leqslant k$ $(**)$.

  We first prove (II): for all $\psi'_1, \dots, \psi'_{|G|} \in \mathcal{L}_{el}$, $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\bigwedge_{i\in G} K_i\psi'_i\rangle\neg\chi$ iff $\langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\bigwedge_{i\in G} K_i\psi'_i\rangle\neg\chi \in x$. Indeed, as $(k-1, n+1, \neg\chi) \ll (k, n, \varphi)$ we have

$\Pi(k-1, n+1, \neg\chi)$. Moreover, $(n+1) + (\Sigma deg(\psi_i) + deg(\bigwedge_{i\in G} K_i\psi'_i)) + deg(\neg\chi) \leqslant k-1$ by $(**)$ and observing that $deg(\bigwedge_{i\in G} K_i\psi'_i) = 0$.

Now $\mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle[G]\chi$

iff $\begin{cases} \mathcal{M}_c, x \models \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \\ \mathcal{M}_c, x \not\models \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle G\rangle\neg\chi \end{cases}$

iff $\begin{cases} \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x \text{ by } \Pi(k, n, \top) \\ \text{for all } \psi'_1, \dots, \psi'_{|G|} \in \mathcal{L}_{el},\ \mathcal{M}_c, x \not\models \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\bigwedge_{i\in G} K_i\psi'_i\rangle\neg\chi \end{cases}$

iff $\begin{cases} \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x \\ \text{for all } \psi'_1, \dots, \psi'_{|G|} \in \mathcal{L}_{el},\ \langle\psi_1\rangle \dots \langle\psi_n\rangle\langle\bigwedge_{i\in G} K_i\psi'_i\rangle\neg\chi \notin x \text{ by } (II) \end{cases}$

iff $\begin{cases} \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x \\ \text{for all } \psi'_1, \dots, \psi'_{|G|} \in \mathcal{L}_{el},\ [\psi_1] \dots [\psi_n][\bigwedge_{i\in G} K_i\psi'_i]\chi \in x \text{ by maximality} \end{cases}$

iff $\begin{cases} \langle\psi_1\rangle \dots \langle\psi_n\rangle\top \in x \\ [\psi_1] \dots [\psi_n][G]\chi \in x \end{cases}$

considering $\left(\begin{array}{l} \text{-the } R^w([G])\text{-cloture of } x \text{ for the direct implication} \\ \text{-Axiom group announcement for the indirect one} \end{array} \right.$

iff$\langle\psi_1\rangle \dots \langle\psi_n\rangle[G]\chi \in x$.

□

Theorem 5.24 *The axiomatization GAL is sound and complete with respect to the class of models $\mathcal{C}_0$.*

Proof Soundness has been proved in Proposition 5.16.

Let $\varphi \in \mathcal{L}_{gal}$ be a a valid formula, then it is valid in the canonical model. Therefore by Lemma 5.22 it is in every maximal consistent theory. Hence, it is a theorem of *GAL*. Indeed, if it were not the case, then there would exists a consistent theory x such that $\neg\varphi \in x$. Therefore, by Lemma 5.18 there exists a *mct* y such that $x \subseteq y$. Therefore $\neg\varphi \in y$. Contradiction.

□

5.2 Expressivity

The notion used in this section have been introduced in Section 2.8. Known results are that $\mathcal{L}_{el}$ is equally expressive as $\mathcal{L}_{pal}$ ([Plaza, 1989]), that in the single-agent situation $\mathcal{L}_{apal}$ is equally expressive as $\mathcal{L}_{pal}$, and that in the multi-agent situation $\mathcal{L}_{apal}$ is more expressive than $\mathcal{L}_{pal}$ ([Balbiani et al., 2007]). In this section we demonstrate that in the single-agent situation $\mathcal{L}_{gal}$ is equally expressive as $\mathcal{L}_{el}$, and that in the multi-agent situation $\mathcal{L}_{gal}$ is more expressive than $\mathcal{L}_{el}$, and $\mathcal{L}_{gal}$ is not more expressive than $\mathcal{L}_{apal}$. We conjecture that $\mathcal{L}_{apal}$ is not as least as expressive as (multi-agent) $\mathcal{L}_{gal}$.

Proposition 5.25 *For a single agent $\mathcal{L}_{gal}$ is equally expressive as $\mathcal{L}_{el}$ and $\mathcal{L}_{pal}$.*

Proof Let a be the unique agent. For all φ in $\mathcal{L}_{gal}$ we have that $\models [a]\varphi \leftrightarrow \varphi$. In fact, in the single agent situation, the restriction of a pointed model $(\mathcal{M}, s)$ to a-equivalence class is the submodel generated from $(\mathcal{M}, s)$ (see Definition 2.12). Thus it is bisimilar to $(\mathcal{M}, s)$ (Proposition 2.15), from which directly follows that $\models [a]\varphi \longleftrightarrow \varphi$.

□

Theorem 5.26 *If $n \geqslant 2$, then $\mathcal{L}_{gal}$ is more expressive than $\mathcal{L}_{el}$ and $\mathcal{L}_{pal}$.*

Proof $\mathcal{L}_{gal}$ is obviously at least as expressive as $\mathcal{L}_{el}$. For the strictness part, consider the formula $\langle b \rangle K_a p$. Assume that there is an *EL* formula ψ equivalent to $\langle b \rangle K_a p$. Formula ψ can only contain a finite number of atoms. Let q be an atom not occurring in ψ. Consider the following models $\mathcal{M}$ and $\mathcal{M}'$ where a and b have common knowledge of their ignorance of p.

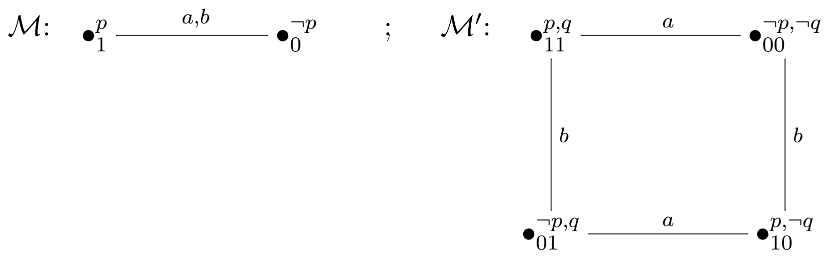


Figure 5.2: Distinguishing $\mathcal{L}_{gal}$ from $\mathcal{L}_{el}$

It is easy to see that $\mathcal{M}, 1 \not\models \langle b \rangle K_a p$, but that $\mathcal{M}', 11 \models \langle K_b q \rangle K_a p$, and thus that $\mathcal{M}', 11 \models \langle b \rangle K_a p$. On the other hand, $(\mathcal{M}, 1)$ and $(\mathcal{M}', 11)$ are bisimilar with respect to the epistemic language not including atom q, thus ψ cannot distinguish between these two pointed models. Therefore, ψ cannot be equivalent to $\langle b \rangle K_a p$.

□

Theorem 5.27 *$\mathcal{L}_{gal}$ is not at least as expressive as $\mathcal{L}_{apal}$.*

Proof Consider the $\mathcal{L}_{apal}$-formula $\Diamond(K_a p \wedge \neg K_b K_a p)$, and suppose there is an equivalent $\mathcal{L}_{gal}$-formula χ. Assume an atomic proposition q not occurring in χ. We prove that the pointed models $(\mathcal{M}, 10)$ and $(\mathcal{M}^a, 10)$ presented below cannot be distinguished by any $\mathcal{L}_{gal}$-formula χ, whereas $\Diamond(K_a p \wedge \neg K_b K_a p)$ is true in the former but false in the latter, thus again deriving a contradiction.

The crucial insight is that in $\mathcal{L}_{gal}$, unlike in $\mathcal{L}_{apal}$, the *only* definable model restrictions of $\mathcal{M}, 10$ are the four models displayed below. And to make the formula $\Diamond(K_a p \wedge \neg K_b K_a p)$ true in $\mathcal{M}, 10$, one needs to be able to define the restriction to domain $\{11, 10, 00\}$. The formal proof is by induction on the structure of ψ, and is formulated in terms also involving other points of other model restrictions of $\mathcal{M}$; of that proof we only give the formal proposition and for state 10 the inductive cases for announcement and for group announcement.

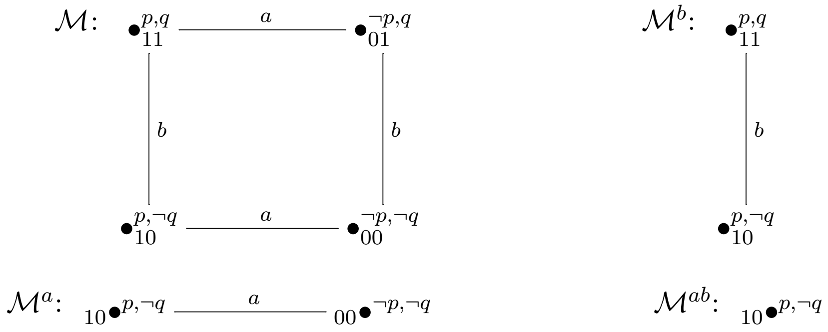


Figure 5.3: Distinguishing $\mathcal{L}_{apal}$ from $\mathcal{L}_{gal}$

Let $\psi \in \mathcal{L}_{gal}$ with $q \notin \Theta_\psi$. Then:

$$\begin{array}{lllllll}
\mathcal{M}, 10 \models \psi & \Leftrightarrow & \mathcal{M}^a, 10 \models \psi & \Leftrightarrow & \mathcal{M}, 11 \models \psi & & (i) \\
\mathcal{M}^{ab}, 10 \models \psi & \Leftrightarrow & \mathcal{M}^b, 10 \models \psi & \Leftrightarrow & \mathcal{M}^b, 11 \models \psi & & (ii) \\
\mathcal{M}, 00 \models \psi & \Leftrightarrow & \mathcal{M}^a, 00 \models \psi & \Leftrightarrow & \mathcal{M}, 01 \models \psi & & (iii)
\end{array}$$

Inductive case announcement:

- $\mathcal{M}, 10 \models [\chi]\psi$
  iff $\mathcal{M}, 10 \models \chi$ implies $\mathcal{M}|\chi, 10 \models \psi$ (*)
  iff $\mathcal{M}, 10 \models \chi$ implies $\begin{cases} \mathcal{M}, 10 \models \psi \text{ if } \mathcal{M}, 00 \models \chi \\ \mathcal{M}^b, 10 \models \psi \text{ otherwise} \end{cases}$
  iff $\mathcal{M}^a, 10 \models \chi$ implies $\begin{cases} \mathcal{M}^a, 10 \models \psi \text{ if } \mathcal{M}, 00 \models \chi \\ \mathcal{M}^{ab}, 10 \models \psi \text{ otherwise} \end{cases}$ (**)
  iff $\mathcal{M}^a, 10 \models \chi$ implies $\mathcal{M}^a|\chi, 10 \models \psi$
  iff $\mathcal{M}^a, 10 \models [\chi]\psi$.

  *: By induction hypothesis: $\mathcal{M}|\chi = \mathcal{M}$ if $\mathcal{M}, 00 \models \chi$, and $\mathcal{M}|\chi = \mathcal{M}^b$ otherwise.

  **: By induction hypothesis: $\mathcal{M}^a|\chi = \mathcal{M}^a$ if $\mathcal{M}^a, 00 \models \chi$, and $\mathcal{M}^a|\chi = \mathcal{M}^{ab}$ otherwise.

Inductive case group announcement (there are four different coalitions):

- $\mathcal{M}, 10 \models [\emptyset]\psi$
  iff $\mathcal{M}, 10 \models \psi$

- $\mathcal{M}, 10 \models [a]\psi$
  iff $\mathcal{M}, 10 \models \psi$ and $\mathcal{M}^a, 10 \models \psi$
  iff $\mathcal{M}^a, 10 \models \psi$ (by IH)
  iff $\mathcal{M}^a, 10 \models [a]\psi$

- $\mathcal{M}, 10 \models [b]\psi$
  iff $\mathcal{M}, 10 \models \psi$ and $\mathcal{M}^b, 10 \models \psi$
  iff $\mathcal{M}^a, 10 \models \psi$ and $\mathcal{M}^{ab}, 10 \models \psi$ (by IH)
  iff $\mathcal{M}^a, 10 \models [b]\psi$
- $\mathcal{M}, 10 \models [a, b]\psi$
  iff $\mathcal{M}, 10 \models \psi$ and $\mathcal{M}^a, 10 \models \psi$ and $\mathcal{M}^{ab}, 10 \models \psi$ and $\mathcal{M}^b, 10 \models \psi$
  iff $\mathcal{M}^a, 10 \models \psi$ and $\mathcal{M}^{ab}, 10 \models \psi$ (by IH)
  iff $\mathcal{M}^a, 10 \models [a, b]\psi$

□

Conjecture 5.28 *$\mathcal{L}_{apal}$ is not at least as expressive as $\mathcal{L}_{gal}$.*

Thus, we conjecture that the two logics are incomparable when it comes to expressivity.The following gives an idea for a possible proof, discussed together with Barteld P. Kooi, even if it does not succeed for now.

Sketch of proof : Let $\beta = \langle a \rangle K_b p$ be a $\mathcal{L}_{gal}$-formula. We already know, by Proposition 5.14, that it is equivalent to $\alpha = \langle a \rangle K_b p \wedge \langle b \rangle K_a p$. The idea of the proof is to show a class of pairs of models $\mathcal{M}_{q,r}, \mathcal{M}'_{q,r}$ ($q, r \in PROP$) such that

1. for all $q, r \in PROP$, $\mathcal{M}_{q,r}, pqr \models \neg\alpha$ and $\mathcal{M}'_{q,r}, pqr \models \alpha$
2. for all $\varphi \in \mathcal{L}_{apal}$, there exists $q, r \in PROP$ such $\mathcal{M}_{q,r}, pqr \models \varphi$ iff $\mathcal{M}'_{q,r}, pqr \models \varphi$

Then the formula α would distinguish any such pair of models, and no $\mathcal{L}_{apal}$-formula would be able to distinguish all of them, so $\mathcal{L}_{apal}$ would not be more expressive than $\mathcal{L}_{gal}$. Our proposal was the following:

Formally, $\mathcal{M}^*_{q,r} = \{S, \sim_a, \sim_b, V\}$ with $S = \{pqr \in \{0,1\}^3\}$ (and $S' = S \backslash \{000\}$). The valuation of p, q and r is defined by the name of the state, and is $\emptyset$ for any atom in $PROP \backslash \{p, q, r\}$. Now for all $s, t \in S^*$, $s \sim_a t$ iff ($\mathcal{M}^*_{q,r}, s \models q \longrightarrow p$ iff $\mathcal{M}^*_{q,r}, t \models q \longrightarrow p$) and $s \sim_b t$ iff ($\mathcal{M}^*_{q,r}, s \models r \longrightarrow p$ iff $\mathcal{M}^*_{q,r}, t \models r \longrightarrow p$).

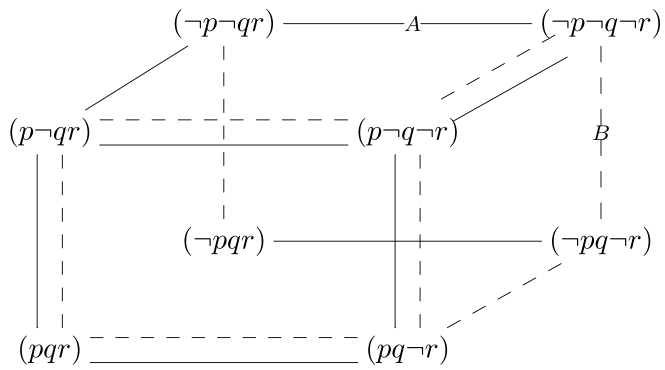
In particular, we have $\mathcal{M}^*_{q,r}, pqr \models K_a(q \longrightarrow p) \wedge K_b(r \longrightarrow p)$

Clearly, 1) is true. Indeed, $\mathcal{M}, pqr \models [a]\hat{K}_b(\neg p \wedge \neg q \wedge \neg r)$ and $\mathcal{M}', pqr \models \langle K_a(\neg q \vee p) \rangle K_b p$. Now to prove 2) it would be enough to prove that

$$(*) \quad \text{for all } \varphi \in \mathcal{L}_{apal}\{p\}, \text{ for all } q, r \in \Theta, \mathcal{M}_{q,r}, pqr \models \varphi \text{ iff } \mathcal{M}'_{q,r}, pqr \models \varphi.$$

Indeed, let φ be an $\mathcal{L}_{apal}$-formula, then let us call $PROP_\varphi = \{q_1, \dots q_n\}$ the atomic propositions appearing in φ, and let q, r be atomic propositions that are not in $PROP_\varphi$. We define $\varphi^* = \varphi(\forall q_i, \bot/q_i)$ and then consider $\mathcal{M}_{q,r}$ and $\mathcal{M}'_{q,r}$. We have that $\mathcal{M}_{q,r} \models \varphi \longleftrightarrow \varphi^*$ and $\mathcal{M}'_{q,r} \models \varphi \longleftrightarrow \varphi^*$ with $\varphi^* \in \mathcal{L}_{apal}\{p\}$. $(*)$ is then sufficient to prove 2).

$\mathcal{M}_{q,r} =$



and $\mathcal{M}'_{q,r} =$

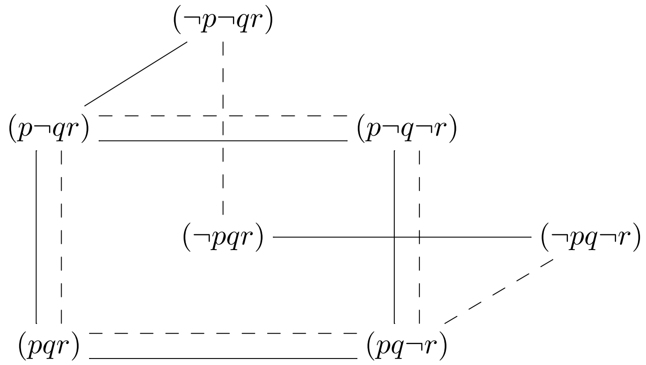


Figure 5.4: Trying to distinguish $\mathcal{L}_{gal}$ from $\mathcal{L}_{apal}$

Unfortunately, $(*)$ is false. The mistake came from the will to obtain a pair of models not able to distinguish $\chi = K_a p \vee K_b p \vee \Diamond(K_a p \wedge \neg K_b p) \wedge \Diamond(\neg K_a p \wedge K_b p)$ which seemed to be a reasonable translation for $\langle a \rangle K_b p$. Indeed $\langle a \rangle K_b p$ means that after some a's announcement b knows p. Indeed, if a knows p she can announce it, if b knows p then a can announce nothing, b will still know p. Now if none of the two agents know p but still $\langle a \rangle K_b p$ is true, then some a's announcement could teach p to b (without a learning anything about p). By Proposition 5.14 the converse would also be true. Well $\mathcal{M}_{q,r}$ and $\mathcal{M}'_{q,r}$ cannot actually distinguish χ (and in fact probably none of the $\mathcal{L}_{apal}$-formula of K-degree 1), but it can distinguish some $\mathcal{L}_{apal}$-formulas. For example the following one, which is true in $\mathcal{M}, pqr$ and not in $\mathcal{M}', pqr$: $\hat{K}_a \Diamond(\hat{K}_a p \wedge \Box(K_a \neg p \longleftrightarrow K_b \neg p))$

However, in spite of the mistake, we think that this may be a good starting point to find bigger classes of pairs of models, able to distinguish every $\mathcal{L}_{apal}$-formula from $\langle a \rangle K_b p$. We would thus get an infinite set $\{\mathcal{C}_i\}_{i \in \mathbb{N}}$ of such classes of pairs, such that each class $\mathcal{C}_i$ is able to distinguish $\langle a \rangle K_b p$ from any $\mathcal{L}_{apal}$-formula of K-degree i.

$\Diamond$

Now consider a very special model class $\mathfrak{M}_g$, namely the class where an agent g has the identity relation on all models (there may be other agents). It is clear that the announcement made by g has the property that $K_g \varphi \longleftrightarrow \varphi$: everything true is known by g. Therefore $\Diamond\varphi$

in $APAL$ is equivalent to $\langle g \rangle \varphi$ in GAL (ignoring a further translation downward in φ). If we restrict the model class of the logic to $\mathfrak{M}_g$, we say that *a super agent g exists*. This makes clear that:

Proposition 5.29 *If a super agent g exists, GAL is at least as expressive as APAL.*

Proof Given a $\varphi \in \mathcal{L}_{apal}$, replace every occurrence of $\Box$ in φ by $[g]$. The resulting formula is in $\mathcal{L}_{gal}$ and it is equivalent to the initial one in every epistemic model.

$\Box$

5.3 Model Checking

If a given system can be modeled as a finite model, one would like to verify if a given property written in a language for specifying desired properties of systems holds in the finite model. We speak of the model checking problem, an area of automated deduction that has been addressed for almost all logical languages, for example modal logic in [Gradel and Otto, 1999], temporal logic in [Clarke et al., 1999], etc. There is a need, on the theoretical side, to provide a sound mathematical basis for the design of algorithms devoted to the model checking problem. Hence, the question arises whether the following decision problem is decidable:

input: a finite structure $\mathcal{M} = (S, \sim_1, \ldots, \sim_n, V)$, a state $x \in S$ and a formula $\varphi \in \mathcal{L}_{gal}$,

output: determine whether φ is satisfied at x in $\mathcal{M}$.

This decision problem, denoted **(MC(GAL))** is a variant of the well-known model checking problem. If one restricts to formulas $\varphi \in \mathcal{L}_{el}$, then the above decision problem is known to be P-complete. The notion of a formula like $[\{1, \ldots, n\}]\varphi$ being satisfied in a structure $\mathcal{M} = (S, \sim_1, \ldots, \sim_n, V)$ at state $x \in S$ relies on the satisfiability of all (infinitely many) formulas like $[K_1\varphi_1 \wedge \ldots \wedge K_n\varphi_n]\varphi$ at x where $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_{el}$. In Theorem 5.31 we show that **(MC(GAL))** is in PSPACE and in Theorem 5.32 we show that it is PSPACE-hard.

5.3.1 Preliminary Results

Let $Z_{\mathcal{M}}$ be the greatest bisimulation relation on $\mathcal{M}$. Note that $Z_{\mathcal{M}}$ is an equivalence relation on S. For all $s \in S$, let $\|s\|$ be the equivalence class of s modulo $Z_{\mathcal{M}}$. The bisimulation contraction of $\mathcal{M}$ is the structure $\|\mathcal{M}\| = (S', \sim'_1, \ldots, \sim'_n, V')$ such that:

- $S' = S|_{Z_{\mathcal{M}}}$, i.e. the quotient of S modulo $Z_{\mathcal{M}}$
- $\|s\| \sim'_i \|t\|$ iff there exist $v, w \in S$ such that $sZ_{\mathcal{M}}v$, $tZ_{\mathcal{M}}w$ and $v \sim_i w$
- $V'(p) = V(p)|_{Z_{\mathcal{M}}}$

The following proposition will be obvious, because:

- the bisimulation contraction is bisimilar to the original structure;
- bisimilar structures have the same logical theory[Blackburn et al., 2001];
- public announcement and group announcement are bisimilation preserving operations.

Proposition 5.30 *For all* $\varphi \in \mathcal{L}_{gal}, \|\mathcal{M}\|, \|x\| \models \varphi$ *iff* $\mathcal{M}, x \models \varphi$*.*

In $\|\mathcal{M}\| = (S', \sim'_1, \ldots, \sim'_n, V')$, every $\|s\| \in S'$ can be distinguished by a pure epistemic formula from all other (non-bisimilar) states. Let us call $\psi_{[\![s]\!]}$ the characteristic formula of $[\![s]\!]$ in $\mathcal{M}'$. Thus, for any $i \in \{1, \ldots, n\}$, $\hat{K}_i\psi_{[\![s]\!]}$ characterizes the class of equivalence for $\sim'_i$. Hence, for any $i \in \{1, \ldots, n\}$, each union $\mathcal{C}'_i$ of classes of equivalence for $\sim'_i$ is distinguished from all other (non-bisimilar) states by a pure epistemic formula of the form $K_i\varphi_i$. Therefore, a pure epistemic formula of the form $\bigwedge_{i \in G} K_i\varphi_i$ defines a restriction $\mathcal{M}'' = (S'', \sim''_1, \ldots, \sim''_n, V'')$ where $S'' = \cap_{i \in G}\mathcal{C}'_i$. We call such a restriction a *definable* restriction.

5.3.2 Model Checking Algorithm

Proposition 5.31 (MC(GAL)) *is in* $PSPACE$*.*

Proof Since $APTIME = PSPACE$ (see [Chandra et al., 1981]), it suffices to prove that **(MC (GAL))** is in $APTIME$. Let us consider the alternating algorithm 1 given on page 115. This algorithm takes as input a finite model $\mathcal{M}$, a state s in $\mathcal{M}$, a formula φ in $\mathcal{L}_{gal}$ and b in $\{0, 1\}$. It stops with a reject iff either $b = 0$ and $\mathcal{M}, s \models \varphi$ or $b = 1$ and $\mathcal{M}, s \not\models \varphi$ whereas it stops with an accept iff either $b = 0$ and $\mathcal{M}, s \not\models \varphi$ or $b = 1$ and $\mathcal{M}, s \models \varphi$. Its execution depends primarily on (φ, b). Each case is either existential or universal. For example, the case $(\varphi_1 \vee \varphi_2, 1)$ is existential. It is an accepting case iff for some $\varphi' \in \{\varphi_1, \varphi_2\}$, the case $(\varphi', 1)$ is accepting, thus corresponding to the fact that $\varphi_1 \vee \varphi_2$ is true at s in $\mathcal{M}$ iff for some $\varphi' \in \{\varphi_1, \varphi_2\}$, φ' is true at s in $\mathcal{M}$. As well, the case $(\varphi_1 \vee \varphi_2, 0)$ is universal. It is an accepting case iff for every $\varphi' \in \{\varphi_1, \varphi_2\}$, the case $(\varphi', 0)$ is accepting, thus corresponding to the fact that $\varphi_1 \vee \varphi_2$ is false at s in $\mathcal{M}$ iff for every $\varphi' \in \{\varphi_1, \varphi_2\}$, φ' is false at s in $\mathcal{M}$. Cases labelled with $(\cdot)$ are both existential and universal.

Obviously,

- $sat(\mathcal{M}, s, \varphi, 1)$ accepts iff $\mathcal{M}, s \models \varphi$,
- $sat(\mathcal{M}, s, \varphi, 1)$ rejects iff $\mathcal{M}, s \not\models \varphi$,
- $sat(\mathcal{M}, s, \varphi, 0)$ accepts iff $\mathcal{M}, s \not\models \varphi$,
- $sat(\mathcal{M}, s, \varphi, 0)$ rejects iff $\mathcal{M}, s \models \varphi$.

Algorithm 1 $sat(\mathcal{M}, s, \varphi, b)$

case (φ, b) of
$(\cdot)$ $(p, 1)$: **if** $s \in V(p)$ **then** accept **else** reject;
$(\cdot)$ $(p, 0)$: **if** $s \in V(p)$ **then** reject **else** accept;
$(\cdot)$ $(\bot, 1)$: reject;
$(\cdot)$ $(\bot, 0)$: accept;
$(\cdot)$ $(\neg\varphi', 1)$: $sat(\mathcal{M}, s, \varphi', 0)$;
$(\cdot)$ $(\neg\varphi', 0)$: $sat(\mathcal{M}, s, \varphi', 1)$;
$(\exists)$ $(\varphi_1 \vee \varphi_2, 1)$: choose $\varphi' \in \{\varphi_1, \varphi_2\}$; $sat(\mathcal{M}, s, \varphi', 1)$;
$(\forall)$ $(\varphi_1 \vee \varphi_2, 0)$: choose $\varphi' \in \{\varphi_1, \varphi_2\}$; $sat(\mathcal{M}, s, \varphi', 0)$;
$(\forall)$ $(K_i\varphi', 1)$: choose $t \in \sim_i (s)$; $sat(\mathcal{M}, t, \varphi', 1)$;
$(\exists)$ $(K_i\varphi', 0)$: choose $t \in \sim_i (s)$; $sat(\mathcal{M}, t, \varphi', 0)$;
$(\cdot)$ $([\varphi_1]\varphi_2, 1)$: compute the φ_1-definable restriction $\mathcal{M}' = (S', \sim_1, \ldots, \sim_n, V')$ of $\mathcal{M}$;
if $s \in S'$ **then** $sat(\mathcal{M}', s, \varphi_2, 1)$ **else** accept;
$(\cdot)$ $([\varphi_1]\varphi_2, 0)$: compute the φ_1-definable restriction $\mathcal{M}' = (S', \sim_1, \ldots, \sim_n, V')$ of $\mathcal{M}$;
if $s \in S'$ **then** $sat(\mathcal{M}', s, \varphi_2, 0)$ **else** reject;
$(\forall)$ $([G]\varphi, 1)$: Compute $\|\mathcal{M}\|$, choose a definable restriction $\mathcal{M}'' = (S'', \sim''_1, \ldots, \sim''_n, V'')$ of $\|\mathcal{M}\|$ s.t. $S'' = \cap_{i \in G}\mathcal{C}_i$ where $\mathcal{C}_i$ are unions of classes of equivalence for $\sim'_i$;
if $s \in S''$ **then** $sat(\mathcal{M}'', s, \varphi, 1)$ **else** accept;
$(\exists)$ $([G]\varphi, 0)$: Compute $\|\mathcal{M}\|$, choose a definable restriction $\mathcal{M}'' = (S'', \sim''_1, \ldots, \sim''_n, V'')$ of $\|\mathcal{M}\|$ s.t. $S'' = \cap_{i \in G}\mathcal{C}_i$ where $\mathcal{C}_i$ are unions of classes of equivalence for $\sim'_i$;
if $s \in S''$ **then** $sat(\mathcal{M}'', s, \varphi, 0)$ **else** reject;
end case

The only difficult case is $([G]\varphi, 1)$. Computing $\|\mathcal{M}\|$ is easy and by Proposition 5.30 we have that $\mathcal{M}, s \models \langle G\rangle\varphi$ iff $\|\mathcal{M}\|, \|s\| \models \langle G\rangle\varphi$. Then we just have to prove it in the case where $\|\mathcal{M}\| = \mathcal{M}$. Let us suppose it, and let us see that, if there is a definable restriction $\mathcal{M}'' = (S'', \sim''_1, \ldots, \sim''_n, V')$ of $\mathcal{M}$ such that $S'' = \cap_{i \in G}\mathcal{C}_i$ where $\mathcal{C}_i$ are unions of classes of equivalence for $\sim_i$, if also $s \in S''$ and $\mathcal{M}'', s \models \varphi$, then $\mathcal{M}, s \models \langle G\rangle\varphi$. Let us then suppose the first part of the implication.
$\mathcal{M}$ is supposed to be bisimulation-contracted, then we know that for all $s \in \mathcal{M}$, there is $\varphi_s \in \mathcal{L}_{gal}$, s.t. for all $t \in \mathcal{M}$, $\mathcal{M}, t \models \varphi_s$ iff $s = t$. It implies that $s \in S''$ iff (for all $i \in G$, $s \in \mathcal{C}_i$) iff $\mathcal{M}, s \models \bigwedge_{i \in G}(\bigvee_{t \in \mathcal{C}_i} \varphi_t)$ which is equivalent to $\mathcal{M}, s \models \bigwedge_{i \in G} K_i(\bigvee_{t \in \mathcal{C}_i} \varphi_t)$. That means that $\mathcal{M}'' = \mathcal{M}| \bigwedge_{i \in G} K_i(\bigvee_{t \in \mathcal{C}_i} \varphi_t)$ and then $\mathcal{M}, s \models \langle \bigwedge_{i \in G} K_i(\bigvee_{t \in \mathcal{C}_i} \varphi_t)\rangle\varphi$ (because $s \in S''$ and $\mathcal{M}'', s \models \varphi$). We obtain $\mathcal{M}, s \models \langle G\rangle\varphi$.

Since *sat* can be implemented in polynomial time, **(MC (GAL))** is in $APTIME$.

□

Proposition 5.32 **(MC(GAL))** *is* $PSPACE$*-hard.*

Proof We prove that **(MC(GAL))** is $PSPACE$-hard. Let $\Psi = Q_1 x_1 ... Q_k x_k \Phi(x_1, ..., x_k)$ be an entry of the problem QBF-SAT:

- $Q_1, ..., Q_k \in \{\forall, \exists\}$
- $x_1, ..., x_k$ are Boolean variables
- $\Phi(x_1, ..., x_k)$ is a Boolean formula

We associate to Ψ a model $\mathcal{M}_{1;k} = (W_{1;k}, \mathcal{R}_1, \dots, \mathcal{R}_k, V)$, a world $x \in W_{1;k}$ and a formula $\psi(\Psi) \in \mathcal{L}_{gal}$ such that the following property (P_k) is true:

$$\models \Psi \text{ iff } \mathcal{M}_{1;k}, x \models \psi(\Psi)$$

Let $1 \leq m \leq k$, $W_{m;k} = \{x\} \cup \{(x_l, 0), (x_l, 1)\}_{l \in \{m, \dots, k\}}$ be the set of possible worlds, $\{p_m^-, p_m^+, ..., p_k^-, p_k^+\}$ be the set of atoms, with $V(p_l^-)$=$\{(x_l, 0)\}$ and $V(p_l^+)$=$\{(x_l, 1)\}$. Let $i, g \in AG$ and let us define $\begin{cases} \mathcal{R}_i = W_{m;k} \times W_{m;k} \\ \mathcal{R}_g = \{(s, s) \text{ such that } s \in W_{m;k}\} \end{cases}$ (We remark that g is omniscient and that i assumes this fact)

$\mathcal{M}_{1;k}$:

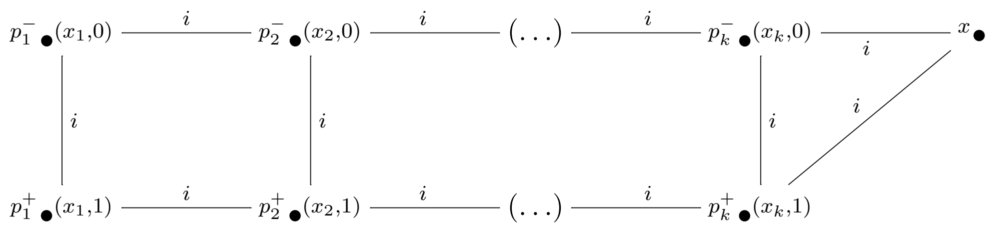


Figure 5.5: A model to prove the hardness of $MC(GAL)$
$\mathcal{R}_i$ is assumed to be reflexive, symmetrical and transitive, and $\mathcal{R}_g$ reflexive

We now define some formulas:

for all $l \in \{1, \dots, k\}, q_l = \hat{K}_i(p_l^- \wedge K_i \neg p_l^+) \vee \hat{K}_i(p_l^+ \wedge K_i \neg p_l^-)$ and $r_l = \hat{K}_i p_l^+ \wedge \hat{K}_i p_l^-$.

Intuitively, $\mathcal{M}_{1,k} \models r_l$ means that $(x_l, 0)$ and $(x_l, 1)$ are still possible worlds of the model (i.e. the truth value of x_l is not fixed) and $\mathcal{M}_{1,k} \models q_l$ means that one and only one of $(x_l, 0)$ and $(x_l, 1)$ is still a possible world (i.e. we have fixed the value of x_l).
We can now define the equivalence recursively:

let $\psi_0 = \Phi(\hat{K}_i p_1^+, \dots, \hat{K}_i p_k^+)$, suppose ψ_l is defined for some $l < k$, then

$$\psi_{l+1} = \begin{cases} K_i[g](q_1 \wedge ... \wedge q_{k-l} \wedge r_{k-l+1} \wedge ... \wedge r_k \rightarrow \psi_l) \text{ if } Q_{l+1} = \forall \\ \hat{K}_i\langle g\rangle(q_1 \wedge ... \wedge q_{k-l} \wedge r_{k-l+1} \wedge ... \wedge r_k \wedge \psi_l) \text{ if } Q_{l+1} = \exists \end{cases}$$

Finally, $\psi(\Psi) = \psi_k$.

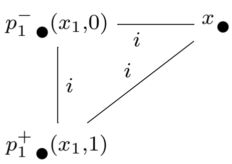
Example: If $\Psi = \forall x_1 \exists x_2 \forall x_3 \Phi(x_1, x_2, x_3)$ then:

$$\psi(\Psi) = K_i[g](q_1 \wedge r_2 \wedge r_3 \rightarrow \hat{K}_i\langle g\rangle(q_1 \wedge q_2 \wedge r_3 \wedge K_i[g](q_1 \wedge q_2 \wedge q_3 \rightarrow \Phi(\hat{K}_i p_1^+, \dots, \hat{K}_i p_k^+))))$$

Intuitively, $K_i[g](q_1 \wedge r_2 \wedge r_3 \rightarrow \varphi)$ means 'After having fixed the value of x_1 only, φ' and $\hat{K}_i\langle g\rangle(q_1 \wedge r_2 \wedge r_3 \wedge \varphi)$ as 'There is a way of fixing the value of x_1 only, such that φ'. We can now prove $\models \Psi \Leftrightarrow \mathcal{M}_{1;k}, x \models \psi(\Psi)$ by induction on k. The induction is quite technical, but the intuition is that something is true after having fixed the value of $k+1$ boolean variables if and only if it is true after having fixed the value of the first k variables, added the final one and then fixed its value. More precisely:

Base case: $k = 1$:

$\Psi = Q_1 x_1 \Phi(x_1)$, and $\mathcal{M}_1$:



- If $Q_1 = \forall$ then $\models \Psi$ iff ($\models \Phi(\top)$ and $\models \Phi(\bot)$)
  iff ($p_1^+ \bullet \overset{i}{\text{———}} \bullet x \models \Phi(\hat{K}_i p_1^+)$ and $p_1^- \bullet \overset{i}{\text{———}} \bullet x \models \Phi(\hat{K}_i p_1^+)$)
  iff $\mathcal{M}_1, x \models K_i[g](q_1 \rightarrow \Phi(\hat{K}_i p_1^+))$ i.e. $\mathcal{M}_1, x \models \psi(\Psi)$

- Else, $Q_1 = \exists$ and $\models \Psi$ iff ($\models \Phi(\top)$ or $\models \Phi(\bot)$)
  iff ($p_1^+ \bullet \overset{i}{\text{———}} \bullet x \models \Phi(\hat{K}_i p_1^+)$ or $p_1^- \bullet \overset{i}{\text{———}} \bullet x \models \Phi(\hat{K}_i p_1^+)$)
  iff $\mathcal{M}_1, x \models \hat{K}_i\langle g\rangle(q_1 \wedge \Phi(\hat{K}_i p_1^+))$ i.e. $\mathcal{M}_1, x \models \psi(\Psi)$

Inductive case: $k \rightarrow k+1$:

Suppose that (P_k) is true, and let us note: $\Psi = Q_1 x_1 ... Q_k x_k Q_{k+1} x_{k+1} \Phi(x_1, ..., x_k, x_{k+1})$. We pose $\tilde{\Psi}(x_1) := Q_2 x_2 .. Q_k x_k Q_{k+1} x_{k+1} \Phi(x_1, ..., x_k, x_{k+1})$ and we have $\models \Psi \Leftrightarrow \models Q_1 x_1 \tilde{\Psi}(x_1)$. Thus:

- If $Q_1 = \forall$ then $\models \Psi$ iff ($\models \tilde{\Psi}(\top)$ and $\models \tilde{\Psi}(\bot)$)

  iff $\mathcal{M}_{2;k+1}, x \models \psi(\tilde{\Psi}(\top))$ and $\mathcal{M}_{2;k+1}, x \models \psi(\tilde{\Psi}(\bot))$ (by IH)

  iff $\mathcal{M}_{1;k+1}, x \models K_i[g](q_1 \wedge r_2 \wedge \ldots \wedge r_{k+1} \rightarrow \psi^*(\tilde{\Psi}(\hat{K}_i p_1^+)))$

  with ψ^* obtained by replacing any succession $q_2 \wedge \ldots$ by $q_1 \wedge q_2 \wedge \ldots$

- If $Q_1 = \exists$ then $\models \Psi$ iff ($\models \tilde{\Psi}(\top)$ or $\models \tilde{\Psi}(\bot)$)

  iff $\mathcal{M}_{2;k+1}, x \models \psi(\tilde{\Psi}(\top))$ or $\mathcal{M}_{2;k+1}, x \models \psi(\tilde{\Psi}(\bot))$ (by IH)

  iff $\mathcal{M}_{1;k+1}, x \models \hat{K}_i\langle g\rangle(q_1 \wedge r_2 \wedge \ldots \wedge r_{k+1} \wedge \psi^*(\tilde{\Psi}(\hat{K}_i p_1^+)))$

$\square$

We conclude:

Theorem 5.33 (MC(GAL)) *is PSPACE-complete.*

We observe that our results also extend to *APAL*: the model checking problem for arbitrary public announcement logic is also PSPACE-complete. A relevant detail in the proof of

Proposition 5.32 is that it involves an omniscient agent g, and that the role of $[g]$ is in *APAL* played by $\Box$, and that of $\langle g\rangle$ by $\Diamond$. (See also the expressivity result involving g, Proposition 5.29.)

5.4 Announcements and Ability

Our initial intuitive interpretation of a formula of the form $\langle C\rangle\varphi$ was that coalition C has the ability to make φ come about by making some public announcement. We now have a better understanding of group announcement logic; let us discuss to what extent that intuition is precise.

Recent work on strategy logics have illuminated the fact that there are many subtly different notions of ability in the context of incomplete information (see [Jamroga, 2003, Jamroga and van der Hoek, 2004, Ågotnes, 2006] or [Jamroga and Ågotnes, 2007] for a recent summary). For example, does ability entail knowledge of ability? In [Jamroga and Ågotnes, 2007, p. 433] three levels of ability in general strategy logics are discussed. We now discuss counterparts of these in the special context of truthful public announcements. In general strategy logics, such as ATL or STIT, agents and coalitions can perform arbitrary state-transforming actions. In our setting the actions are truthful *announcements*, and there is thus an intimate relationship between knowledge and ability. There are two main questions of interest related to the mentioned different variants of ability here: are they indeed different in this special context, and are they expressible in the logical language of *GAL*?

5.4.1 Singleton Coalitions

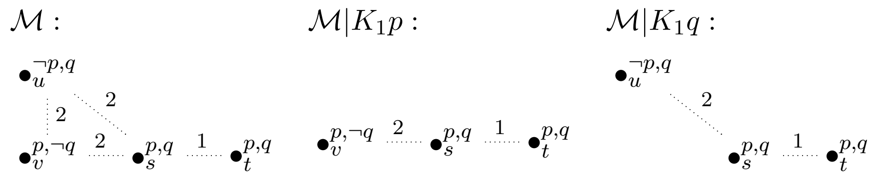
For simplicity we first consider a singleton coalition $\{a\}$. What does it mean that agent a has the ability to make a goal φ come about by making a public announcement? Let us begin with the weakest form of ability.

Being able to, but not necessarily knowing it The formula $\langle a\rangle\varphi$ means that there is something which a knows, and if the fact that a knows it is announced, φ is a consequence. However, it might be the case that *a doesn't know this*, i.e., that $K_a\langle a\rangle\varphi$ is not true. As an example, first observe that $\langle K_a\psi\rangle\varphi \rightarrow K_a\langle K_a\psi\rangle\varphi$ is not a principle of public announcement logic. As a counter-example take state s of the following model

$$\bullet^{p}_{s} \overset{a}{\cdots\cdots} \bullet^{\neg p}_{t}$$

and take $\psi_a = \top$ and $\varphi = p$. However, this does not mean that a cannot achieve φ in all her accessible states by some *other* announcements (possibly different ones in different states). But in group announcement logic, we have in the model above that $s \models \langle a\rangle p$ (a can announce $K_a\top$), but $t \not\models \langle a\rangle p$ and thus, $s \models \neg K_a\langle a\rangle p$. So, $\langle a\rangle\varphi \rightarrow K_a\langle a\rangle\varphi$ is not a principle of group announcement logic. This is a first illustration of the fact that we must be careful when using the term "ability": in some (but not necessarily all) circumstances it might be counter-intuitive to say that a has the ability to make φ come about, when she is not aware that she is; when she cannot discern between the actual situation and a situation in which she does not have this ability.

Being able to, knowing that, but not knowing how Consider the following model $\mathcal{M}$ (some model updates are also shown):



and let

$$\varphi = K_2q \wedge (\neg K_2p \vee \hat{K}_1(K_2p \wedge \neg K_2q))$$

If we take the current state to be s, we have a situation where 1 is able to make φ come about and where she in addition knows this; a stronger type of ability than in the example above. Formally: $s \models \langle 1 \rangle \varphi$, because $s \models \langle K_1q \rangle \varphi$, and $t \models \langle 1 \rangle \varphi$ because $t \models \langle K_1p \rangle \varphi$. Thus, $s \models K_1\langle 1 \rangle \varphi$. However, we argue, it might still be counter-intuitive to say that 1 can make φ come about in this situation. The reason is that she has to use *different announcements in indiscernible states.* Observe that $s \models \langle K_1p \rangle \neg \varphi$ and $t \models \langle K_1q \rangle \neg \varphi$: while the same announcements can be made in both states, they don't have the same consequences. In fact, there exists *no* single announcement agent 1 can make which will ensure that φ will be true in both s and t. To see this, we can enumerate the possible models resulting from 1 making an announcement in s or t. Because such a model must include 1's equivalence class $\{s, t\}$, there are four possibilities. First, the starting model itself (e.g., 1 announces a tautology), in which φ does not hold in s. Second, the model where only state u is removed (e.g., 1 announces K_1p), in which φ does not hold in s (as we saw above). Third, the model where only state v is removed (e.g., 1 announces K_1q), in which φ does not hold in t (as we saw above). Fourth, the model where both u and v are removed, in which φ holds in neither s nor t.

Since agent 1 cannot discern state s from state t, she has the ability to make φ come about only in the sense that she depends on guessing the correct announcement. In other words, she can make φ come about, knows that she can make φ come about, but does not know *how* to make φ come about.

Being able to, knowing that, knowing how Thus, we can formulate a strong notion of the ability of a to achieve φ by public announcements: there exists a formula ψ such that a

knows ψ and in any state a considers possible, $\langle K_a\psi\rangle\varphi$ holds.

Compare this version of ability, "there is an announcement which a knows will achieve the goal", with the previous version above, "a knows that there is an announcement which will achieve the goal". We can call these notions, respectively, *knowing de re* and *knowing de dicto* that the goal can be achieved, following [Jamroga and van der Hoek, 2004] who use the same terminology for general strategy logics, after the corresponding notion used in quantified modal logic. In our framework these notions are more formally defined as follows:

Knowledge de dicto: Agent i knows *de dicto* that she can achieve the goal φ in state s of model $\mathcal{M}$ iff

$$\forall t \sim_i s \; \exists\psi \in \mathcal{L}_{el} \; (\mathcal{M}, t) \models \langle K_i\psi\rangle\varphi \tag{5.1}$$

Knowledge de re: Agent i knows *de re* that she can achieve the goal φ in state s of model $\mathcal{M}$ iff

$$\exists\psi \in \mathcal{L}_{el} \; \forall t \sim_i s \; (\mathcal{M}, t) \models \langle K_i\psi\rangle\varphi \tag{5.2}$$

Note, however, that it is not prima facie clear that there is a distinction between these notions in *GAL*, because of the intimate interaction between knowledge and possible actions (announcements), but the model and formula above show that there indeed is.

We have seen how to express knowledge *de dicto*. In the most popular general strategy logics such as ATL, where actions are not necessarily truthful announcements, extended with epistemics, knowledge *de re* is not expressible. Several recent works have focussed on extending such logics in order to be able to express knowledge *de re* and other interaction properties between knowledge and ability : see [Jamroga and van der Hoek, 2004, Ågotnes, 2006, Jamroga and Ågotnes, 2007, Broersen, 2008]. In the special case of *GAL*, however, it turns out that knowledge *de re* in fact *is* already expressible (in the single agent case, at least), as the following proposition shows.

Proposition 5.34

1. *Knowledge* de dicto *(5.1) is expressed by the formula* $K_i\langle i\rangle\varphi$
2. *Knowledge* de re *(5.2) is expressed by the formula* $\langle i\rangle K_i\varphi$

Proof

1. Immediate.
2. Let $\mathcal{M}$ be a model and s a state in $\mathcal{M}$.

Agent i knows *de re* that she can achieve φ

iff $\exists\psi \in \mathcal{L}_{el}, (\mathcal{M}, s \models K_i\psi$ and $\forall t \in S$ (if $s \sim_i t$ then $\mathcal{M}, t \models \langle K_i\psi\rangle\varphi$))

iff $\exists\psi \in \mathcal{L}_{el}, (\mathcal{M}, s \models K_i\psi$ and $\forall t \in S$ (if $s \sim_i t$ then $\mathcal{M}|K_i\psi, t \models \varphi$))

(since $\mathcal{M}, s \models K_i\psi$ and $s \sim_i t$ implies that $\mathcal{M}, t \models K_i\psi$)

iff $\exists\psi \in \mathcal{L}_{el}, (\mathcal{M}, s \models K_i\psi$ and $\forall t \in [\![K_i\psi]\!]$, $\mathcal{M}|K_i\psi, t \models \varphi$)

iff $\exists\psi \in \mathcal{L}_{el}, (\mathcal{M}, s \models K_i\psi$ and $\forall t \in [\![K_i\psi]\!]$, (if $s \sim_i t$ then $\mathcal{M}|K_i\psi, t \models \varphi$))

iff $\exists\psi \in \mathcal{L}_{el}((\mathcal{M}, s) \models K_i\psi$ and $\mathcal{M}|K_i\psi, s \models K_i\varphi$)

iff $(\mathcal{M}, s) \models \langle i\rangle K_i\varphi$.

□

Thus, i knows *de re* that she can achieve φ iff she can achieve the fact that she knows φ. This depends crucially on the fact that by "achieve" we mean achieve by truthful public announcements; it is not true if we allow general actions. As an illustration of the latter case, take the following example. An agent i is in front of a combination lock safe. The agent does not know the combination. The available actions correspond to dialling different codes. The agent is *able* to open the safe, $\langle i\rangle open$, because there is a successful action (dial the correct code). She knows *de dicto* that she can open the safe, $K_i\langle i\rangle open$, because this is true in all the states she considers possible (a possible state correspond to a possible correct code). But she does not know *de re* that she can open the safe, because there is no code that will open the safe in all the states she considers possible. However, $\langle i\rangle K_i open$ *does* hold: there is some action she can perform (dial the correct code) after which she will know that the safe is open. In *GAL*, the fact that $\langle i\rangle K_i\varphi$ expresses (5.2) is a result of the inter-dependence between knowledge and actions (announcements) and the S5 properties of knowledge. The following are some properties of knowledge *de dicto* and *de re* in *GAL*.

Proposition 5.35 *The following are valid.*

1. $K_i\langle i\rangle\varphi \to \langle i\rangle\varphi$. *Knowledge* de dicto *of ability implies ability; if you* know *that you can do it then you can do it.*

2. $\langle i\rangle K_i\varphi \to K_i\langle i\rangle\varphi$. *Knowledge* de re *implies knowledge* de dicto*; if you know* how *to do it you know* that *you can do it.*

3. $\langle i\rangle K_i\varphi \leftrightarrow K_i\langle i\rangle K_i\varphi$. *Knowledge* de re *holds iff knowledge of knowledge* de re *holds; you know how to do it iff you know that.*

Proof The first point is immediate from reflexivity of the accessibility relations. The second point is also immediate; let ψ be fixed by (5.2). For the third point, the direction to the left is immedate by point 1, so consider the direction to the right. Assume that $\mathcal{M}, s \models \langle i\rangle K_i\varphi$, i.e., that (5.2) holds. Let $u \sim_i s$. We must show that $\exists\psi \in \mathcal{L}_{el}\ \forall t \sim_i u\ (\mathcal{M}, t) \models \langle K_i\psi\rangle\varphi$. Let ψ be as in (5.2), and let $t \sim_i u$. By transitivity of $\sim_i$ we have that $t \sim_i s$, and thus that $(\mathcal{M}, t) \models \langle K_i\psi\rangle\varphi$ by (5.2).

□

On first sight the expression $\langle i\rangle K_i\varphi$ of knowledge *de re* might seem to suffer from a similar problem as the expression of "mere" ability of the first type we discussed above, $\langle i\rangle\psi$, namely that while i has the ability to make ψ come about she does not necessarily know this (*de dicto*). However, as the last point in the proposition above shows, if ψ is of the special form $K_i\varphi$ (for the *same* agent i), then ability does in fact imply knowledge of ability. In every circumstance where you can achieve a state where you know φ, you know that you can.

As illustrated above, the other direction of the second property in Prop. 5.35 does not hold; knowledge *de dicto* does not imply knowledge *de re*. Given our expressions of these two properties, we thus have that

$$K_i\langle i\rangle\varphi \rightarrow \langle i\rangle K_i\varphi \text{ is not valid}$$

– that you know that you can achieve φ does not necessarily mean that you can achieve a state where you know φ.

5.4.2 More Than One Agent

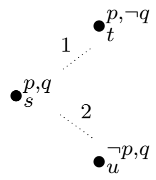
In the case of more than one agent, there are even more subtleties. In particular, what does it mean that a group knows how to achieve something, i.e., knows which joint announcement will be effective? That everybody knows it? That they have common knowledge of it?

In [Jamroga and van der Hoek, 2004] it is argued that the answer depends on the situation. It might be the case that the agents have common knowledge (although they then need some resolution mechanism for cases when there are more than one effective announcement, in order to coordinate); that every agent knows the effective announcement; that the agents have distributed knowledge about the effective announcement and thus can pool their knowledge together to find out what they should do; that a particular agent (the "leader") knows the effective announcement and can communicate it to the others.

In *GAL* we do not have distributed or common knowledge in the language, but "everybody knows" can be defined: $E_G\varphi \equiv \bigwedge_{i\in G} K_i\varphi$, where G is a coalition. The following generalisation of (5.2) says that in state s coalition G can make a truthful announcement which all the members of G know will achieve the goal φ:

$$\exists\{\psi_i\}_{i\in G} \subseteq \mathcal{L}_{el}\ \forall(t,s) \in \bigcup_{i\in G} \sim_i \ \ (\mathcal{M},t) \models \langle \bigwedge_{i\in G} K_i\psi_i\rangle\varphi \tag{5.3}$$

However, while the single agent case (5.2) is expressed by $\langle i\rangle K_i\varphi$, it is *not* in general the case that (5.3) is expressed by $\langle G\rangle E_G\varphi$. The following is a counter-example. Let $\mathcal{M}$ and φ be the following model and formula.



$$\varphi = p \wedge q$$

Let $G = \{1, 2\}$. It is easy to see that group G in s does not know *de re* that they can achieve φ in the sense of (5.3): it would imply, for instance, that it is possible to make an announcement in state t which at the same time eliminates state t – which is impossible. However, $\langle 1, 2\rangle E_G\varphi$ holds in s – $\{1, 2\}$ can announce $K_1p \wedge K_2q$.

Let us consider distributed and common knowledge. Assume for a moment that the language is extended with operators C_G and D_G where G is a coalition, such that $\mathcal{M}, s \models D_G\varphi$ iff for all $(s, t) \in \bigcap_{i\in G} \sim_i \mathcal{M}, t \models \varphi$ and $M, s \models C_G\varphi$ iff for all $(s, t) \in {\sim_G}^* \mathcal{M}, t \models \varphi$, where ${\sim_G}^*$ is the reflexive transitive closure of $\bigcup_{i\in G} \sim_i$. The following version of (5.3) says that in s, G can make a truthful announcement which G distributively know will achieve the goal φ:

$$\exists\{\psi_i\}_{i\in G} \subseteq \mathcal{L}_{el}\ \forall t \in S\ \left((s, t) \in \bigcap_{i\in G} \sim_i \Rightarrow (\mathcal{M}, t) \models \langle \bigwedge_{i\in G} K_i\psi_i\rangle\varphi\right) \tag{5.4}$$

Contrary to the case for "everybody knows", this property *is* in fact expressed by the analogue to the expression for the single-agent case. This can be shown similarly to Prop. 5.35 – observe that $(s, t) \in \bigcap_{i\in G}$ and $\mathcal{M}, s \models \bigwedge_{i\in G} K_i\psi_i$ implies that $\mathcal{M}, t \models \bigwedge_{i\in G} K_i\psi_i$:

Proposition 5.36 *The property (5.4) is expressed by the formula $\langle G\rangle D_G\varphi$.*

The situation for common knowledge is, however, similar to that of "everybody knows". The following version of (5.4) says that in s G can make a truthful announcement which G commonly know will achieve the goal φ:

$$\exists\{\psi_i\}_{i\in G} \subseteq \mathcal{L}_{el}\ \forall t {\sim_G}^* s\ (\mathcal{M}, t) \models \langle \bigwedge_{i\in G} K_i\psi_i\rangle\varphi \tag{5.5}$$

The model $\mathcal{M}$, formula φ and coalition $G = \{1, 2\}$ above is a counterexample showing that (5.5) is not expressed by $\langle G\rangle C_G\varphi$: (5.5) does not hold in state s, but $\mathcal{M}, s \models \langle G\rangle C_G\varphi$.

Summing up, it can be argued that *all* of the different notions of ability discussed in this section are *useful*. For example, in different contexts it might be useful to reason about what an agent can achieve by *guessing* the right actions to perform, while in others what she can achieve by *identifying* the correct actions with certainty. It is, however, of vital importance to discriminate between these different notions, for example in the analysis of security protocols.

5.5 Security Protocols

Consider a sender and a receiver who attempt to communicate a secret to each other without an eavesdropper learning it. A very powerful eavesdropper is one that intercepts all communications. This creates the setting where sender, receiver, and eavesdropper are three agents that can be modelled in a multi-S5 system and where all communications are public announcements by sender and receiver. One specific example of such a setting is known as the Russian Cards Problem (see [van Ditmarsch, 2003]). The setting is one where a pack of all different cards are distributed over the three 'players', where every player only knows his own cards, where sender and receiver have an informational advantage over the eavesdropper because they hold more cards, and where the 'secrets' that should not be divulged are about card ownership. Posed as a riddle it looks as follows—Alex and Brune are sender and receiver, Cha the eavesdropper:

> From a pack of seven known cards $0, 1, 2, 3, 4, 5, 6$ Alex and Brune each draw three cards and Cha gets the remaining card. How can Alex and Brune openly (publicly) inform each other about their cards, without Cha learning from any of their cards who holds it?

To simplify matters, assume that Alex has drawn $\{0, 1, 2\}$, that Brune has drawn $\{3, 4, 5\}$ and that Cha therefore has card 6, as in Figure 5.6.

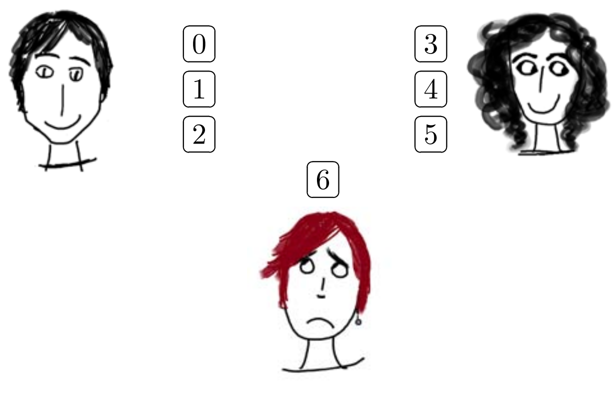


Figure 5.6: Three moody children playing in the *Russian cards* problem

The initial Kripke model $\mathcal{D}$ describing this setting consists of all possible card deals (valuations). In that model an epistemic class for an agent can be identified with the hand of cards of that agent. For example, given that Alex holds $\{0, 1, 2\}$, he cannot distinguish the four deals—allow us to use some suggestive notation—012.345.6, 012.346.5, 012.356.4, and 012.456.3 from one another.

Given that all announcements that can be made by a player are known by that player, they consist of unions of equivalence classes for that player and can therefore be identified with sets of alternative hands for that player. One solution is where

> Alex says "My hand of cards is one of $012, 034, 056, 135, 246$" after which Brune says "My hand of cards is one of $345, 125, 024$."

The last is equivalent in that information state to Brune saying "Cha has card 6." Alex and Brune in fact execute a protocol here, not in the sense of sets of sequences of announcements but in the sense of functions from local states of agents to nondeterministic choice between announcements. For example, Alex is executing "given cards i, j, k, the first of my five hands is that actual hand ijk; the second of my five hands to announce is ikl where k, l are chosen from the five remaining cards; the third is imn where m, n are the remaining two cards; etc...; shuffle the hands before announcing them."

We can describe this solution in logic. Agent a stands for Alex, b for Brune, and c for Cha. Let q_i stand for 'agent i holds card q' and let klm_i stand for $k_i \wedge l_i \wedge m_i$. The information and safety requirements are as follows — the conjunction in the formula suggests a conjunction over all hands of cards, 'Cha does not learn any card' means 'Cha does not learn the ownership of any card except her own card.'

Alex learns Brune's cards	$\bigwedge_{ijk}(ijk_b \longrightarrow K_a ijk_b)$	(one)
Brune learns Alex's cards	$\bigwedge_{ijk}(ijk_a \longrightarrow K_b ijk_a)$	(two)
Cha does not learn any card	$\bigwedge_{q=0}^{6}((q_a \longrightarrow \neg K_c q_a) \wedge (q_b \longrightarrow \neg K_c q_b))$	$(three)$

These requirements should hold throughout the model after protocol completion (i.e., they should be common knowledge between Alex and Brune). The safety requirement should be satisfied both at the end and in all intermediate stages: after any announcement that forms part of such a protocol.

All protocols are finite, because the model is finite and all informative announcements result in proper model restriction. But it is unclear how long such protocols need to be. The above solution was of length two, but settings that require strictly longer protocols are also known. The uncertain but finite length cannot be described in public announcement logic, but it can be described in group announcement logic. The diamond in $\langle ab \rangle \varphi$ refers to arbitrarily finite length protocols taking place between sender a and receiver b in the presence of other agents, such as the eavesdropper, as was discussed in Section 5.1.3.

Let us see how this works for the length-two protocol above that solves the Russian Cards Problem. First, we model the solution in public announcement logic. In the solution, first Alex announces $012_a \vee 034_a \vee 056_a \vee 135_a \vee 246_a$ $(alex)$. Then Brune announces $345_b \vee 125_b \vee 024_b$

($brune$). After these two announcements the solution requirements are satisfied. This can now be described in various ways: as a sequence of two announcements by different agents, as a sequence of two simultaneous announcements by Alex and Brune, or as a single announcement by Alex and Brune.

$$\mathcal{D}, 012.345.6 \models \langle K_a alex\rangle\langle K_b brune\rangle(one \wedge two \wedge three)$$
$$\mathcal{D}, 012.345.6 \models \langle K_a alex \wedge K_b \mathcal{T}\rangle\langle K_a \mathcal{T} \wedge K_b brune\rangle(one \wedge two \wedge three)$$
$$\mathcal{D}, 012.345.6 \models \langle K_a(alex \wedge [K_a alex \wedge K_b \mathcal{T}]\mathcal{T}) \wedge K_b(\mathcal{T} \wedge [K_a alex \wedge K_b \mathcal{T}]brune)\rangle(one \wedge two \wedge three)$$

The last one implies that we have, in this case:

$$\mathcal{D}, 012.345.6 \models \langle ab\rangle(one \wedge two \wedge three)$$

Given that we should be able to realize the three postconditions after any execution of the underlying protocol, and regardless of the initial card deal, the existence of a successful protocol to realize them can be expressed all at once by the model validity

$$\mathcal{D} \models \langle ab\rangle(one \wedge two \wedge three)$$

or in other words

$$\text{“}\langle ab\rangle(one \wedge two \wedge three) \text{ is valid in the initial model for card deals”} \tag{5.6}$$

In principle, we can now model check this formula in that model, thus establishing that a secure exchange is possible under the uncertainty conditions about card ownership in a fully automated way.

We have so far overlooked one aspect of the meaning of announcements executing such protocols. The security requirement *three* should be an invariant: its validity throughout the model should be preserved after every good announcement. In this particular case we can enforce that, because its negation is a positive formula: if it is ever not preserved, then it is lost forever afterwards. Therefore, it suffices to guarantee it after the execution of the protocol. Thus the above expression also incorporates that invariance.

One must be careful when interpreting the meaning of the existence of sequences of announcements. If we can replace the two successive announcements: Alex says “My hand of cards is one of $012, 034, 056, 135, 246$” after which Brune says “My hand of cards is one of $345, 125, 024$”, by a single one, does that not mean that all protocols can be reduced to length 1? And what would in this case that single simultaneous announcement be? Well: as both agents are announcing facts and not knowledge, their single announcement is simply the conjunction of their successive announcements. As the second one for Alex and the first one for Brune was ‘true’ (vacuous), this means that they could *simultaneously* have made their successive announcements: Alex says "My hand of cards is one of $012, 034, 056, 135, 246$" and simultaneously Brune says "My hand of cards is one of $345, 125, 024$". Unfortunately, even though this indeed solves the problem, the agents do not know the public consequences of their joint action merely from the public consequences of their individual part in it. This situation was discussed in the previous section: there *is* a simultaneous announcement by Alex and Brune which will achieve the goal, but Alex and Brune do not *know* that their respective announcements will achieve the goal – they will not achieve the goal in all the states they consider possible. A different execution of the protocol for Alex, when he holds cards $\{0, 1, 2\}$, is the announcement "My hand of cards is one of $012, 035, 046, 134, 256$". From that with Brune's above announcement Cha can deduce straightaway that the card deal is 012.345.6. And, obviously, Brune does not know whether Alex is going to announce the original or the alternative set of five hands, or any of many others. In epistemic terms,we can sum up our achievements for this security setting as follows, also using the discussion and results of Section 5.4.

$$\mathcal{D} \models \langle ab \rangle (one \wedge two \wedge three) \tag{5.7}$$

$$\mathcal{D} \not\models \langle ab \rangle K_a (one \wedge two \wedge three) \tag{5.8}$$

$$\mathcal{D} \not\models \langle ab \rangle K_b (one \wedge two \wedge three) \tag{5.9}$$

$$\mathcal{D} \models \langle a \rangle K_a (two \wedge three \wedge \langle b \rangle K_b (one \wedge two \wedge three)) \tag{5.10}$$

Recall (Proposition 5.34.2) that a formula of the form $\langle i \rangle K_i \varphi$ expresses the fact that agent i knows *de re* that she can achieve φ; that she can make an announcement that will ensure that φ is true in any state that i considers possible. Thus, the last formula above, (5.10), expresses the fact that there is an announcement that Alex can make after which Brune has learnt his cards and Cha remains ignorant, no matter which of the four card deals Alex considers possible is the actual one, and such that Brune then can make an announcement after which all three requirements hold. Thus, it is rational for Alex to make that announcement, and for Brune to make a proper counter-announcement in the resulting state. Unlike the property *(5.6), (5.10) shows that Alex and Brune know* how *to execute a successful protocol.*

Chapter 6

Permission and Public Announcements

> Consider an art school examining works at an exhibition. A student is supposed to select one of the displayed works and is then permitted to make a number of intelligent observations about it, sufficient to impress the examiners with the breadth of her knowledge. Now in such cases it never hurts to be more informative than necessary, in order to pass the exam, but a certain minimum amount of intelligent information has to be passed on. This particular museum has both the Night Watch by Rembrandt and Guernica by Picasso on display in the same room! You pass the exam if you observe about the Night Watch that a big chunk of a meter or so is missing in the left corner, that was cut off in order to make the painting fit in the Amsterdam Townhall (a_1), and that the painter was Rembrandt van Rijn (a_2). Clearly, this is not a very difficult exam. You also pass the exam if you make two of the three following observations: that Guernica depicts the cruelties of the Spanish Civil war (b_1), that it is painted in black and white and not in colour (b_2), and that the painter was Pablo Picasso (b_3). It is not permitted to make observations about different paintings at the same time, so any conjunction of a_i's and b_j's is not permitted: it would amount to bad judgement if you cannot focus on a single painting. You are obliged to make two observations about the Rembrandt and in that case say nothing about the Picasso, or to make at least two of the three possible observations about the Picasso and in that case say nothing about the Rembrandt. We can treat the permissions and obligations in this setting in an extension of public announcement logic.

To formalize the concept of "having the permission to say" we extend Plaza's public announcement logic [Plaza, 1989] with a modal operator P of permission, where $P\varphi$ expresses that it is permitted to say (i.e., announce) φ.

Our proposal can be seen as an adaption of the dynamic logic of permission proposed by [van der Meyden, 1996]. Van der Meyden's proposal was later elaborated on by [Pucella and Weissman, 2004]). In Van der Meyden's work, $\Diamond(\alpha, \varphi)$ means "*there is a way to execute α which is permitted and after which φ is true.*" We treat the particular case where actions are public announcements. Thus, for α in van der Meyden's $\Diamond(\alpha, \varphi)$ we take an announcement $\psi!$ such that $\Diamond(\psi!, \varphi)$ now means "*there is a way to execute the announcement ψ which is permitted and after which φ is true.*" The executability precondition for an announcement ('truthful public announcement') is the truth of the announcement formulas, therefore, the latter is equivalent to "*ψ is true and it is permitted to announce ψ, after which φ is true*". This suggests an equivalence of $\Diamond(\psi!, \varphi)$ with, in our setting, $P\psi \wedge \langle\psi\rangle\varphi$, but our operator behaves slightly different. This is because we assume that if you have the permission to say something, you also have the permission to say something weaker, and because our binary permission operator allows update of permissions after an announcement.

[van der Meyden, 1996] also introduces a weak form of obligation. The meaning of $\mathcal{O}(\alpha, \varphi)$ is "*after any permitted execution of α, φ is true*". Similarly, we also introduce an obligation operator $O\varphi$, meaning "*the agents are obliged to announce φ.*"

This chapter further relates to the extension of public announcement logic with protocols by [van Benthem et al., 2009, Wang et al., 2009]. In their approach, one cannot just announce anything that is true, but one can only announce a true formula that is part of the *protocol*, i.e., that is the first formula in a sequence of formulas (standing for a sequence of successive announcements) that is a member of a set of such sequences called the protocol. In other words, one can only announce *permitted* formulas.

In the setting of informative actions like announcements we leave the beaten track for permission in one important aspect. Cha is given permission by her parents to invite uncle Jean for her 8th birthday party with his children friends and for a delightful canoe trip on the river Rhône, but not for the family dinner afterwards. When seeing uncle Jean, she only mentions the canoe trip but not the children's party. She does not mention the family dinner. Has she transgressed the permissions given? Of course not. Permission to say $p \wedge q$ implies permission to say only q. She has also not transgressed the permission if she were not to invite him at all. Permission to say $p \wedge q$ implies permission to say nothing, i.e., to say the always true and therefore uninformative statement $\top$. Similarly, an obligation to say φ entails the obligation for anything entailed by φ. If you are obliged to say $p \wedge q$ you are also obliged to say q. Now saying q does not therefore mean you have fulfilled the original obligation of $p \wedge q$, you have only partially fulfilled the entailed weaker obligation of q. It may be worth to already point out as this stage that the weakening of announcement formulas is unrelated to Ross's Paradox (see [Ross, 1941]): this is about the obligation to do one of two possible *actions*—the alternative to that in public announcement logic would be the obligation to make one of two possible announcements (announcement of) φ and (announcement of) ψ, completely different from the obligation to make an announcement of (the disjunctive formula) $\varphi \vee \psi$. As we saw in Section 3.2, in dynamic epistemic logics, there is a clear distinction between actions and formulas. We comment the validity of the classical deontic paradoxes in Section 6.5.1.

We present first the syntax and the semantics of our logic, continue with various validities and semantics observations, and conclude with the completeness of the axiomatization and the decidability of the problem of satisfiability. After that we present an example in detail: the card game La Belote. We conclude with some observations relating to standard deontic logical topics, and a more detailed comparison of our proposal with the relevant dynamic logical literature, i.e. with [van der Meyden, 1996, Pucella and Weissman, 2004, van Benthem et al., 2009].

Most parts of this chapter have been published in the Journal of Philosophical Logic ([Balbiani and Seban, 2011]).

6.1 The Logic of Permission and Obligation to Speak

6.1.1 Syntax

The logic $POPAL$ of permitted announcements is an extension of the multi-agent epistemic logic of public announcements ([Plaza, 1989]).

Definition 6.1 (Language $\mathcal{L}_{popal}$) *The language $\mathcal{L}_{popal}$ over a countable set of agents AG and a countable set of propositional atoms $PROP$ is defined as follows:*

$$\varphi ::= p|\bot|\neg\varphi|\psi \vee \varphi|K_i\varphi|[\psi]\varphi|P(\psi,\varphi)|O(\psi,\varphi)$$

where $i \in AG$ and $p \in PROP$. The language $\mathcal{L}_{poel}$ is the fragment without announcement construct $[\psi]\varphi$, the language $\mathcal{L}_{pal}$ is the fragment without O and P, and the language $\mathcal{L}_{el}$ is the fragment restricted to the Boolean and epistemic operators.

The intuitive reading of $K_i\varphi$ is *"agent i knows that φ is true"* whereas $[\psi]\varphi$ is read as *"after announcing ψ, it is true that φ"*. We read $P(\psi,\varphi)$ as "*(ψ is true and) after announcing ψ, it is permitted to announce φ*". Similarly, $O(\psi,\varphi)$ stands for "*(ψ is true and) after announcing ψ, it is obligatory to announce φ*". Note that announcements are assumed to be public and truthful. Definitions by abbreviation of other Boolean operators are standard. Moreover, we define by abbreviation:

- $\langle\psi\rangle\varphi := \neg[\psi]\neg\varphi$;
- $P\varphi := P(\top,\varphi)$;
- $O\varphi := O(\top,\varphi)$.

Formula $P\varphi$ stands for "It is permitted to announce φ" and $O\varphi$ stands for "It is obligatory to announce φ" (the semantics also entails the *truth* of φ, in both cases); $\langle\psi\rangle\varphi$ stands for "ψ is true and after announcing ψ, φ is true." Note the difference with $[\psi]\varphi$: "*if* ψ is true, then after announcing it, φ is true." The latter is vacuously true if the announcement cannot be made.

The degree *deg* of a formula is a concept that will be used in the completeness proof, in Section 6.2.2. It keeps count of the number of P and O operators in a given formula.

Definition 6.2 (Degree) *The degree of a formula $\varphi \in \mathcal{L}_{popal}$ is defined inductively on the structure of φ as follows:*

$$\begin{array}{lllllll} deg(p) & = & 0 & deg(\psi_1 \vee \varphi_2) & = & max(deg(\psi_1), deg(\psi_2)) \\ deg(\bot) & = & 0 & deg([\psi]\varphi) & = & deg(\psi) + deg(\varphi) \\ deg(\neg\psi) & = & deg(\psi) & deg(P(\psi,\varphi)) & = & deg(\psi) + deg(\varphi) + 1 \\ deg(K_i\psi) & = & deg(\psi) & deg(O(\psi,\varphi)) & = & deg(\psi) + deg(\varphi) + 1 \end{array}$$

This is therefore not the usual *modal* degree function, that counts K_i operators. For all formulas $\varphi \in \mathcal{L}_{popal}$, $deg(\varphi) = 0$ iff φ does not contain any occurrence of P or O iff $\varphi \in \mathcal{L}_{pal}$.

6.1.2 Semantics

The models of our logic are Kripke models with an additional *permission relation* $\mathcal{P}$ between states and pairs of sets of states, that represents, for each state, the announcements that are permitted to be done in this state.

Definition 6.3 (Permission Kripke Model) *Given a set of agents AG and a set of atoms $PROP$,* permission Kripke models *have the form $\mathcal{M} = (S, \{\sim_i\}_{i\in AG}, V, \mathcal{P})$ with S a non-empty set of states, for each $i \in AG$, $\sim_i$ an equivalence relation between states of S, valuation function V mapping propositional atoms to subsets of S, and $\mathcal{P} \subseteq S \times 2^S \times 2^S$ such that if $(s, S', S'') \in \mathcal{P}$ then $s \in S''$ and $S'' \subseteq S'$.*

If the equivalence relation $\sim_i$ holds between states $s, t \in S$, this means that, as far as agent i is concerned, s and t are indiscernible. The membership of (s, S', S'') in $\mathcal{P}$ can be interpreted as follows: in state s, after an announcement that restricts the set of possible states to S', a further announcement in S' that restricts that set to S'' is permitted. We will explain this in more detail after giving the semantics.

We simultaneously define the restriction $\mathcal{M}_\psi$ of a model $\mathcal{M}$ after the public announcement of ψ, and the satisfiability relation $\models$. In the definitions we use the abbreviation $[\![\psi]\!]_\mathcal{M} = \{s \in S \mid \mathcal{M}, s \models \psi\}$. If no ambiguity results, we occasionally write $[\![\psi]\!]$ instead of $[\![\psi]\!]_\mathcal{M}$.

Definition 6.4 (Restricted model) *For any model $\mathcal{M}$ and any $\psi \in \mathcal{L}_{popal}$, we define the* restriction *$\mathcal{M}_\psi = (S_\psi, \sim_i^\psi, V_\psi, \mathcal{P}_\psi)$ where:*

- $S_\psi = [\![\psi]\!]_\mathcal{M}$
- *for all i, $\sim_i^\psi = \sim_i \cap (S_\psi \times S_\psi)$*
- *for all $p \in PROP$, $V_\psi(p) = V(p) \cap S_\psi$*
- $\mathcal{P}_\psi = \{(s, S', S'') \in \mathcal{P} \mid s \in S_\psi,\ S' \subseteq S_\psi, S'' \subseteq S_\psi\}$

Definition 6.5 (Satisfiability relation) *Let $\mathcal{M}$ be a model and let s be a state of S. The satisfiability relation $\models$ is defined inductively on the structure of φ:*

$\mathcal{M}, s \models p$ *iff* $s \in V(p)$

$\mathcal{M}, s \not\models \bot$

$\mathcal{M}, s \models \neg\psi$ *iff* $\mathcal{M}, s \not\models \psi$

$\mathcal{M}, s \models \psi_1 \vee \psi_2$ *iff* $(\mathcal{M}, s \models \psi_1$ *or* $\mathcal{M}, s \models \psi_2)$

$\mathcal{M}, s \models K_i\psi$ *iff for all* $t \sim_i s$, $\mathcal{M}, t \models \psi$

$\mathcal{M}, s \models [\psi]\chi$ *iff* $(\mathcal{M}, s \models \psi \Rightarrow \mathcal{M}_\psi, s \models \chi)$

$\mathcal{M}, s \models P(\psi, \chi)$ *iff for some* $(s, [\![\psi]\!]_\mathcal{M}, S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_\mathcal{M}$

$\mathcal{M}, s \models O(\psi, \chi)$ *iff for all* $(s, [\![\psi]\!]_\mathcal{M}, S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_\mathcal{M}$.

For all $\varphi \in \mathcal{L}_{popal}$, $\mathcal{M} \models \varphi$ iff for all $s \in S$, $\mathcal{M}, s \models \varphi$; and $\models \varphi$ iff for all models $\mathcal{M}$ we have $\mathcal{M} \models \varphi$.

We do not impose that S' and S'' are denotations of formulas in the language for (s, S', S'') to be in $\mathcal{P}$. This semantics is thus more general than the intuitive one for "having the permission to say". Indeed, if S' or S'' do not correspond to a restriction of S made by an announcement, then $(s, S', S'') \in \mathcal{P}$ does not correspond to some announcement being permitted.

The semantics of $P(\psi, \chi)$ expresses that after announcement of ψ it is permitted to announce a χ weaker than the restriction given in the relation $\mathcal{P}$. If the S'' in $(s, [\![\psi]\!], S'')$ is the denotation of some $[\![\langle\psi\rangle\varphi]\!]$, we get that after announcement of ψ it is permitted to announce a χ weaker than (implied by) φ.

Remark 6.6 *For any finite list of formulas $\sigma = (\sigma_1, \ldots, \sigma_n)$ we can define $\mathcal{M}_\sigma$ by a direct induction on the length of σ. Similarly, for every $\varphi \in \mathcal{L}_{popal}$ we abbreviate $[\sigma]\varphi := [\sigma_1] \ldots [\sigma_n]\varphi$. This will be particularly useful in Section 6.6.*

We shall make another important observation before going ahead. In Proposition 6.7 we notice that $\mathcal{L}_{popal}$ may be reduced to a language with a unary operator $P\varphi$ that we would read simply 'it is permitted to say φ'. We get the equivalence of their expressivity through the translation $P(\psi, \varphi) := \langle\psi\rangle P\varphi$. We prove this result after discussing it a while.

The language with unary operator (let us call it $\mathcal{L}^1_{popal}$) has an advantage and an inconvenience with respect to $\mathcal{L}_{popal}$. Its advantage is that it is easier to read and to translate into natural language: sentences such as 'after the announcement of ψ it is permitted to say φ' are quite boring and not intuitive. Its inconveniences are on one hand that the unary nature of its operators is farther to the semantics made of triplets that we just saw, on the other hand that technical results are more difficult to obtain, because the language cannot be reduced to a language without announcements. In the following chapter we decided to use a kind of language with unary operator, but in this one let us use binary ones.

As we show now, both languages are expressively equivalent, and the reader may prefer to translate binary operators into unary ones using the translation $P(\psi,\varphi) := \langle\psi\rangle P\varphi$ used in the following proof.

Proposition 6.7 *The language $\mathcal{L}_{popal}$ is expressively equivalent to the language $\mathcal{L}^1_{popal}$.*

Proof Clearly, $\mathcal{L}_{popal}$ is at least as expressive as $\mathcal{L}^1_{popal}$, indeed the latter is a sublanguage of the former, considering that the $P\varphi$ and $O\varphi$ are defined by the following abbreviations $P\varphi := P(\top,\varphi)$ and $O\varphi := O(\top,\varphi)$. To prove the equivalence, it is thus sufficient to prove that for all $\psi,\varphi \in \mathcal{L}_{popal}$, $\models P(\psi,\varphi) \longleftrightarrow \langle\psi\rangle P(\top,\varphi)$ and $\models O(\psi,\varphi) \longleftrightarrow \langle\psi\rangle O(\top,\varphi)$. We refer to proposition 6.9 to see that $\models [\psi]P(\top,\varphi) \longleftrightarrow (\psi \longrightarrow P(\psi,\varphi))$ and $\models [\psi]O(\top,\varphi) \longleftrightarrow (\psi \longrightarrow O(\psi,\varphi))$. It remains to prove that $\models P(\psi,\varphi) \longrightarrow \psi$ (and *idem* for O). But by definition of $\mathcal{T}$, if $(s, [\![\psi]\!]_\mathcal{M}, S'') \in \mathcal{P}$ then $s \in [\![\psi]\!]_\mathcal{M}$. Then we have the wanted result.

□

6.1.3 Example: Art School

Consider the example presented in the introduction. In an art school examination you are asked to "describe precisely one (and only one) of the presented pictures". There are two distinct sets of intelligent observations to make (modelled as atomic propositional variables): $A = \{a_1, a_2\}$, $B = \{b_1, b_2, b_3\}$, with $A \cup B = PROP$. The domain of discourse consists of all possible valuations $S = 2^{PROP}$, in the actual state s all atoms are in fact true, and our student is in fact an omniscient agent g (i.e. $\sim_g = id_S$) that can announce anything she likes. The set $\mathcal{P}$ is given as $\mathcal{P} = \{(s, [\![\top]\!], [\![a_1 \wedge a_2]\!]), (s, [\![\top]\!], [\![b_1 \wedge b_2]\!]), (s, [\![\top]\!], [\![b_1 \wedge b_3]\!]), (s, [\![\top]\!], [\![b_2 \wedge b_3]\!]), (s, [\![\top]\!], [\![b_1 \wedge b_2 \wedge b_3]\!])\}$. Note that $[\![\top]\!] = S$. We now have that

- It is permitted to say a_1 ($\mathcal{M}, s \models P(\top, a_1)$), because $(s, [\![\top]\!], [\![a_1 \wedge a_2]\!]) \in \mathcal{P}$ and $[\![a_1 \wedge a_2]\!] \subseteq [\![\langle\top\rangle a_1]\!]_\mathcal{M}$ (where $[\![\langle\top\rangle a_1]\!]_\mathcal{M} = [\![a_1]\!]_\mathcal{M}$): it is permitted to say something weaker than $a_1 \wedge a_2$.
- It is not permitted to say $a_1 \wedge b_2$ ($\mathcal{M}, s \models \neg P(\top, a_1 \wedge b_2)$) because the denotation of that formula is not contained in either of the members of the set $\mathcal{P}$.
- It is not obligatory to say a_1 ($\mathcal{M}, s \not\models O(\top, a_1)$), because it is permitted to say $b_1 \wedge b_2$, and $[\![a_1]\!] \not\subseteq [\![b_1 \wedge b_2]\!]$.
- It is obligatory to say $ob := (a_1 \wedge a_2) \vee (b_1 \wedge b_2) \vee (b_2 \wedge b_3) \vee (b_1 \wedge b_3)$ as all members of $\mathcal{P}$ are stronger.

This last obligation is also the strongest obligation in this setting. It is, e.g., also obligatory to say $a_1 \vee b_1 \vee b_2$ ($\mathcal{M}, s \models O(\top, a_1 \vee b_1 \vee b_2)$) because this is weaker than *ob*. However, as already mentioned, this does not mean that a student has fulfilled her obligation when saying $a_1 \vee b_1 \vee b_2$ - she then only fulfills *part* of her obligation (and will therefore fail the exam!). We observe that our intuition of what an obligation is corresponds to the strongest obligation under our definition—reasons to prefer the current definition are technical, such as getting completeness right. A different definition, nearer to our intuition, is proposed in Chapter 7.

6.1.4 Valid Principles and Other Semantic Results

The O and P operators are not interdefinable. This is because the obligation to say φ means that anything not entailing φ may not be permitted to say, and not only that it is not permitted to say $\neg\varphi$. As an example, consider the following two models that have the same domain $S = \{s_1, s_2\}$, the same valuation $V(p) = \{s_1\}$, the same epistemic relation $\sim_i = S \times S$, but that differ on the permission relation: $\mathcal{M} = (S, V, \sim_i, \mathcal{P})$ and $\mathcal{M}' = (S, V, \sim_i, \mathcal{P}')$ where $\mathcal{P} = \{(s_1, S, \{s_1\}), (s_2, S, S)\}$ and $\mathcal{P}' = \{(s_1, S, \{s_1\}), (s_1, S, S), (s_2, S, S)\}$.

Let $\mathcal{L}^-_{popal}$ be the language without the obligation operator O. The models $\mathcal{M}$ and $\mathcal{M}'$ satisfy the same formulas in that language: for all $k \in \{1, 2\}$ and all $\varphi \in \mathcal{L}^-_{popal}$, ($\mathcal{M}, s_k \models \varphi$ iff $\mathcal{M}', s_k \models \varphi$). The proof is obvious for all inductive cases of φ except when φ takes shape $P(\psi, \varphi)$. In that case, observe from the semantics of P and the given relations $\mathcal{P}$ and $\mathcal{P}'$ that *only* formulas of type $P(\top, \varphi_2)$ (or simply $P(\varphi_2)$) can be true in these models, as the second argument of all triples in $\mathcal{P}$ and $\mathcal{P}'$ is the entire domain S. Further these two properties: first anything that is true in s_1 is permitted to be said, formally for all $\varphi \in \mathcal{L}_{popal}$, $\mathcal{M}, s_1 \models \varphi \leftrightarrow P(\varphi)$ and $\mathcal{M}', s_1 \models \varphi \leftrightarrow P(\varphi)$. Second anything that is permitted to be said in s_2 is a validity of the model, formally for all $\varphi \in \mathcal{L}_{popal}$, $\mathcal{M}, s_2 \models P(\varphi) \rightarrow (\varphi \leftrightarrow \top)$ and $\mathcal{M}', s_2 \models P(\varphi) \rightarrow (\varphi \leftrightarrow \top)$. So $\mathcal{M}$ and $\mathcal{M}'$ are modally equivalent in $\mathcal{L}^-_{popal}$. On the other hand, as (s, S, S) is not in $\mathcal{P}$ we have that $\mathcal{M}, s_1 \models O(\top, p)$ but $\mathcal{M}', s_1 \models \neg O(\top, p)$, so the models are not modally equivalent in $\mathcal{L}_{popal}$. We conclude that:

Proposition 6.8 *$\mathcal{L}^-_{popal}$ is strictly less expressive than $\mathcal{L}_{popal}$.*

The standard validities for public announcement logic are preserved in this extension of the logic with permission and obligation (for details, see a standard introduction like [van Ditmarsch et al., 2007]):

- $\models [\psi]p \leftrightarrow (\psi \longrightarrow p)$
- $\models [\psi]\bot \leftrightarrow \neg\psi$
- $\models [\psi]\neg\varphi \leftrightarrow (\psi \longrightarrow \neg[\psi]\varphi)$
- $\models [\psi](\varphi_1 \vee \varphi_2) \leftrightarrow ([\psi]\varphi_1 \vee [\psi]\varphi_2)$

- $\models [\psi]K_i\varphi \leftrightarrow (\psi \longrightarrow K_i[\psi]\varphi)$

- $\models [\psi_1][\psi_2]\varphi \leftrightarrow [\langle\psi_1\rangle\psi_2]\varphi$

For example, $[\psi]p \leftrightarrow (\psi \longrightarrow p)$ says that p is true after announcement of ψ iff ψ implies p (is true). As ψ is the condition to be able to make the announcement, this principle merely says that an announcement cannot change the valuation of atoms. Of course, for other formulas than atoms we cannot get rid of the announcement that way. A typical counterexample (the Moore-sentence) is that $(p \wedge \neg K_i p) \longrightarrow (p \wedge \neg K_i p)$ is a trivial validity whereas $[p \wedge \neg K_i p](p \wedge \neg K_i p)$ is false, because whenever $p \wedge \neg K_i p$ can be announced, p is known afterwards: $K_i p$.

Additional to the principles for public announcement logic, two principles address how to treat a permission or obligation operator after an announcement.

Proposition 6.9

For all $p \in POPAL$, all $\psi, \varphi, \psi_1, \psi_2, \varphi_1, \varphi_2 \in \mathcal{L}_{popal}$:

1. $\models [\psi_1]P(\psi_2, \varphi) \leftrightarrow (\psi_1 \longrightarrow P(\langle\psi_1\rangle\psi_2, \varphi))$

2. $\models [\psi_1]O(\psi_2, \varphi) \leftrightarrow (\psi_1 \longrightarrow O(\langle\psi_1\rangle\psi_2, \varphi))$

Proof For all $\mathcal{M}$, all $s \in S$ and all $\psi_1, \psi_2, \varphi \in \mathcal{L}_{popal}$ we have:

1. ($\Rightarrow$) Suppose that $\mathcal{M}, s \models [\psi_1]P(\psi_2, \varphi)$ and $s \in [\![\psi_1]\!]_{\mathcal{M}}$. Then for some $S'' \subseteq [\![\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}_{\psi_1}}, (s, [\![\psi_2]\!]_{\mathcal{M}_{\psi_1}}, S'') \in \mathcal{P}_{\psi_1}$. This implies that for some $S'' \subseteq [\![\langle\psi_1\rangle\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}}$, $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}_{\psi_1}$, i.e. for some $S'' \subseteq [\![\langle\psi_1\langle\psi_2\rangle\rangle\varphi]\!]_{\mathcal{M}}$, $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}$. Finally $\mathcal{M}, s \models P(\langle\psi_1\rangle\psi_2, \varphi)$.

   ($\Leftarrow$) Suppose that $\mathcal{M}, s \models (\psi_1 \longrightarrow P(\langle\psi_1\rangle\psi_2, \varphi))$. If $\mathcal{M}, s \not\models \psi_1$ then obviously $\mathcal{M}, s \models [\psi_1]P(\psi_2, \varphi)$. Otherwise $s \in [\![\psi_1]\!]_{\mathcal{M}}$ and $\mathcal{M}, s \models P(\langle\psi_1\rangle\psi_2, \varphi)$. Therefore, there exists $S'' \subseteq [\![\langle\langle\psi_1\rangle\psi_2\rangle\varphi]\!]_{\mathcal{M}}$ such that $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}$. Thus $s \in [\![\psi_1]\!]_{\mathcal{M}}$, $S'' \subseteq [\![\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}_{\psi_1}}$ and $(s, [\![\psi_2]\!]_{\mathcal{M}_{\psi_1}}, S'') \in \mathcal{P}_{\psi_1}$. Finally $\mathcal{M}, s \models [\psi_1]P(\psi_2, \varphi)$.

2. ($\Rightarrow$) Suppose that $\mathcal{M}, s \models [\psi_1]O(\psi_2, \varphi)$ and $s \in [\![\psi_1]\!]_{\mathcal{M}}$. Then for all $(s, [\![\psi_2]\!]_{\mathcal{M}_{\psi_1}}, S'') \in \mathcal{P}_{\psi_1}$, $S'' \subseteq [\![\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}_{\psi_1}}$. This implies that for all $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}_{\psi_1}$, $S'' \subseteq [\![\langle\psi_1\rangle\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}}$ i.e. for all $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi_1\rangle\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}}$. Finally $\mathcal{M}, s \models O(\langle\psi_1\rangle\psi_2, \varphi)$.

   ($\Leftarrow$) Suppose that $\mathcal{M}, s \models (\psi_1 \longrightarrow O(\langle\psi_1\rangle\psi_2, \varphi))$. If $\mathcal{M}, s \not\models \psi_1$ then obviously $\mathcal{M}, s \models [\psi_1]O(\psi_2, \varphi)$. Otherwise $s \in [\![\psi_1]\!]_{\mathcal{M}}$ and $\mathcal{M}, s \models O(\langle\psi_1\rangle\psi_2, \varphi)$. Therefore, for all $(s, [\![\langle\psi_1\rangle\psi_2]\!]_{\mathcal{M}}, S'') \in \mathcal{P}$ we have $S'' \subseteq [\![\langle\langle\psi_1\rangle\psi_2\rangle\varphi]\!]_{\mathcal{M}}$. Thus $s \in [\![\psi_1]\!]_{\mathcal{M}}$ and for all $(s, [\![\psi_2]\!]_{\mathcal{M}_{\psi_1}}, S'') \in \mathcal{P}_{\psi_1}$ we have $S'' \subseteq [\![\langle\psi_2\rangle\varphi]\!]_{\mathcal{M}_{\psi_1}}$. Finally $\mathcal{M}, s \models [\psi_1]O(\psi_2, \varphi)$.

□

For example, principle $[\psi_1]P(\psi_2, \varphi) \leftrightarrow (\psi_1 \longrightarrow P(\langle\psi_1\rangle\psi_2, \varphi))$ of Proposition 7.18 says the following: *"(After announcing ψ_1 we have that (ψ_2 is true and after announcing ψ_2 it is permitted to announce φ)) iff (On condition that ψ_1 is true ($\langle\psi_1\rangle\psi_2$ is true and after announcing $\langle\psi_1\rangle\psi_2$ it is permitted to say φ))."*

Using the meaning of the public announcement operator, the right part is the same as *"On condition that ψ_1 is true, after announcing ψ_1, ψ_2 is true and after then announcing ψ_2 it is permitted to say φ."* Which gets us back to the left part of the original equivalence.

Another validity of the logic spells out that equivalent announcements lead to equivalent permissions.

Proposition 6.10 *For all models $\mathcal{M}$ and all formulas $\psi, \psi', \varphi, \varphi' \in \mathcal{L}_{popal}$: If $\mathcal{M} \models (\psi \longleftrightarrow \psi') \wedge ([\psi]\varphi \longrightarrow [\psi']\varphi')$ then $\mathcal{M} \models P(\psi, \varphi) \longrightarrow P(\psi', \varphi')$ and $\mathcal{M} \models O(\psi, \varphi) \longrightarrow O(\psi', \varphi')$.*

Proof For all $\psi, \psi', \varphi, \varphi' \in \mathcal{L}_{popal}$, if $\mathcal{M} \models (\psi \longleftrightarrow \psi')$ and $\mathcal{M} \models \langle\psi\rangle\varphi \longrightarrow \langle\psi'\rangle\varphi'$, then $[\![\psi]\!]_{\mathcal{M}} = [\![\psi']\!]_{\mathcal{M}}$ and $[\![\langle\psi\rangle\varphi]\!]_{\mathcal{M}} \subseteq [\![\langle\psi'\rangle\varphi']\!]_{\mathcal{M}}$. It implies that for all $(s, [\![\psi]\!]_{\mathcal{M}}, S'') \in \mathcal{P}$, we have $(s, [\![\psi']\!]_{\mathcal{M}}, S'') \in \mathcal{P}$ and if $S'' \subseteq [\![\langle\psi\rangle\varphi]\!]$ then $S'' \subseteq [\![\langle\psi\rangle'\varphi']\!]$.

□

We continue with a proposition on allowed logical compositions of permitted and obliged announcements.

Proposition 6.11 *For all $\psi, \varphi, \varphi_1, \varphi_2 \in \mathcal{L}_{popal}$:*

1. $\models (O(\psi, \varphi_1) \wedge O(\psi, \varphi_2)) \leftrightarrow O(\psi, \varphi_1 \wedge \varphi_2)$
2. $\models (P(\psi, \varphi_1) \wedge O(\psi, \varphi_2)) \longrightarrow P(\psi, \varphi_1 \wedge \varphi_2)$
3. $\models (\psi \wedge O(\psi, \varphi) \wedge \neg P(\psi, \varphi)) \leftrightarrow (\psi \wedge \neg P(\psi, \top))$

Proof For all models $\mathcal{M}$ and all state $s \in S$ we have

1. $\mathcal{M}, s \models O(\psi, \varphi_1) \wedge O(\psi, \varphi_2)$ iff for all $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi_1]\!]$ & $S'' \subseteq [\![\langle\psi\rangle\varphi_2]\!]$ iff for all $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi_1]\!] \cap [\![\langle\psi\rangle\varphi_2]\!] = [\![\langle\psi\rangle(\varphi_1 \wedge \varphi_2)]\!]$ iff $\mathcal{M}, s \models O(\psi, \varphi_1 \wedge \varphi_2)$.

2. Suppose $\mathcal{M}, s \models P(\psi, \varphi_1) \wedge O(\psi, \varphi_2)$. Then for some $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi_1]\!]$ and for all $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi_2]\!]$. Thus, for some $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi_1]\!] \cap [\![\langle\psi\rangle\varphi_2]\!] = [\![\langle\psi\rangle(\varphi_1 \wedge \psi_2)]\!]$, which is equivalent to $\mathcal{M}, s \models P(\psi, \varphi_1 \wedge \varphi_2)$.

3. $\mathcal{M}, s \models \psi \wedge O(\psi, \varphi) \wedge \neg P(\psi, \varphi)$ if and only if $\mathcal{M}, s \models \psi$ and for all $(s, [\![\psi]\!], S'') \in \mathcal{P}$, $S'' \subseteq [\![\langle\psi\rangle\varphi]\!]$ and $S'' \not\subseteq [\![\langle\psi\rangle\varphi]\!]$. This is equivalent to $\mathcal{M}, s \models \psi$ and the fact that there is no S'' such that $(s, [\![\psi]\!], S'') \in \mathcal{P}$, which means that $\mathcal{M}, s \models \psi \wedge \neg P(\psi, \top)$.

$\square$

We also have that if φ is permitted, then any $\varphi \vee \psi$ is *also* permitted (namely anything weaker than φ is also permitted) and similarly, if φ is obligatory, than any $\varphi \vee \psi$ is *also* obligatory. In the example in the previous section we already illustrated that this notion of weakened obligation is not intuitive—one might rather see the announcement of $\varphi \vee \psi$ as something *towards* fulfilling an obligation. The weakened permission of $\varphi \vee \psi$ we find intuitive in the setting of permitted announcements. Unlike in the Ross Paradox [Ross, 1941], note that this is not a choice between two different announcements, but the single announcement of a disjunction.

Proposition 6.9 suggests the following translation $tr : \mathcal{L}_{popal} \longrightarrow \mathcal{L}_{poel}$:

Definition 6.12 (the translation *tr*) *We define $tr(\varphi)$ by induction on the complexity of φ as follows:*

- $tr(p) = p$
- $tr(\bot) = \bot$
- $tr(\neg\varphi) = \neg tr(\varphi)$
- $tr(\psi \vee \varphi) = tr(\psi) \vee tr(\varphi)$
- $tr(K_i\varphi) = K_i tr(\varphi)$
- $tr(P(\psi, \varphi)) = P(tr(\psi), tr(\varphi))$
- $tr(O(\psi, \varphi)) = O(tr(\psi), tr(\varphi))$
- $tr([\psi]p) = tr(\psi) \longrightarrow p$
- $tr([\psi]\bot) = \neg tr(\psi)$
- $tr([\psi]\neg\varphi) = tr(\psi) \longrightarrow \neg tr([\psi]\varphi)$
- $tr([\psi](\varphi_1 \vee \varphi_2)) = tr([\psi]\varphi_1) \vee tr([\psi]\varphi_2)$
- $tr([\psi]K_i\varphi) = tr(\psi) \longrightarrow K_i tr([\psi]\varphi)$
- $tr([\psi_1][\psi_2]\varphi) = tr([\langle\psi_1\rangle\psi_2]\varphi)$
- $tr([\psi_1]P(\psi_2, \varphi)) = tr(\psi_1) \longrightarrow P(tr(\langle\psi_1\rangle\psi_2), tr(\varphi))$
- $tr([\psi_1]O(\psi_2, \varphi)) = tr(\psi_1) \longrightarrow O(tr(\langle\psi_1\rangle\psi_2), tr(\varphi))$

An elementary proof by induction on the structure of φ, using Proposition 6.9, now delivers:

Proposition 6.13 *For all* $\varphi \in \mathcal{L}_{popal}$*,* $\models \varphi \longleftrightarrow tr(\varphi)$.

In other words, adding public announcements to a logical language with permitted and obligatory announcement does not increase the expressivity of the logic.

Finally, we need to show a property of the degree function: for all $\varphi \in \mathcal{L}_{popal}$, $\Pi(\varphi)$: $deg(tr(\varphi)) = deg(\varphi)$. This property will be used in the completeness proof. To prove it, we first introduce a preliminary lemma:

Lemma 6.14 *For all* $\psi, \psi' \in \mathcal{L}_{popal}$*,*

1. $deg(\psi) \leqslant deg([\psi]\psi')$ *and* $deg(tr(\psi)) \leqslant deg(tr([\psi]\psi'))$
2. $deg(\langle\psi\rangle\psi') = deg([\psi]\psi')$ *and* $deg(tr(\langle\psi\rangle\psi')) = deg(tr([\psi]\psi'))$

Proof

1. By a simple induction on the structure of ψ'.
2. Let us look at the second one, the first being similar and easier:
$deg(tr(\langle\psi\rangle\psi')) = deg(tr(\neg[\psi]\neg\psi')) = deg(\neg tr([\psi]\neg\psi'))$
$= deg(tr([\psi]\neg\psi')) = deg(tr(\psi) \longrightarrow \neg tr([\psi]\psi'))$
$= max(deg(tr(\psi)), deg(\neg tr([\psi]\psi'))) = max(deg(tr(\psi)), deg(tr([\psi]\psi')))$
$= deg(tr([\psi]\psi'))$ by 1.

$\square$

Lemma 6.15 *For all* $\theta \in \mathcal{L}_{popal}$*: for all* $\psi \in \mathcal{L}_{popal}$*, if* $\Pi(\psi)$ *then* $\Pi([\psi]\theta)$ *and* $\Pi(\theta)$.

Proof Let us prove it by induction on the structure of θ. We denote by $=^*$ the equalities that come from $\Pi(\psi)$.

Base cases

$-\theta = p$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$

$$\Pi([\psi]p) \begin{cases} deg(tr([\psi]p)) \\ = deg(tr(\psi) \longrightarrow p) \\ = max(deg(tr(\psi)), deg(p)) \\ = deg(tr(\psi)) \\ =^* deg(\psi) \\ = deg([\psi]p) \end{cases} \text{and } \Pi(p) \begin{cases} deg(tr(p)) \\ = deg(p) \end{cases}$$

$\theta = \bot$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$

$$\Pi([\psi]\bot) \begin{cases} deg(tr([\psi]\bot)) \\ = deg(\neg tr(\psi)) \\ = deg(tr(\psi)) \\ =^* deg(\psi) \\ = deg([\psi]\bot) \end{cases} \text{and } \Pi(\bot) \begin{cases} deg(tr(\bot)) \\ = deg(\bot) \end{cases}$$

Inductive cases: let us suppose that for all ψ, $\Pi(\psi)$ implies $\Pi([\psi]\chi)$ and $\Pi(\chi)$ for all χ that are subformula of θ. We denote by $=_{IH1}$ the equalities that come from $\Pi([\psi]\chi)$ and by $=_{IH2}$ the ones that come from $\Pi(\chi)$.

- $\theta = \neg\chi$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$

$$\Pi([\psi]\neg\chi) \begin{cases} deg(tr([\psi]\neg\chi)) \\ = deg(tr(\psi) \longrightarrow \neg tr([\psi]\chi)) \\ = max(deg(tr(\psi)), deg(tr([\psi]\chi))) \\ =^*_{IH1} max(deg(\psi), deg([\psi]\chi)) \\ = deg(\psi) + deg(\chi) \\ = deg([\psi]\neg\chi) \end{cases} \text{and } \Pi(\neg\chi) \begin{cases} deg(tr(\neg\chi)) \\ = deg(\neg tr(\chi)) \\ = deg(tr(\chi)) \\ =_{IH2} deg(\chi) \\ = deg(\neg\chi) \end{cases}$$

- $\theta = K_i\chi$: Identical

- $\theta = \chi_1 \vee \chi_2$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$

$$\begin{cases} deg(tr([\psi](\chi_1 \vee \chi_2))) \\ = deg(tr([\psi]\chi_1) \vee tr([\psi]\chi_2)) \\ = max(deg(tr([\psi]\chi_1)), deg(tr([\psi]\chi_2))) \\ =_{IH1} max(deg([\psi]\chi_1), deg([\psi]\chi_2)) \\ = deg(\psi) + max(deg(\chi_1), deg(\chi_2)) \\ = deg([\psi](\chi_1 \vee \chi_2)) \end{cases} \text{and} \begin{cases} deg(tr(\chi_1 \vee \chi_2)) \\ = deg(tr(\chi_1) \vee tr(\chi_2)) \\ =_{IH2} max(deg(\chi_1), deg(\chi_2)) \\ = deg(\chi_1 \vee \chi_2) \end{cases}$$

- $\theta = P(\psi', \chi)$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$. We denote by $=^@$ the equalities that comes from Lemma 6.14

$$\Pi([\psi]P(\psi', \chi)) \begin{cases} deg(tr([\psi]P(\psi', \chi))) \\ = deg(tr(\psi) \longrightarrow P(tr(\langle\psi\rangle\psi'), tr(\chi))) \\ = max(deg(tr(\psi)), deg(P(tr(\langle\psi\rangle\psi'), tr(\chi)))) \\ =^* max(deg(\psi), deg(tr(\langle\psi\rangle\psi')) + deg(tr(\chi)) + 1) \\ =^@_{IH1} max(deg(\psi), deg(\langle\psi\rangle\psi') + deg(tr(\chi)) + 1) \\ =^@ deg(\langle\psi\rangle\psi') + deg(tr(\chi)) + 1 \\ =^@_{IH2} deg(\psi) + deg(\psi') + deg(\chi) + 1 \\ = deg([\psi]P(\psi', \chi)) \end{cases}$$

$$\text{and } \Pi(P(\psi', \chi)) \begin{cases} deg(tr(P(\psi', \chi))) \\ = deg(P(tr(\psi'), tr(\chi))) \\ = deg(tr(\psi')) + deg(tr(\chi)) + 1 \\ =_{IH2} deg(\psi') + deg(\chi) + 1 \\ = deg(P(\psi', \chi)) \end{cases}$$

- $\theta = O(\psi', \chi)$: Identical

- $\theta = [\psi']\chi$: let $\psi \in \mathcal{L}_{popal}$ be such that $\Pi(\psi)$

$$\Pi([\psi][\psi']\chi)\begin{cases} deg(tr([\psi][\psi']\chi)) \\ = deg(tr([\langle\psi\rangle\psi']\chi)) \text{ by definition} \\ =_{IH1} deg([\langle\psi\rangle\psi']\chi) \\ \text{because } \Pi(\langle\psi\rangle\psi') \text{ (by IH1) and then } \Pi([\langle\psi\rangle\psi']\chi) \\ = deg(\langle\psi\rangle\psi') + \text{deg}(\chi) \\ = deg(\psi) + deg(\psi') + deg(\chi) \\ = deg([\psi][\psi']\chi) \end{cases}$$

$$\text{and } \Pi([\psi']\chi)\begin{cases} deg(tr([\psi']\chi)) \\ =_{IH1} deg([\psi']\chi) \text{ because } \Pi(\psi') \text{ by IH2} \end{cases}$$

□

Therefore we have:

Proposition 6.16 *For all $\varphi \in \mathcal{L}_{popal}$, $\Pi(\varphi) : deg(tr(\varphi)) = deg(\varphi)$.*

Proof Since $\Pi(\top)$, the previous proposition guarantees this one.

□

6.2 Axiomatization

We define the axiomatization $POPAL$ and prove its soundness and completeness. Let $POPAL$ be the least set of formulas in our language that contains the axiom schemata and is closed under the inference rules in Table 6.1. We write $\vdash_{POPAL} \varphi$ for $\varphi \in POPAL$.

The axiom of 'obligation and prohibition' and the inference rule of 'substitution' deserve some explanation. This last inference rule simply express the fact that the announcements of two equivalent formulas give the same result, and that if announcing φ gives more information than announcing φ', if the first is permitted then the second also is. This intuition has been explained in the introduction of this chapter. Axiom 'obligation and prohibition' is a quite complicated way to express the intuitive implication $O(\psi,\varphi) \longrightarrow P(\psi,\varphi)$: everything that is obligatory to be said is also permitted to be said. In fact, this last property is not valid in our models, because it may happen that the set $\mathcal{P}$ is empty. In this case, nothing is permitted (not even the fact to say nothing), and thus everything is obligatory. This axiom allows to consider such borderline cases. If we want to avoid them, we can consider the class of models in which for all ψ, $\psi \longrightarrow P(\psi,\top)$ is valid, that we call 'permissive models'. In these models 'obligation and prohibition' is equivalent to $O(\psi,\varphi) \longrightarrow P(\psi,\varphi)$. Moreover we have the following:

Remark 6.17 *Let $\mathcal{M}, s$ be a pointed model, and let $\chi \in \mathcal{L}_{popal}$. If $\mathcal{M}, s \models P(\psi,\chi) \vee \neg O(\psi,\chi)$ then for all $\varphi \in \mathcal{L}_{popal}$ we have $\mathcal{M}, s \models O(\psi,\varphi) \longrightarrow P(\psi,\varphi)$.*

all instances of propositional tautologies	
$K_i(\psi \longrightarrow \varphi) \longrightarrow (K_i\psi \longrightarrow K_i\varphi)$	distribution
$K_i\varphi \longrightarrow \varphi$	truth
$K_i\varphi \longrightarrow K_iK_i\varphi$	positive introspection
$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi$	negative introspection
$[\psi]p \leftrightarrow (\psi \longrightarrow p)$	atomic permanence
$[\psi]\bot \leftrightarrow \neg\psi$	ann. and false
$[\psi]\neg\varphi \leftrightarrow (\psi \longrightarrow \neg[\psi]\varphi)$	ann. and negation
$[\psi](\varphi_1 \vee \varphi_2) \longleftrightarrow ([\psi]\varphi_1 \vee [\psi]\varphi_2)$	ann. and disjunction
$[\psi]K_i\varphi \leftrightarrow (\psi \longrightarrow K_i[\psi]\varphi)$	ann. and knowledge
$[\psi_1][\psi_2]\varphi \leftrightarrow [\langle\psi\rangle_1\psi_2]\varphi$	ann. composition
$[\psi]P(\psi', \varphi) \leftrightarrow (\psi \longrightarrow P(\langle\psi\rangle\psi', \varphi))$	ann. and permission
$[\psi]O(\psi', \varphi) \leftrightarrow (\psi \longrightarrow O(\langle\psi\rangle\psi', \varphi))$	ann. and obligation
$P(\psi, \varphi) \longrightarrow \langle\psi\rangle\varphi$	permission and truth
$O(\top, \top)$	
$(O(\psi, \varphi_1) \wedge O(\psi, \varphi_2)) \leftrightarrow O(\psi, \varphi_1 \wedge \varphi_2)$	obligations composition
$(P(\psi, \varphi_1) \wedge O(\psi, \varphi_2)) \longrightarrow P(\psi, \varphi_1 \wedge \varphi_2)$	obligation and permission comp.
$(\psi \wedge O(\psi, \varphi) \wedge \neg P(\psi, \varphi)) \leftrightarrow (\psi \wedge \neg P(\psi, \top))$	obligation and prohibition
From φ and $\varphi \longrightarrow \psi$ infer ψ	modus ponens
From φ infer $K_i\varphi$	necessitation of K_i
From φ infer $[\psi]\varphi$	necessitation of announcement
From $(\psi \longleftrightarrow \psi') \wedge (\langle\psi\rangle\varphi \longrightarrow \langle\psi'\rangle\varphi')$ infer	
$(P(\psi, \varphi) \longrightarrow P(\psi', \varphi'))$ and $(O(\psi, \varphi) \longrightarrow O(\psi', \varphi'))$	substitution

Table 6.1: Axiomatization of $POPAL$

We define the consistency and the maximality of a set x of formulas as usual: x is $POPAL$-consistent iff for all nonnegative integers n and for all formulas $\varphi_1, \ldots, \varphi_n \in x$, $\neg(\varphi_1 \wedge \ldots \wedge \varphi_n) \notin POPAL$ whereas x is maximal iff for all formulas φ, $\varphi \in x$ or $\neg\varphi \in x$.

6.2.1 Soundness

Proposition 6.18 *$POPAL$ is sound on the class of all permission Kripke models.*

Proof By Propositions 6.9, 6.10 and 6.11.

□

Note that we have in particular that

Proposition 6.19 *For all $\varphi \in \mathcal{L}_{popal}, \vdash_{POPAL} \varphi \longleftrightarrow tr(\varphi)$.*

6.2.2 Completeness

To prove the completeness result, we use a classical method through the canonical model. To do so, we first define the following:

Definition 6.20 *Let S^c be the set of all $\vdash_{POPAL}$-maximal consistent sets, $x \in S^c$, and $\psi \in \mathcal{L}_{popal}$. We define $A^x_\psi := \{y \in S^c \mid \langle\psi\rangle\chi \in y$ for all $O(\psi, \chi) \in x\}$.*

How to interpret such a set A^x_ψ? A $\vdash$-maximal consistent set y is in A^x_ψ if it contains $\langle\psi\rangle\varphi$ for every announcement φ that is obligatory in x after the announcement of ψ.

Remark 6.21 *By Remark 6.17, we get that if $P(\psi, \top) \in x$ then :*

- *$x \in A^x_\psi$. Indeed we then have $O(\psi, \varphi) \in x$ only if $P(\psi, \varphi) \in x$ only if $\langle\psi\rangle\varphi \in x$.*
- *for all formula $\chi \in \mathcal{L}_{popal}$, $O(\psi, \chi) \longrightarrow P(\psi, \chi) \in x$.*

It is also the case if, for some formula $\mu \in \mathcal{L}_{popal}$, $P(\psi, \mu) \in x$ or $\neg O(\psi, \mu) \in x$.

We are now able to define the canonical model for $POPAL$:

Definition 6.22 (Canonical Model)
The canonical model $\mathcal{M}^c = (S^c, \sim^c_i, V^c, \mathcal{P}^c)$ is defined as follows:

- *S^c is the set of all $\vdash_{POPAL}$-maximal consistent sets*
- *for any $p \in PROP$, $V^c(p) = \{x \in S^c \mid p \in x\}$*
- *$x \sim^c_i y$ iff $K_i x = K_i y$, where $K_i x = \{\varphi | K_i\varphi \in x\}$*
- *$\mathcal{P}^c = \{(x, S', S'') : \exists P(\psi, \varphi) \in x|\ S' = \{y \in S^c : \psi \in y\}\ ,\ S'' = \{y \in S^c : \langle\psi\rangle\varphi \in y\} \cap A^x_\psi\} \bigcup \{(x, S', S'') : \exists(\psi \wedge \neg O(\psi, \varphi)) \in x|\ S' = \{y \in S^c : \psi \in y\}\ ,\ S'' = A^x_\psi\}$.*

The definition of $\mathcal{P}^c$ requires some explanation. The main idea is that we put in $\mathcal{P}^c$ the triplets (x, S', S'') such that S'' is as big as possible (*i.e.* corresponds to the least restriction). For that purpose, considering that S' corresponds to the announcement of ψ, two kinds of transitions (x, S', S'') (*i.e.* two kinds of announcements) are in $\mathcal{P}^c$:

1. if some $P(\psi, \varphi) \in x$, we take for S'' the expression of $\varphi \wedge \bigwedge \chi_i$ where the χ_i are such that $O(\psi, \chi_i) \in x$: as every restriction in $\mathcal{P}^c$ has to restrict more than every restriction corresponding to the χ_i, then S'' is the least restriction that insure that $P(\psi, \varphi) \in x$.
2. if some $\neg O(\psi, \varphi)) \in x$, we take for S'' the set A^x_ψ, which is the biggest set we can take to insure that the obligations in x are satisfied.

Proposition 6.23 *The canonical model is a model.*

Proof The set of states and the valuation are clearly well defined, and as the equality is an equivalence relation between sets of formulas, $\sim^c_i$ is an equivalence relation. $\mathcal{P}^c$ is a set of triplets of the expected form, the only thing we have to verify is that for every $(x, S', S'') \in \mathcal{P}^c$, we have $x \in S'' \subseteq S'$. Indeed, let $(x, S', S'') \in \mathcal{P}^c$, thus there are two possibilities:

1. There exists a $P(\psi,\varphi) \in x$ such that $S' = \{y \in S^c : \psi \in y\}$ and $S'' = \{y \in S^c : \langle\psi\rangle\varphi \in y\} \cap A^x_\psi$. In this case, clearly $S'' \subseteq S'$ because for all $y \in S^c$, $\langle\psi\rangle\varphi \in y$ only if $\psi \in y$. Now $x \in \{y \in S^c : \langle\psi\rangle\varphi \in y\}$ comes from the axiom 'permission and truth' considering that $P(\psi,\varphi) \in x$. It remains to show that $x \in A^x_\psi$, which is proved by Remark 6.21.

2. There exists a $(\psi \wedge \neg O(\psi,\varphi)) \in x$ such that $S' = \{y \in S^c : \psi \in y\}$ and $S'' = \{y \in S^c : \forall O(\psi,\chi) \in x, \langle\psi\rangle\chi \in y\}$. In this case, $S'' \subseteq S'$ comes from the fact that $[\psi]O(\top,\top) \in x$, and thus $\psi \in x$ implies that $O(\psi,\top) \in x$. Therefore if $y \in S''$ then $\langle\psi\rangle\top \in y$, which means that $y \in S'$. Now to show that $x \in S''$ let us consider $O(\psi,\chi) \in x$ and let us show that $\langle\psi\rangle\chi \in x$. Indeed, by Remark 6.21, $\neg O(\psi,\varphi) \in x$ implies $O(\psi,\chi) \longrightarrow P(\psi,\chi) \in x$, with $O(\psi,\chi) \in x$ we get $P(\psi,\chi) \in x$ by 'modus ponens', and finally $\langle\psi\rangle\chi \in x$ by 'permission and truth'.

□

In the canonical model, a state is a set of formulas. The link between the fact that a formula φ is in a set x and the fact that $\mathcal{M}^c, x \models \varphi$ is given by the Truth Lemma. In the proof of the Truth Lemma, we need the following

Lemma 6.24 *For any $x \in \mathcal{M}^c$ and any $\psi,\varphi,\alpha,\beta \in \mathcal{L}_{poel}$,*

1. *if $A^x_\psi \subseteq \{y : \langle\psi\rangle\varphi \in y\}$, then $O(\psi,\varphi) \in x$,*

2. *if $P(\alpha,\beta) \in x$ and $\{y : \langle\alpha\rangle\beta \in y\} \cap A^x_\alpha \subseteq \{y : \langle\alpha\rangle\varphi \in y\}$, then $P(\alpha,\varphi) \in x$*

Proof

1. By hypothesis, any maximal consistent set that contains $\langle\psi\rangle\chi$ for all $O(\psi,\chi) \in x$ contains also $\langle\psi\rangle\varphi$, thus $\{\langle\psi\rangle\chi : O(\psi,\chi) \in x\} \cup \{[\psi]\neg\varphi\}$ is inconsistent. By definition, it has a finite subset $\{\langle\psi\rangle\chi_1, \ldots, \langle\psi\rangle\chi_n, [\psi]\neg\varphi\}$ that is inconsistent. Thus $\vdash \langle\psi\rangle\chi_1 \wedge \ldots \wedge \langle\psi\rangle\chi_n \longrightarrow \langle\psi\rangle\varphi$, i.e. $\vdash \langle\psi\rangle \bigwedge \chi_i \longrightarrow \langle\psi\rangle\varphi$ and then $\vdash O(\psi, \bigwedge \chi_i) \longrightarrow O(\psi,\varphi)$ by the inference rule of substitution. By axiom 'obligations composition' $O(\psi, \bigwedge \chi_i) \in x$, and by 'modus ponens' $O(\psi,\varphi) \in x$.

2. By hypothesis, any maximal consistent set that contains $\langle\alpha\rangle\beta$ and $\langle\alpha\rangle\chi$ for all $O(\alpha,\chi) \in x$ contains also $\langle\alpha\rangle\varphi$. Thus $\{\langle\alpha\rangle\beta\} \cup \{\langle\alpha\rangle\chi : O(\alpha,\chi) \in x\} \cup \{[\alpha]\neg\varphi\}$ is inconsistent. By definition, this set has a finite subset $\{\langle\alpha\rangle\beta, \langle\alpha\rangle\chi_1, \ldots, \langle\alpha\rangle\chi_n, \alpha\neg\varphi\}$ that is inconsistent. Thus $\vdash (\langle\alpha\rangle\beta \wedge \langle\alpha\rangle\chi_1 \wedge \ldots \wedge \langle\alpha\rangle\chi_n) \longrightarrow \langle\alpha\rangle\varphi$, i.e. $\vdash \langle\alpha\rangle(\beta \wedge \bigwedge \chi_i) \longrightarrow \langle\alpha\rangle\varphi$ and then $\vdash P(\alpha, \beta \wedge \bigwedge \chi_i) \longrightarrow P(\alpha,\varphi)$. $O(\alpha, \bigwedge \chi_i) \in x$ is true by axiom 'obligation composition' and $P(\alpha,\beta) \in x$ by hypothesis. Thus $P(\alpha, \beta \wedge \bigwedge \chi_i) \in x$ is true by axiom 'obligation and permission comp.'. Finally, $P(\alpha,\varphi) \in x$ by 'modus ponens'.

□

Proposition 6.25 (Truth Lemma for $\mathcal{L}_{poel}$) *For all $\varphi \in \mathcal{L}_{poel}$ we have:*

$$\Pi(\varphi): \text{ for all } x \in S^c, \mathcal{M}^c, x \models \varphi \text{ iff } \varphi \in x$$

Proof The proof is by induction on the degree of φ.

Base case If $deg(\varphi) = 0$ then $\varphi \in \mathcal{L}_{el}$ and $\Pi(\varphi)$ is a known result (See Proposition 4.1 or [Blackburn et al., 2001] or [Fagin et al., 1995] for details). Note that $(S^c, \sim_i^c, V^c)$ is the classical canonical model for $\mathcal{L}_{el}$.

Induction steps Let $k \in \mathbb{N}$, let us suppose that $\Pi(\psi)$ is true for all $\psi \in \mathcal{L}_{poel}$ such that $deg(\psi) \leqslant k$.

Note that it follows that $\Pi(\psi)$ is true for all $\psi \in \mathcal{L}_{popal}$ such that $deg(\psi) \leqslant k$. Indeed, for all such ψ, for all $x \in S^c$, $\mathcal{M}^c, x \models \psi$ iff $\mathcal{M}^c, x \models tr(\psi)$ iff $tr(\psi) \in x$ iff $\psi \in x$.

Let φ be such that $deg(\varphi) \leq k + 1$ and let us reason by induction on the structure of φ.

- $\varphi = p; \bot; \neg\psi; \varphi_1 \vee \varphi_2; K_i\psi$: See the proof of the truth lemma for $\mathcal{L}_{el}$ in [Blackburn et al., 2001] or [Fagin et al., 1995].
- $\varphi = P(\psi, \chi)$:
  ($\Rightarrow$) Suppose that $\mathcal{M}^c, x \models P(\psi, \chi)$. This implies that $\mathcal{M}^c, x \models \psi$ (and thus $\psi \in x$ by IH) and that there exists $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^c}$ such that $(x, [\![\psi]\!]_{\mathcal{M}^c}, S'') \in \mathcal{P}^c$.
  Now, there are two possibilities:
  - There exists $P(\alpha, \beta) \in x$ s.t. $(*)[\![\psi]\!]_{\mathcal{M}^c} = \{y \in S^c : \alpha \in y\}$ and $S'' = \{y \in S^c : \langle\alpha\rangle\beta \in y\} \cap A^x_\alpha$. In this case $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^c}$ by hypothesis, and $[\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^c} = [\![\langle\alpha\rangle\chi]\!]_{\mathcal{M}^c}$ by $(*)$. Thus $S'' \subseteq [\![\langle\alpha\rangle\chi]\!]_{\mathcal{M}^c}$. By lemma 6.24.2, with $P(\alpha, \beta) \in x$, this implies that $P(\alpha, \chi) \in x$ by substitution rule. By $(*)$ again we obtain that $P(\psi, \chi) \in x$.
  - There exists $\neg O(\alpha, \beta) \in x$ s.t. $[\![\psi]\!]_{\mathcal{M}^c} = \{y \in S^c : \alpha \in y\}$ and $S'' = A^x_\alpha$. On one hand, this implies that $\vdash \psi \longleftrightarrow \alpha$ and then $\neg O(\psi, \beta) \in x$ by substitution rule. On the other hand, with the fact that $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^c}$ we obtain, by lemma 6.24.1, that $O(\psi, \chi) \in x$. Now, if we suppose $P(\psi, \chi) \notin x$ then $\neg P(\psi, \top) \in x$ by Remark 6.21. Therefore, using the fact that $\psi \in x$,we obtain $O(\psi, \beta) \in x$ by the same Remark. This leads to a contradiction which proves that $P(\psi, \chi) \in x$.

  ($\Leftarrow$) If $P(\psi, \chi) \in x$ then let us define $S' = [\![\psi]\!]_{\mathcal{M}^c}$ and $S'' = [\![\langle\psi\rangle\chi]\!] \cap A^x_\psi$. We obtain $(x, S', S'') \in \mathcal{P}^c$ by definition of $\mathcal{P}^c$. Therefore, as $S'' \subseteq [\![\langle\psi\rangle\chi]\!] = [\![tr(\langle\psi\rangle\chi)]\!]$, $\mathcal{M}_c, x \models P(\psi, \chi)$.
- $\varphi = O(\psi, \chi)$:
  ($\Leftarrow$) Suppose that $O(\psi, \chi) \in x$ and $\mathcal{M}_c, x \not\models O(\psi, \chi)$. Thus $\mathcal{M}_c, x \models \psi$ otherwise we would have $\mathcal{M}_c, x \models O(\psi, \chi)$. Now $\mathcal{M}_c, x \not\models O(\psi, \chi)$ implies that there exists

$(x, [\![\psi]\!], S'') \in \mathcal{P}^c$ such that $S'' \not\subseteq [\![\langle\psi\rangle\chi]\!]$. That is impossible, because by definition $S'' \subseteq A^x_\psi \subseteq [\![\langle\psi\rangle\chi]\!]$.

($\Rightarrow$) Suppose that $O(\psi, \chi) \notin x$ and $\mathcal{M}_c, x \models O(\psi, \chi)$. Then $\neg O(\psi, \chi) \in x$ and, by definition of $\mathcal{P}^c$, $(x, [\![\psi]\!], A^x_\psi) \in \mathcal{P}^c$. But then $\mathcal{M}_c, x \models O(\psi, \chi)$ leads to $A^x_\psi \subseteq [\![\langle\psi\rangle\chi]\!]$ by the semantics. Now, if we prove that $[\![\langle\psi\rangle\chi]\!] = \{y : \langle\psi\rangle\chi \in y\}$ we obtain $O(\psi, \chi) \in x$ by Lemma 6.24.1. It would lead to a contradiction and then to the wanted result. Now $[\![\langle\psi\rangle\chi]\!] = [\![tr(\langle\psi\rangle\chi)]\!]$ by Proposition 6.13. $[\![tr(\langle\psi\rangle\chi)]\!] = \{y : tr(\langle\psi\rangle\chi) \in y\}$ by IH (because $tr(\langle\psi\rangle\chi) \in \mathcal{L}_{poel}$ and $deg(tr(\langle\psi\rangle\chi)) \leq k$). Finally, $\{y : tr(\langle\psi\rangle\chi) \in y\} = \{y : \langle\psi\rangle\chi \in y\}$ by Proposition 6.19.

□

Proposition 6.26 *POPAL is sound and complete with respect to the class of all permission Kripke models.*

Proof The soundness has been shown in Proposition 6.18. By Proposition 6.25 we can show the completeness with respect to the class of all permission Kripke models. Indeed, for all $\varphi \in \mathcal{L}_{popal}$:

$$\models \varphi \Rightarrow \models tr(\varphi) \Rightarrow \mathcal{M}^c \models tr(\varphi) \Rightarrow \vdash tr(\varphi) \Rightarrow \vdash \varphi.$$

□

6.3 Decidability

We prove in this section that *POPAL* is decidable by proving a *small model property*. To do so, we use a filtration method, extending the notion of filtration introduced in Definition 2.16. We first introduce to notions that are useful in this method:

Definition 6.27 (Closed set) *Let $X \subseteq \mathcal{L}_{poel}$. We shall say that X is* closed *if the following properties are satisfied:*

- *X is closed under subformulas*
- *for all $P(\psi, \varphi) \in X$, $tr(\langle\psi\rangle\varphi) \in X$*
- *for all $O(\psi, \varphi) \in X$, $tr(\langle\psi\rangle\varphi) \in X$*

Definition 6.28 *Let $\mathcal{M} = (S, \sim_i, V, \mathcal{P})$ be a model and Γ be a closed set of formulas. Let $\leftrightsquigarrow_\Gamma$ be the relation on S defined, for all $s, t \in S$, by:*

$$s \leftrightsquigarrow_\Gamma t \quad \textit{iff} \quad \forall\, \varphi \in \Gamma : (\mathcal{M}, s \models \varphi \textit{ iff } \mathcal{M}, t \models \varphi)$$

Note that $\leftrightsquigarrow_\Gamma$ is an equivalence relation. For all $s \in S$, let us denote by $|s|_\Gamma$ (or simply $|s|$) the equivalence class of s with respect to $\leftrightsquigarrow_\Gamma$; and for $S' \subseteq S''$, we write $\leftrightsquigarrow_\Gamma(S')$ for $\{t \in S \mid \exists s \in S' : s \leftrightsquigarrow_\Gamma t\}$.

We can now generalize the notion of filtration in the context of permission models:

Definition 6.29 (Filtration) *We call* the filtration of $\mathcal{M}$ through Γ *(or simply* the filtration of $\mathcal{M}$*) the model* $\mathcal{M}^\Gamma = (S^\Gamma, \sim_i^\Gamma, V^\Gamma, \mathcal{P}^\Gamma)$ *where:*

- $S^\Gamma = S/_{\leftrightsquigarrow_\Gamma}$
- $|s| \sim_i^\Gamma |t|$ *iff for all* $K_i\varphi \in \Gamma$, $(\mathcal{M}, s \models K_i\varphi$ *iff* $\mathcal{M}, t \models K_i\varphi)$
- $V^\Gamma(p) = \begin{cases} \emptyset \ if \ p \notin \Gamma \\ V(p)/_{\leftrightsquigarrow_\Gamma} \ if \ p \in \Gamma) \end{cases}$
- $\mathcal{P}^\Gamma = \{(|s|, S^1, S^2)$*: there exists* $t \in |s|$ *and* $S'' \subseteq S$ *s.t.* $S''/_{\leftrightsquigarrow_\Gamma} = S^2$ *and* $(t, \bigcup(S^1), S'') \in \mathcal{P}\}$

In this definition, S^1 is a set of equivalence classes, and $\bigcup S^1$ is the set of all states that are represented by an element of S^1. Here is a useful lemma:

Lemma 6.30 *Let* $\Gamma \subset \mathcal{L}_{poel}$ *be a finite closed set. For any model* $\mathcal{M}$*, its filtration* $\mathcal{M}^\Gamma$ *contains at most* 2^m *nodes, where* $m = Card(\Gamma)$.

Proof Let $\mathcal{M}$ be a model. Let $g : S^\Gamma \longrightarrow 2^\Gamma$ be defined by $g(|s|) = \{\psi \in \Gamma : \mathcal{M}, s \models \psi\}$. It follows from the definition of $\leftrightsquigarrow_\Gamma$ that g is well-defined and injective. Thus the size of S^Γ is at most 2^m.

□

The epistemic relations of a model and their filtrations over a set Γ are linked by the following property:

Proposition 6.31 *Let* $\mathcal{M}$ *be a model and let* Γ *be a closed set of formulas. Then for all* $s, t \in S$*, for all* $\varphi \in \Gamma$*:*

1. $s \sim_i t \Rightarrow |s| \sim_i^\Gamma |t|$.
2. $|s| \sim_i^\Gamma |t|$ *and* $K_i\varphi \in \Gamma$ *and* $\mathcal{M}, s \models K_i\varphi \Rightarrow \mathcal{M}, t \models \varphi$.

Proof

1. *Let* $s, t \in S$ *such that* $s \sim_i t$*, and let* $K_i\varphi \in \Gamma$*. Then we have* $\mathcal{M}, s \models K_i\varphi$ *iff for all* $u \sim_i s, \mathcal{M}, u \models \varphi$ *iff for all* $u \sim_i t$*,* $\mathcal{M}, u \models \varphi$ *iff* $\mathcal{M}, t \models K_i\varphi$*. Then by definition of* $\sim_i^\Gamma$ *we obtain* $|s| \sim_i^\Gamma |t|$.
2. *Let us suppose the first part of the implication. Since* $|s| \sim_i^\Gamma |t|$*,* $K_i\varphi \in \Gamma$ *and* $\mathcal{M}, s \models K_i\varphi$ *then* $\mathcal{M}, t \models K_i\varphi$*. Since* $\sim_i$ *is reflexive* $\mathcal{M}, t \models \varphi$.

$\square$

Proposition 6.31 is sufficient to prove the following:

Proposition 6.32 (Filtration lemma) *Let $\mathcal{M}$ be a model and let Γ be a closed set of formulas. For all $\varphi \in \Gamma$ we have:*

$$(F^o\varphi) \quad \forall s \in S, (\mathcal{M}, s \models \varphi \textit{ iff } \mathcal{M}^\Gamma, |s| \models \varphi).$$

Proof By induction on the degree of φ.

base case If $deg(\varphi) = 0$ then $\varphi \in \mathcal{L}_{el}$ and the proof of $(F^o\varphi)$ is done by induction on the complexity of φ (see [Blackburn et al., 2001] or [Fagin et al., 1995] for details, note that Γ is in particular closed under subformulas).

induction steps Let $k \in \mathbb{N}$. Suppose that $(F^o\psi)$ is true for all $\psi \in \mathcal{L}_{poel}$ such that $deg(\psi) \leqslant k$. Let φ be such that $deg(\varphi) \leq k+1$ and let us reason on the structure of φ.

- $\varphi = p; \bot; \neg\psi; \varphi_1 \vee \varphi_2, K_i\varphi$**:** See the proof of the filtration lemma in [Blackburn et al., 2001] or [Fagin et al., 1995].

- $\varphi = P(\psi, \chi)$**:** Let $s \in S$. By construction of Γ we know that
$$\begin{cases} \leftrightsquigarrow_\Gamma([\![\psi]\!]_\mathcal{M}) = [\![\psi]\!]_\mathcal{M} \text{ and} \\ (*) \leftrightsquigarrow_\Gamma([\![\langle\psi\rangle\chi]\!]_\mathcal{M}) = \leftrightsquigarrow_\Gamma([\![tr(\langle\psi\rangle\chi)]\!]_\mathcal{M}) = [\![tr(\langle\psi\rangle\chi)]\!]_\mathcal{M} = [\![\langle\psi\rangle\chi]\!]_\mathcal{M}. \end{cases}$$
($\Rightarrow$) Suppose $\mathcal{M}, s \models P(\psi, \chi)$. Let $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_\mathcal{M} = [\![tr(\langle\psi\rangle\chi)]\!]_\mathcal{M}$ be such that $(s, [\![\psi]\!], S'') \in \mathcal{P}$, and let $S^{oo} = S''/_{\leftrightsquigarrow_\Gamma}$. We have (by IH) that $S^{oo} \subseteq [\![tr(\langle\psi\rangle\chi)]\!]_{\mathcal{M}^\Gamma}$ and we obtain that $(|s|, [\![\psi]\!]_{\mathcal{M}^\Gamma}, S^{oo}) \in \mathcal{P}^\Gamma$ by definition of the filtration and $(*)$. Finally, $\mathcal{M}^\Gamma, |s| \models P(\psi, \chi)$

  ($\Leftarrow$) Suppose $\mathcal{M}^\Gamma, |s| \models P(\psi, \chi)$. Let $S^{oo} \subseteq [\![tr(\langle\psi\rangle\chi)]\!]_{\mathcal{M}^\Gamma}$ be such that $(|s|, [\![\psi]\!]_{\mathcal{M}^\Gamma}, S^{oo}) \in \mathcal{P}^\Gamma$. Then by definition of $\mathcal{P}^\Gamma$, there exists $t \in |s|$ and S'' such that $S''/_{\leftrightsquigarrow_\Gamma} = S^{oo}$ and $(t, [\![\psi]\!], S'') \in \mathcal{P}$. By IH, $S^{oo} \subseteq [\![tr(\langle\psi\rangle\chi)]\!]_{\mathcal{M}^\Gamma}$ implies that $S'' \subseteq [\![tr(\langle\psi\rangle\chi)]\!]_\mathcal{M}$. Therefore, $\mathcal{M}, t \models P(\psi, \chi)$. Finally, as $s \leftrightsquigarrow_\Gamma t$, $\mathcal{M}, s \models P(\psi, \chi)$.

- $\varphi = O(\psi, \chi)$**:** Let $s \in S$,

  ($\Rightarrow$) Suppose $\mathcal{M}, s \models O(\psi, \chi)$ and let S^{oo} be such that $(|s|, [\![\psi]\!]_{\mathcal{M}^\Gamma}, S^{oo}) \in \mathcal{P}^\Gamma$, we want to show that $S^{oo} \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^\Gamma}$. By definition of the filtration, we can construct S'' such that $S''/_{\leftrightsquigarrow_\Gamma} = S^{oo}$ and $(t, [\![\psi]\!], S'') \in \mathcal{P}$ for some $t \in |s|$. Thus $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_\mathcal{M}$, because $\mathcal{M}, t \models O(\psi, \chi)$ (as $s \leftrightsquigarrow_\Gamma t$). Finally $S^{oo} \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^\Gamma}$ by $(*)$ and IH.

  ($\Leftarrow$) Suppose $\mathcal{M}^\Gamma, |s| \models O(\psi, \chi)$ and let S'' be such that $(s, [\![\psi]\!], S'') \in \mathcal{P}$. We show that $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_\mathcal{M}$. Let $S^{oo} = S''/_{\leftrightsquigarrow_\Gamma}$, then by definition of the filtration, $(|s|, [\![\psi]\!]_{\mathcal{M}^\Gamma}, S^{oo}) \in \mathcal{P}^\Gamma$. Thus $S^{oo} \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^\Gamma}$ and then $S'' \subseteq [\![\langle\psi\rangle\chi]\!]_{\mathcal{M}^\Gamma}$ by $(*)$ and IH.

□

Definition 6.33 (Closure) *For all $\varphi \in \mathcal{L}_{poel}$, we construct the P-Closure of φ, noted $Cl(\varphi)$, inductively on the structure of φ:*

- $Cl(p) = \{p\}$
- $Cl(\bot) = \{\bot\}$
- $Cl(\neg\varphi) = \{\neg\varphi\} \cup Cl(\varphi)$
- $Cl(\psi \vee \varphi) = \{\psi \vee \varphi\} \cup Cl(\psi) \cup Cl(\varphi)$
- $Cl(K_i\varphi) = \{K_i\varphi\} \cup Cl(\varphi)$
- $Cl(P(\psi, \varphi)) = \{P(\psi, \varphi)\} \cup Cl(\psi) \cup Cl(\varphi) \cup Cl(tr(\langle\psi\rangle\varphi))$.
- $Cl(O(\psi, \varphi)) = \{O(\psi, \varphi)\} \cup Cl(\psi) \cup Cl(\varphi) \cup Cl(tr(\langle\psi\rangle\varphi))$.

Proposition 6.34 *For all $\varphi \in \mathcal{L}_{poel}, Cl(\varphi)$ is well-defined and it is a finite closed set.*

Proof The proof is by induction on the degree of φ.

base case If $deg(\varphi) = 0$ then $\varphi \in \mathcal{L}_{el}$ and we only need to prove that $Cl(\varphi)$ is a well-defined finite set closed under subformulas, which is straightforward.

inductive cases Let $k \in \mathbb{N}$, let us suppose that $Cl(\psi)$ is a well-defined finite closed set for any ψ such that $deg(\psi) \leq k$. Let φ be such that $deg(\varphi) \leq k + 1$ and let us reason inductively on the structure of φ.

- $\varphi = p; \bot; \neg\psi; \varphi_1 \vee \varphi_2; K_i\psi$: Trivial.
- $\varphi = P(\psi, \chi)$ or $O(\psi, \chi)$: By IH, $Cl(\psi), Cl(\chi)$ and $Cl(tr(\langle\psi\rangle\chi))$ are well-defined finite closed sets, so $Cl(P(\psi, \chi))$ and $Cl(O(\psi, \chi))$ are well-defined finite sets. We only need to prove that they are closed, which is straightforward.

□

Proposition 6.35 (Finite model property)

Let $\varphi \in \mathcal{L}_{poel}$, if φ is satisfiable then φ is satisfiable in a model containing at most 2^m nodes, where $m = Card(Cl(\varphi))$.

Proof Suppose that $\mathcal{M}$ and s are such that $\mathcal{M}, s \models \varphi$. Let $\Gamma = Cl(\varphi)$. Then by Proposition 6.32, $\mathcal{M}^\Gamma, |s| \models \varphi$. By Lemma 6.30, $\mathcal{M}^\Gamma$ contains at most 2^m states.

□

Theorem 6.36 *POPAL is decidable.*

Proof Let $\varphi \in \mathcal{L}_{popal}$ be a formula. The following procedure decides whether φ is satisfiable or not:

1. Compute $\Phi = tr(\varphi)$
2. Compute $\Gamma = Cl(\Phi)$
3. For all permission Kripke models $\mathcal{M}$ of size $\leq 2^{Card(\Gamma)}$ check if there exists $s \in \mathcal{M}$ such that $\mathcal{M}, s \models \Phi$.

□

6.4 Extended Example: La Belote

We now consider the French card game "la Belote". For a full description of the game, see `http://en.wikipedia.org/wiki/Belote`. The game is played with four players, who form two teams, and with 32 cards of a regular full deck of cards (the ranks 2 to 6 are eliminated). The name of the game, "belote", is also used in the game to designate a pair of a King and a Queen of a trump suit.

After the deal, and after the choice of a trump suit, the first player chooses and plays a card of her hand, followed by the other players in clockwise order. The player who played the highest trump card or the highest of the same color as the first player's card wins the round and starts the next round. Except for the first player of a round, each player has to follow suit or, if she cannot, to play trump. Moreover, when a trump has been played, it is forbidden to play a lower trump.

The act of playing a card can be seen as the public announcement that the corresponding card belonged to the corresponding player. We model the game with the set of propositional atoms $PROP$ expressing card ownership, namely $PROP = \{RC_i \mid R \in \{7, 8, 9, 10, J, Q, K, A\}, C \in \{\clubsuit, \heartsuit, \diamondsuit, \spadesuit\}, i \in \{A, B, C, D\}\}$. An atom RC_i stands for 'player i holds the card with rank R of suit C'. For any suit C and player i, we introduce the abbreviations $C_i = \bigvee_R RC_i$, and $C_i^{>R} = \bigvee_{R'>R} R'C_i$.

A model $\mathcal{M} = (S, \{\sim_i\}, V, \mathcal{P})$ is called a "model of La Belote" if

1. for each state s, for any R and C, there is exactly one i such that $s \in V(RC_i)$ (i.e. the states of $\mathcal{M}$ are deals of cards);

2. for any $s, t \in S$ and any i, $s \sim_i t$ implies that for all R, C: $s \in V(RC_i)$ iff $t \in V(RC_i)$ (i.e. each player can distinguish different cards);

3. $\mathcal{P}$ is constructed from $(S, \{\sim_i\}, V)$ according to the rules of the game.

The last item means that in a given deal s, for all the cards p held by an agent i that are permitted by the rules to be played, $(s, S, S_{p_i}) \in \mathcal{P}$. If after p has been played by player i it is permitted for player j to play q, then we also need that $(s, S_{p_i}, S_{\langle p_i \rangle q_j}) \in \mathcal{P}$. And so on, for all possible moves.

Let $\mathcal{M}$ be a model of La Belote. The trump suit has been selected before the game starts, we will suppose that it is clubs. The set of atoms is partially ordered as follows ($*$ can be one of the players A, B, C, D). First, any trump is higher than any non-trump. But the cards are also ordered in the following way: for non-trumps (i.e. for any $C \neq \clubsuit$):

$$7C_* < 8C_* < 9C_* < JC_* < QC_* < KC_* < 10C_* < AC_*$$

For trumps:

$$7\clubsuit_* < 8\clubsuit_* < Q\clubsuit_* < K\clubsuit_* < 10\clubsuit_* < A\clubsuit_* < 9\clubsuit_* < J\clubsuit_*$$

For more details, see the mentioned website.

We now list a number of model validities of La Belote. These formulas are valid at the beginning of each round of the game, in other words, the models $\mathcal{M}$ considered below result from any iteration of a sequence of four permitted announcements. We will call 1 the player that opens the round, followed by 2, etc.

1-One player at once:

For all $\psi \in \mathcal{L}_{popal}$, all $i \neq j$, all $p_i, q_j \in PROP$, $\mathcal{M} \models P(\psi, p_i) \longrightarrow \neg P(\psi, q_j)$.

Two different players are not allowed to play simultaneously.

2-Each card is played once:

For all $p \in PROP$, all $\psi \in \mathcal{L}_{popal}$, $\mathcal{M} \models \neg P(\langle p \rangle \psi, p)$.

If a card has been played once, it cannot be played again.

3-Obligation to follow suit:

For all ranks R and all suits C,

$\mathcal{M} \models C_2 \longrightarrow O(RC_1, C_2)$.

If the player 2 can follow the suit asked by 1, she is obliged to do so.

4-Obligation to play trump:

For all ranks R, all suit $C \neq \clubsuit$, $\mathcal{M} \models \clubsuit_2 \longrightarrow O(RC_1, C_2 \vee \clubsuit_2)$.

If the player 2 can follow the suit asked by 1 or play trump, she is obliged to do so.

5-Permission to say "belote et rebelote":

For all players i, $\mathcal{M} \models K\clubsuit_i \wedge Q\clubsuit_i \wedge (P(\psi, Q\clubsuit_i) \vee P(\psi, K\clubsuit_i)) \longrightarrow P(\psi, Q\clubsuit_i \wedge K\clubsuit_i)$.

If the player one is allowed to play the queen of the trump suit, she is allowed to announce that she has the royal couple (called the "belote"). This does not mean that she is allowed to play both cards, but playing one of them she is allowed to announce that she also has the other one.

6-Obligation to go up at trump:

For all $\psi \in \mathcal{L}_{popal}$, all player i and all R, $\mathcal{M} \models \clubsuit_i^{>R} \longrightarrow O(\langle\psi\rangle R\clubsuit_{i-1}, \clubsuit_i^{>R})$.
This says that if the previous player played trump and if you have higher cards than the played trump, then you are obliged to play one of them.

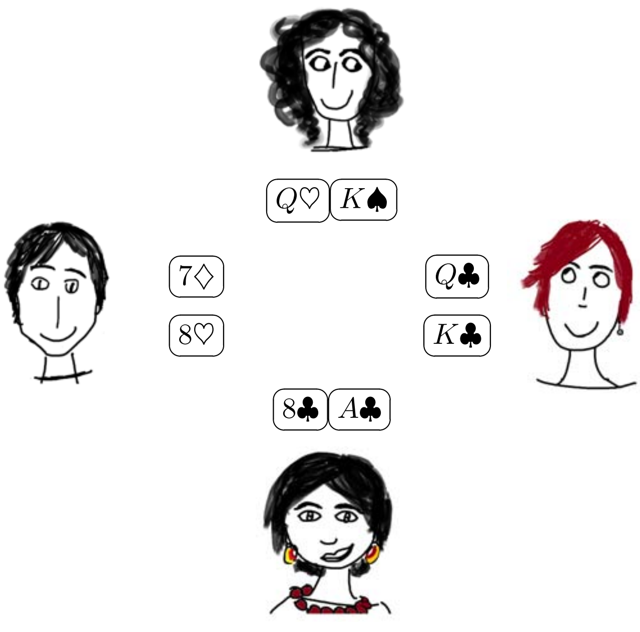


Figure 6.1: The moody children playing *la belote*

We apply these conditional rules about the permission to speak to the state (deal) s, presented in Figure 6.1, in which each player has 2 cards. Alex starts the game. According to the rule, our model validates the following formulas:

- $\mathcal{M}, s \models P(8\heartsuit_A) \wedge P(7\diamondsuit_A) \wedge \neg P(8\heartsuit_A \wedge 7\diamondsuit_A)$:

  Alex has the permission to play one of his cards, but not both.

- $\mathcal{M}, s \models O(8\heartsuit_A, Q\heartsuit_B)$:

  If Alex plays the $8\heartsuit$ card, Brune is obliged to play a card of the same suit, thus she cannot play her $K\spadesuit$ card (rule (3)).

- $\mathcal{M}, s \models P(\langle 8\heartsuit_A\rangle Q\heartsuit_B, Q\clubsuit_C \wedge K\clubsuit_C)$:

  Cha has the permission to announce that she has both cards of the "belote" (rule (5)).

- $\mathcal{M}, s \models O(\langle\langle 8\heartsuit_A\rangle Q\heartsuit_B\rangle Q\clubsuit_C, A\clubsuit_D)$:

  The 'go up at trump"applies if Cha plays the queen of clubs. As Dan has a unique higher trump, she has the obligation to play it.

6.5 Comparison to the Literature

6.5.1 Classic Deontic Principles and Paradoxes

As we reviewed before, deontic logic started out with Von Wright's operators P and O binding formulas in expressions $P\varphi$ and $O\varphi$, then came Meyer's and Van der Meyden's mind-frame switch to operators P and O binding actions, and finally we treated communicative actions that are represented by the announced formulas. Recall that in our framework we treat the obligation and permission to speak φ (as $P\varphi$ and $O\varphi$, using the abbreviation) and not the obligation and permission that φ. Well, if we end up with such expressions, how do its validities relate to the standard and historical Von Wright approach? In this subsection, we summarily treat that matter.

First, a disappointment: the P and O operators we have introduced are *not normal modal operators* (the triples in the $\mathcal{P}$ relation rather suggest a modality with a neighbourhood-type of semantics). They do not satisfy necessitation! A formula may be valid, but that does not make it an obligation, or permitted; if you are not allowed to announce p nor $\neg p$, it does not help you a great deal that $p \vee \neg p$ is a validity!

Something else has to be underlined once again: our formalism allows to consider situations in which nothing is permitted to be said, which is equivalent to the fact that everything is obligatory. To avoid such borderline cases, we will often consider the class of models in which for all ψ, $\psi \longrightarrow P(\psi, \top)$ is valid, called the 'permissive models'. Our comments on classical deontic paradoxes is in this context of permissive models.

Obligation distributes over conjunction (and implication), as $O(\varphi \wedge \psi) \longleftrightarrow (O\varphi \wedge O\psi)$ is a special case (in the case where the first argument is $\top$) of Proposition 6.11.1. Permission does not distribute over conjunction: we may have that p and q are both permissible announcements, such that Pp and Pq are true, but not at the same time, $P(p \wedge q)$ may be false.

This reflects that for a given Kripke model with domain S and actual state s the relation $\mathcal{P}$ may contain $(s, S, [\![p]\!])$ and $(s, S, [\![q]\!])$ but not $(s, S, [\![p \wedge q]\!])$. However, given weakening of permitted announcements, a valid principle indeed is $P(\varphi \wedge \psi) \longrightarrow (P\varphi \wedge P\psi)$.

Permitted announcements are true, obligatory ones also in the permissive models: $P\varphi \longrightarrow \varphi$ and $O\varphi \longrightarrow \varphi$, a principle obviously false in classic deontic logic. But one has to realize the special reading of such implications in our setting! $P\varphi \longrightarrow \varphi$ is valid because a precondition for a permitted announcement is the truth of the announcement formula. It does not formalize that all permitted actions always take place. A similar slip of the deontic mind occurs when observing that $P\varphi \longrightarrow P(\varphi \vee \psi)$ is valid. Doesn't this conflict with Ross's Paradox [Ross, 1941]? We addressed this matter in the introduction, let us go over the details. Ross's Paradox is about the reading (for permission and for obligation) that 'to be permitted to do a or b' entails 'to be permitted to do a' *and* 'to be permitted to do b'. In the setting of permitted announcements we have to clearly distinguish the action of announcing from the formula being announced. Permission to announce a or b indeed entails permission to perform either announcement, and choose between them. This is a nondeterministic action. This is different from the permission to make an announcement weaker than the announcement of a, such as $a \vee b$. In other words, permission to announce a or b is not the same as permission to announce $a \vee b$. Possibly, "permission to announce a or b" might be called *ambiguous*, as the 'or' may mean logical disjunction of formulas or non-deterministic choice between programs. But once the reading has been chosen, the course is clear.

We already observed that obligation and permission are not interdefinable. In Proposition 6.8 we showed that obligation adds to the expressivity of the logic. So $O\varphi \longleftrightarrow \neg P\neg\varphi$ is not valid. Now, Clearly, $O\varphi \longrightarrow \neg O\neg\varphi$ is valid in the class of permissive models. The norm is thus still considered as non-contradictory, we may want, to avoid this 'paradox', to include different norms in the same framework. We leave this for further research. But then again, even in the permissive cases, $P\varphi \vee P\neg\varphi$ is not valid: there is nothing against both p and $\neg p$ being forbidden announcements at the same time! For yet another example, consider the schema $O(O\varphi \longrightarrow \varphi)$, formalizing the requirement that obligations are fulfilled. In our setting, either we are in a non-permissive case and thus this obligation is satisfied, or it is a permissive one and thus as $O\varphi \longrightarrow \varphi$ is valid, this is equivalent to the validity of $O\top$, which indeed is a validity (note that $\top$ is weaker than any obligatory announcement, and that weakening holds for obligation).

A more recent development in deontic logic is the interaction between obligations and permissions and explicit agency (see [Chisholm, 1963, Horty, 2001]). The well-known Meinong-Chisholm reduction of "The agent is obliged to do a" to "It is obligatory that the agent does a" seems to have an interesting parallel in the logic of permitted announcements. In the logic of public announcements, the announcement by agent a is typically reduced to 'the (public) announcement of 'agent a knows φ'. It is relevant to recall at this stage that public announcements are supposedly made by outsiders of the system, not by agents modelled explicitly in the logical language. This observation can be applied in the logic of permitted and obligatory announcements! A Meinongian turn to permitted announcements seems to interpret $OK_i\varphi$—"It is obligatory that agent i announces φ" (announcements of φ by an agent i in the system are known to be true by that agent, so in fact have form $K_i\varphi$)—as an indirect form of agency in our logic, namely, we can let it stand for "Agent i is obliged to announce φ."

6.5.2 Deontic Action Logics

For the purpose of comparing our work with the existing literature we present a variant of the semantics for permission. Our current understanding of $P(\psi, \varphi)$ is that "*after the announcement of* ψ *it is permitted to give* at most *the information* φ". Any weakening of φ is also permitted. Instead, it was until now "*after the announcement of* ψ *it is permitted to give* exactly *the information* φ". We will write $P^=$ for that modality. It has the semantics: for all $\mathcal{M}$ and s in the domain of $\mathcal{M}$:

$$\mathcal{M}, s \models P^=(\psi, \varphi) \text{ iff } (s, [\![\psi]\!], [\![\langle\psi\rangle\varphi]\!]) \in \mathcal{P}.$$

The logic with P subsumes the one with $P^=$: let us expand a given relation $\mathcal{P}$ with all supersets for the third argument of a triple in that relation: for all subsets S''' of the domain of a given model $\mathcal{M}$, if $(s, S', S'') \in \mathcal{P}$ and $S'' \subseteq S'''$, then add (s, S', S''') to $\mathcal{P}$. Call the resulting relation $\mathcal{P}^=$ and let $\mathcal{M}^=$ be the model with $\mathcal{P}^=$ instead of $\mathcal{P}$. On the language without obligation, inductively define a translation $\bullet^=$ that replaces all occurrences of P by $P^=$. We now have that $\mathcal{M}, s \models P(\psi, \varphi)$ iff $\mathcal{M}^=, s \models P^=(\psi, \varphi)$. We compare the proposal using the operator $P^=$ with the related works presented in Section 3.3.

The Dynamic Logic of Permission Recall the framework proposed by [van der Meyden, 1996] and presented in Section 3.2.2. Our semantics for $P^=(\psi, \varphi)$ consists of the particular case where actions are public announcements. Thus, for α in Van der Meyden's $\Diamond(\alpha, \varphi)$ we take an announcement $\psi!$ such that $\Diamond(\psi!, \varphi)$ now means 'it is permitted to announce ψ, after which φ is true'. The precise correspondence is:

Proposition 6.37 *$\Diamond(\varphi!, \theta)$ is equivalent to $P^=(\top, \varphi) \wedge \langle\varphi\rangle\theta$*

Proof Given a model $\mathcal{M}$ with domain S, we can see the announcement $\varphi!$ as an atomic action which links each state $s \in [\![\varphi]\!]_{\mathcal{M}}$ to the same state $s \in S_\varphi$. This is a permitted action in Van der Meyden's semantics if and only if $(s, S, S_\varphi) \in \mathcal{P}$. By definition, $\mathcal{M}, s \models P^=(\top, \varphi)$ iff $(s, S, S_\varphi) \in \mathcal{P}$. The formula θ should then hold after the permitted announcement of φ.

□

Van der Meyden's $\Diamond(\varphi!, \theta)$ is found in a syntactic variant $Perm(\varphi)\theta$ in [Pucella and Weissman, 2004]. Now, we have that $P^=(\top, \varphi)$ is equivalent to $Perm(\varphi)$. Given the abbreviation $P(\varphi)$ in our language for $P^{=}(\top, \varphi)$, the correspondence is therefore very close.

Merging Frameworks for Interaction Recall the logical language $\mathcal{L}_{tpal}$ of Van Benthem et al.'s protocol logic $TPAL$ ([van Benthem et al., 2009]) presented in Section 3.3. We have seen that the domain was a set of histories, a history h being a succession of announcements, with the following semantics for the dynamic operator: $\mathcal{M}_\Pi, h \models \langle\psi\rangle\varphi$ iff

- $\mathcal{M}_\Pi, h \models \psi$
- $h' = h\psi \in \Pi$
- $\mathcal{M}_\Pi, h' \models \varphi$

This *suggests* to translate $P^{=}(\psi, \varphi)$ in $\mathcal{L}_{popal}$ by $[\psi]\langle\varphi\rangle\top$ in $\mathcal{L}_{tpal}$. (For convenience, we write announcements $\psi!$ and $\varphi!$ instead of singleton event models with precondition ψ and φ, respectively.) Unfortunately, this translation is imprecise. Consider executing these two announcements in a state s of an initial model $\mathcal{M}$. If $s\psi \notin H$ then $\mathcal{M}_\Pi, s \models [\psi]\langle\varphi\rangle\top$: after a non-permitted announcement, anything is permitted to be said, because anything holds after a necessity-type modal operator that cannot be executed. But $\mathcal{M}, s \not\models P^{=}(\psi, \varphi)$, because $(s, [\![\psi]\!], [\![\langle\psi\rangle\varphi]\!])$ is not in the $\mathcal{P}$ relation to validate it. In other words, in our logic we get the *full* forest produced by the protocol of all truthful public announcements, but some branches are coloured with *permitted* and others are coloured with not-permitted. The Van Benthem et al. approach produces a forest restricted to the protocol (i.e., restricted to permitted announcements only).

A more serious problem with such a translation is as follows. Our semantics allows that if something is later permitted to be said, we are already permitted to say something now in a different way, a consequence of the axiom "announcement and permission" $[\psi]P(\psi', \varphi) \leftrightarrow (\psi \longrightarrow P(\langle\psi\rangle\psi', \varphi))$. (This axiom holds for $P^{=}$ as well.) In $TPAL$ this would amount to requiring that (announcement) protocols are *postfix*-closed in the restricted sense that if $\pi'\pi'' = \pi \in \Pi$, then there is a single announcement ξ (combining all the announcements in the initial π' part in one complex announcement) such that $\xi\pi'' \in \Pi$.

Our logic with P instead of $P^{=}$ and with obligation O as well makes the comparison even more problematic.

As we now know, the notion of "*obligation to say* φ" cannot be captured only by the negation of permission to say anything else than φ (except in a very radical dictatorship), but much more by the fact that all that does not say *at least* φ is not permitted. This notion of obligation we consider a strong point of our logic $POPAL$, in which it differs from known other proposals.

6.6 Tableaux

We introduce here a proof method for $PPAL$, the logic reduced to the language wihtout obligation operator. This proof uses analytic tableaux to construct, given a formula φ, a model that satisfies φ (or prove its inconsistency if it fails). In order to do so we use notations near to the one proposed by [Balbiani et al., 2010] for a tableau method for PAL.

6.6.1 Definitions

The 'formulas' appearing in the tableau are $\mathcal{L}^1_{ppal}$-formulas prefixed by a natural number (n) that stands for a possible world in the model, and by a list of formulas (σ) representing successive updates.

Definition 6.38 (Labelled formula) *A* labelled formula *is a triple of the form $\langle \sigma, n, \varphi \rangle$ such that: σ is a (possibly empty) finite list of formulas of $\mathcal{L}^1_{ppal}$, $n \in \mathbb{N}$ and $\varphi \in \mathcal{L}^1_{ppal}$.*

The intuition behind this notation is that $\langle \sigma, n, \varphi \rangle$ appears in the tableau if the state n is still a state of the model after having announced successively the elements of the list σ, and the resulting state (after the announcements) satisfies φ. In other words $\langle \sigma, n, \varphi \rangle$ appears in the tableau if $n \in W_\sigma$ and $\mathcal{M}_\sigma, n \models \varphi$. If $n \in W_\sigma$ we say that 'n survives (the announcement) σ'.

To define what is a branch in this context, we denote by $\mathcal{L}^{(\mathbb{N})}$ the set of all the finite lists of $\mathcal{L}^1_{ppal}$-formulas. For any list $\sigma = (\sigma_1, \ldots, \sigma_n)$ we write $hd(\sigma) = \sigma_n$, $tl(\sigma) = (\sigma_1, \ldots, \sigma_{n-1})$, and for any formula $\psi \in \mathcal{L}^1_{ppal}$, $\sigma \circ \psi = (\sigma_1, \ldots, \sigma_n, \psi) \in \mathcal{L}^{(\mathbb{N})}$. We write $\varepsilon \in \mathcal{L}^{(\mathbb{N})}$ to designate the empty list. Finally, § is a countable set of formal symbols whose aim is to represent subsets of the initial set of worlds in a model.

Definition 6.39 (Branch) *Let § be a countable infinite set of symbols. A branch is a set of terms t of the form:*

1. *"$\langle \sigma, n, \varphi \rangle$" (labelled formula);*
2. *"nR_im" where $i \in AG$, $n, m \in \mathbb{N}$;*
3. *"$\Pi(n, S, S')$" where $n \in \mathbb{N}$ and $S, S' \in$ §;*
4. *"$n \in S$", "$n \notin S$" where $n \in \mathbb{N}$ and $S \in$ §;*
5. *"$S \vdash \sigma$", "$S \dashv \sigma$", "$S \nvdash \sigma$" and "$S \not\dashv \sigma$" where $S \in$ § and $\sigma \in \mathcal{L}^{(\mathbb{N})}$.*

We denote by $[\Diamond]$ the set of terms of the first category, $[R]$ for the second category, $[\Pi]$ for the third, $[\in]$ for the fourth and finally $[\vdash]$ for the fifth.

nR_im corresponds to the fact that the corresponding worlds of n and m are linked by $\mathcal{R}$ in a model. Π corresponds to $\mathcal{P}$. The term "$n \in S$" ("$n \notin S$") means that the world n

belongs (does not belong) to the set represented by S. The term "$S \vdash \sigma$" means "all worlds in the set represented by S survive the announcement of σ". The term "$S \dashv \sigma$" means "all worlds which survive the announcement of σ are in S". The term "$S \nvdash \sigma$" means "there exists a world n in S that does not survive the announcement of σ". The "$S \not\dashv \sigma$" means "there exists a world n which survives the announcement of σ and which is not in S".

Definition 6.40 (Initial Tableau for φ) *Given a formula $\varphi \in \mathcal{L}^1_{ppal}$, the set containing a unique branch $\{\langle \varepsilon, 0, \varphi \rangle\}$ is called the* initial tableau *for φ. We represent it as $\mathcal{T}_0(\varphi)$.*

We define now the notion of 'tableau for φ'. Informally, a tableau for φ is a set of branches obtained from the initial tableau by applying some rules. These rules are presented in a concise way in Table 6.2. We read them in the following way, considering a branch b of a given tableau $\mathcal{T}$:

$\dfrac{\alpha_1 \ ; \ \ldots ; \alpha_n}{\beta_1 ; \ldots ; \beta_p} \ (R1)$:

If the patterns $\alpha_1 ; \ldots \alpha_n$ are unifiable with a subset of terms of b and terms of the form $\beta_1, \ldots, \beta_p$ can not be found in b, then we add the instantiated instances of $\beta_1, \ldots, \beta_p$ in b. Formally if there exists a substitution s such that $s(\alpha_1), \ldots, s(\alpha_n) \in b$ and for all s', $\{s's(\beta_1), \ldots, s's(\beta_p)\} \not\subseteq b$ then $R1(b) = b \cup \{s's(\beta_1), \ldots, s's(\beta_p)\}$ where s' is a substitution for free variables of $s(\beta_1), \ldots, s(\beta_p)$.

$\dfrac{\alpha_1 ; \ldots \alpha_n}{\beta_1, \ldots, \beta_p \mid \gamma_1, \ldots, \gamma_r} \ (R2)$

If the patterns $\alpha_1 ; \ldots \alpha_n$ are unifiable with a subset of terms of b and no terms of the form $\beta_1, \ldots, \beta_p$ or $\gamma_1, \ldots, \gamma_r$ can be found in b, then we create one branch where we add $\beta_1, \ldots, \beta_p$ in b and another branch where we add $\gamma_1, \ldots, \gamma_r$. Formally if there exists a substitution s such that $s(\alpha_1), \ldots, s(\alpha_n) \in b$, for all s', $\{s's(\beta_1), \ldots, s's(\beta_p)\} \not\subseteq b$ and $\{s's(\gamma_1), \ldots, s's(\gamma_r)\} \not\subseteq b$ then $R1(b) = b \cup \{s's(\beta_1), \ldots, s's(\beta_p)\}; b \cup \{s's(\gamma_1), \ldots, s's(\gamma_r)\}$ where s' is a substitution for free variables of $s(\beta_1), \ldots, s(\beta_p)$, $s(\gamma_1), \ldots, s(\gamma_r)$.

A deterministic rule (such as $R1$) modifies a given branch, a non deterministic one (such as R_2) makes copies of a given branch and modifies them. We give some details about the rules appearing in Table 6.2 to make precise the understanding of this table:

- The rule $(R\vee)$ means that if a node n survives the list of announcements σ and satisfies $\psi \vee \varphi$ after the announcements, then we will consider one branch where n survives the list of announcements σ and satisfies ψ and another branch where n survives the list of announcements σ and satisfies φ.
- The rules (RT), (RS) and $(R4)$ capture respectively the reflexivity, symmetry and transitivity of the relation R_i.

- The rule $(R\sigma)$ explains the mechanism of announcement. If a world n survives after the announcement of σ, where σ is not the empty list, then it means the world survives after the announcement of the tail of σ and satisfies the head of σ after the announcement.
- The rule (RSB) concerns the behaviour of literals (propositions or negation of propositions) towards announcements: if a literal l is true in n after the announcements of σ then the literal l is already true in the world n without any announcement.
- The rule $(R \in)$ is a "cut" rule and choose for all nodes n and all already used symbols $S \in \S$ whether n should belong to the set represented by S or not.
- The rule (RP) creates a new permission relation $\Pi(n, S_1, S_2)$, S_1 and S_2 representing subsets of the domain. $S_1 \vdash \sigma$ and $S_1 \dashv \sigma$ ensures that S_1 corresponds to the domain of $\mathcal{M}_\sigma$ and $S_2 \vdash \sigma \circ \varphi$ ensures that S_2 corresponds to a subset of $[\![\varphi]\!]_{\mathcal{M}_\sigma}$.
- The rule $(R\neg P)$ guarantees that if a permission relation exists, it does not satisfy the conditions that would make the announcement of φ permitted.
- The rule $(R \vdash)$ explains the meaning of $S \vdash \sigma$: if a state belongs to S it 'survives' after the successive announcements of σ. Similarly $(R \dashv)$, $(R \nvdash)$ and $(R \not\dashv)$ explains the meaning of $S \dashv \sigma$, $S \nvdash \sigma$ and $S \not\dashv \sigma$.

Definition 6.41 (Tableau for φ) *Given a formula $\varphi \in \mathcal{L}^1_{ppal}$ we define the set $\mathcal{T}ab(\varphi)$ of 'tableaux for φ' inductively as follows:*

- $\mathcal{T}_0 = \{\{\langle \varepsilon, 0, \varphi \rangle\}\}$ *is in* $\mathcal{T}ab(\varphi)$
- *If $\mathcal{T}$ is obtained from $\mathcal{T}' \in \mathcal{T}ab(\varphi)$ by applying one of the rules of Table 6.2 to one of the branch of $\mathcal{T}'$, then $\mathcal{T} \in \mathcal{T}ab(\varphi)$*

Definition 6.42 (Closed Tableau) *Let b be a branch. We say that b is* closed *if it contains $\langle \sigma, n, \bot \rangle$ for some σ, n. We say that a tableau $\mathcal{T}$ for φ is closed if every branch in $b \in \mathcal{T}$ is closed. We say that a tableau (resp. branch) is open if it is not closed.*

Definition 6.43 (Satisfiability) *The branch b is said* satisfiable *iff there exists a model $\mathcal{M} = \langle W, \sim, V, \mathcal{P} \rangle$ and a function $f : \mathbb{N} \longrightarrow W$ such that*

1. *for all $i \in AG$, if $(nR_in') \in b$ then $f(n) \sim_i f(n')$*
2. *for all $\langle \sigma, n, \varphi \rangle \in b$: $f(n) \in W_\sigma$ and $\mathcal{M}_\sigma, f(n) \models \varphi$*
3. *for all $\Pi(n, S_1, S_2)$ in b, there exists $S^1, S^2 \subseteq W$ such that $(f(n), S^1, S^2) \in \mathcal{P}$ and*

$$\text{(i) for all } m \in \mathbb{N}, \begin{cases} \text{if } (m \in S_*) \text{ is in } b \text{ then } f(m) \in S^* \\ \text{if } (m \notin S_*) \text{ is in } b \text{ then } f(m) \notin S^* \end{cases}$$

$$\frac{\langle\sigma, n, \psi \vee \varphi\rangle}{\langle\sigma, n, \psi\rangle \mid \langle\sigma, n, \varphi\rangle}\ (R\vee) \qquad \frac{\langle\sigma, n, \neg(\psi \vee \varphi)\rangle}{\begin{array}{c}\langle\sigma, n, \neg\varphi\rangle\\ \langle\sigma, n, \neg\psi\rangle\end{array}}\ (R\wedge)$$

$$\frac{\langle\sigma, n, K_i\varphi\rangle\ ;\ nR_im}{\langle\varepsilon, m, [\sigma]\varphi\rangle}\ (RK_i) \qquad \frac{\langle\sigma, n, \neg K_i\varphi\rangle}{\langle\sigma, m, \neg\varphi\rangle\ ;\ nR_im}\ (R\hat{K}_i)$$

$$\frac{\langle\sigma, n, [\psi]\varphi\rangle}{\langle\sigma, n, \neg\psi\rangle \mid \langle\sigma \circ \psi, n, \varphi\rangle}\ (R[.]) \qquad \frac{\langle\sigma, n, \neg[\psi]\varphi\rangle}{\langle\sigma \circ \psi, n, \neg\varphi\rangle}\ (R\langle.\rangle)$$

$$\frac{\langle\sigma, n, \varphi\rangle}{\langle tl(\sigma), n, h(\sigma)\rangle}\ (R\sigma) \qquad \frac{\langle\sigma, n, \neg\neg\varphi\rangle}{\langle\sigma, n, \varphi\rangle}\ (R\neg)$$

$$\frac{\cdot}{nR_in}\ (RT) \qquad \frac{nR_im}{mR_in}\ (RS) \qquad \frac{\langle\sigma \circ \psi, n, l\rangle}{\langle\varepsilon, n, l\rangle}\ (RSB)$$

$$\frac{nR_im\ ;\ mR_io}{nR_io}\ (R4) \qquad \frac{\langle\sigma, n, \varphi\rangle\ ;\ \langle\sigma, n, \neg\varphi\rangle}{\langle\sigma, n, \bot\rangle}\ (R\bot)$$

$$\frac{\langle\sigma, n, P\varphi\rangle}{\Pi(n, S_1, S_2)\ ;\quad \begin{array}{c}S_1 \vdash \sigma\\ S_1 \dashv \sigma\\ S_2 \vdash \sigma \circ \varphi\end{array}}\ (RP) \qquad \frac{\langle\sigma, n, \neg P\varphi\rangle\ ;\ \Pi(n, S_1, S_2)}{S_1 \not\vdash \sigma \mid S_1 \not\dashv \sigma \mid S_2 \not\vdash \sigma \circ \varphi}\ (R\neg P)$$

$$\frac{\cdot}{n \in S \mid n \notin S}\ (R \in) \qquad \frac{\Pi(n, S', S'')}{n \in S''}\ (R\Pi)$$

$$\frac{S \vdash \sigma\ ;\ n \in S}{\langle\sigma, n, \top\rangle}\ (R \vdash) \qquad \frac{S \not\vdash \sigma}{n \in S\ ;\ \langle\varepsilon, n, [\sigma]\bot\rangle}\ (R \not\vdash)$$

$$\frac{S \dashv \sigma\ ;\ n \notin S}{\langle\varepsilon, n, [\sigma]\bot\rangle}\ (R \dashv) \qquad \frac{S \not\dashv \sigma}{n \notin S\ ;\ \langle\sigma, n, \top\rangle}\ (R \not\dashv)$$

Table 6.2: Tableau rules for $PPAL$

$$(ii)\ for\ all\ \sigma \in \mathcal{L}^{(\mathbb{N})}, \begin{cases} if\ (S_* \vdash \sigma)\ is\ in\ b\ then\ S^* \subseteq W_\sigma \\ if\ (S_* \not\vdash \sigma)\ is\ in\ b\ then\ S^* \not\subseteq W_\sigma \\ if\ (S_* \dashv \sigma)\ is\ in\ b\ then\ S^* \supseteq W_\sigma \\ if\ (S_* \not\dashv \sigma)\ is\ in\ b\ then\ S^* \not\supseteq W_\sigma \end{cases}$$

A tableau is said to be satisfiable *if it contains a satisfiable branch.*

6.6.2 Properties

We show in this section soundness and completeness of the tableau method:

Proposition 6.44 (Soundness) *If φ is satisfiable, then there exists no closed tableau for φ.*

Proof It is enough to see that any tableau rule preserves satisfiability of a given tableau, *i.e* if b is a satisfiable branch, then the set B of branches generated from b by applying a rule contains a satisfiable branch.

Indeed, if φ is satisfiable then the initial tableau for φ is satisfiable. Therefore every tableau would be satisfiable. Hence every tableau would contain an open branch (otherwise $\langle \sigma, n, \bot \rangle \in L$ and thus $\mathcal{M}_\sigma, f(n) \models \bot$) *Q.E.D.*

For every tableau rule we prove that it preserves satisfiability, by showing that it preserves the three constraints of the definition of satisfiability (Definition 6.43).

- **R$\wedge$,R$\vee$, R$\neg$,RK,R$\hat{K}$,RSB, R[.],R$\langle . \rangle$:** We let the reader prove that conditions 1 and 2 are preserved. Similar proofs can be found in [Balbiani et al., 2010]. Furthermore $[\Pi]$, $[\in]$ and $[\vdash]$ do not change by applying these rules. Therefore condition 3 is clearly preserved.

- **R$\bot$:** No satisfiable branch can satisfy the conditions to apply this rule, therefore it necessarily preserves satisfiability.

- **RT,R4,RS:** Condition 1 is preserved by the fact that the relation in the constructed model is an equivalence relation. Conditions 2 and 3 are preserved because only $[R]$ is modified by this rule.

- **Rσ:** If $f(n) \in W_\sigma$ then $f(n) \in W_{tl(\sigma)}$ and $\mathcal{M}_{tl(\sigma)}, f(n) \models hd(\sigma)$. Therefore condition 2 is preserved. Conditions 1 and 3 are preserved because only $[\Diamond]$ is modified by this rule.

- **RP:** Conditions 1 and 2 are preserved because $[\Diamond]$ and $[R]$ are not modified by this rule. Condition 3.i is preserved because $[\in]$ is not modified by this rule, and the sets S_1, S_2 created are new. For condition 3.ii recall that by hypothesis $\mathcal{M}_\sigma, f(n) \models P\varphi$ (and $f(n) \in W_\sigma$). Then there exists $(f(n), S^1, S^2) \in \mathcal{P}$ such that $S^1 = W_\sigma$ and $S^2 \subseteq W_{\sigma\circ\varphi}$. Let us consider the $\Pi(n, S_1, S_2)$ created by the rule RP, we can easily verify that $(f(n), S^1, S^2) \in \mathcal{P}$ is a good candidate to satisfy the requisites of 3.ii.

•**R¬P:** Conditions 1 and 2 and 3.i are preserved because $[\Diamond]$, $[R]$, $[\in]$ and $[\Pi]$ are not modified by this rule. For condition 3.ii, by hypothesis, $\mathcal{M}^\sigma, f(n) \models \neg P\varphi$ (and $f(n) \in W_\sigma$) and there exists a $(f(n), S^1, S^2) \in \mathcal{P}$ satisfying conditions 3.i and 3.ii. But $\mathcal{M}^\sigma, f(n) \models \neg P\varphi$ imposes that $S^1 \not\subseteq W_\sigma$ or $S^1 \not\supseteq W_\sigma$ or $S^2 \not\subseteq W_{\sigma\circ\varphi}$. Therefore, by choosing the corresponding branch 3.ii is preserved by the rule.

•**R∈:** Conditions 1, 2 and 3.ii are preserved because only $[\in]$ is modified by this rule. By hypothesis, n is a node and S_i is a letter appearing in an element of $[\Pi]$. Thus by hypothesis $f(n) \in W$ and there exists $S^i \subseteq W$ satisfying 3.i and 3.ii. Therefore either $f(n) \in S^i$ and we choose $n \in S_i$ in $[\in]$ or $f(n) \notin S^i$ and we choose $n \notin S_i$ in $[\in]$.

•**R⊢ :** Conditions 1. and 3. are evidently satisfied. The hypothesis imposes that there exists $S' \subseteq W_\sigma$ with $f(n) \in S'$. Therefore $f(n) \in W_\sigma$ and $\mathcal{M}_\sigma, f(n) \models \top$, which is condition 2.

•**R⊣ :** Conditions 1. and 3. are evidently satisfied. The hypothesis imposes that there exists $S' \supseteq W_\sigma$ with $f(n) \notin S'$. Therefore $f(n) \notin W_\sigma$, which means that $\mathcal{M}, f(n) \models [\sigma]\bot$. Condition 2 is thus satisfied.

•**R⊬ :** Condition 1. and 3.ii are evidently satisfied. By hypothesis, there exists $S' \not\subseteq W_\sigma$ satisfying 3.i. Therefore there exists a $f(n) \in S'$ such that $\mathcal{M}, f(n) \models [\sigma]\bot$. Therefore the rule preserves 2. and 3.i.

•**R$\not\dashv$:** Condition 1. and 3.ii are evidently satisfied. By hypothesis, there exists $S' \not\supseteq W_\sigma$ satisfying 3.i. Therefore there exists $f(n) \in W_\sigma$ s.t. $f(n) \notin S'$ (and $\mathcal{M}_\sigma, f(n) \models \top$). Therefore the rule preserves 2. and 3.i.

□

Definition 6.45 *Let b a branch and $R = \frac{\alpha_1;\ldots;\alpha_n}{\beta_1;\ldots;\beta_p|\gamma_1\ldots\gamma_r}$ a rule. We say that R is* applicable *on b iff there exists a substitution s such that $s(\alpha_1), \ldots, s(\alpha_n) \in b$ and for all substitutions s' we have $\{s's(\beta_1), \ldots, s's(\beta_p)\} \not\subseteq b$ and $\{s's(\gamma_1), \ldots, s's(\gamma_r)\} \not\subseteq b$.*

Let us now prove completeness. We first need the notion of saturated tableau:

Definition 6.46 *Let b a branch. We say that b is* saturated *under a rule R iff R is not applicable on b.*

A branch b is said saturated *if and only if it is saturated under all tableau rules.*

Let $\mathcal{T}$ be a tableau for φ_0. $\mathcal{T}$ is said saturated *if and only if for all branch $b \in \mathcal{T}$, b is either closed or saturated.*

Proposition 6.47 *If there exists an open saturated tableau for φ_0, then φ_0 is satisfiable.*

In order to prove it, we require the following notion of degree of a formula:

Definition 6.48 (Degree) *We define the degree of a formula* $\varphi \in \mathcal{L}_{popal}$ *inductively as follows:* $deg(p) = 0$, $deg(\neg\varphi) = deg(\varphi)$, $deg(\varphi_1 \vee \varphi_2) = max(deg(\varphi_1), deg(\varphi_2))$, $deg(K_i\varphi) = deg(\varphi)$, $deg([\psi]\varphi) = deg(\psi) + \mathrm{deg}(\varphi) + 2$, $deg(P\varphi) = deg(\varphi) + 2$.

Proof (of Proposition 6.47) Let T be an open saturated tableau for φ_0, and b be an open branch of T. b is saturated under every tableau rule. We then construct the model $\mathcal{M} = (W, \sim, V, \mathcal{P})$ with:

- $W = \{n \in \mathbb{N} \mid \langle \sigma, n, \varphi \rangle \in b \text{ for some } \sigma, \varphi\}$
- for every agent i, $\sim_i = \{(n, n') \mid nR_in' \in b\}$
- for every propositional atom p, $V(p) = \{n \in \mathbb{N} \mid \langle \varepsilon, n, p \rangle \in b\}$
- $\mathcal{P} = \{(n, g(S_1), g(S_2)) \in W \times 2^W \times 2^W \mid \Pi(n, S_1, S_2) \in b\}$, where by definition $g(S) = \{n \in \mathbb{N} \mid (n \in S) \in b\}$.

First of all, it is easy to see that $\mathcal{M}$ is a model. In particular

- for every agent i $\sim_i$ is an equivalence relation by rules RT, $R4$ and RS
- for every $(n, S^1, S^2) \in \mathcal{P}$, $n \in S^2$ by rule $R\Pi$ and $S^1 \subseteq S^2$ by rule RP with $R\vdash$ and $R\dashv$.

For all $a \in \mathbb{N}$, $k \in \mathbb{N}$, $\varphi \in \mathcal{L}^1_{ppal}$, we call $\rho(a, k, \varphi)$ the following property:

$\forall n \in W,\ \forall \sigma = (\sigma_1, \dots, \sigma_k)$, if $(*) \begin{cases} k + \Sigma deg(\sigma_i) + deg(\varphi) \leqslant a \\ \langle (\sigma_1, \dots, \sigma_k), n, \varphi \rangle \in b \end{cases}$ then $\begin{cases} n \in W_\sigma \\ \mathcal{M}_\sigma, n \models \varphi \end{cases}$

Note that $\rho(a, k, \varphi)$ is true if $k + deg(\varphi) > a$ (†).

For any pair of triplets (a', k', φ') and (a, k, φ) in $\mathbb{N} \times \mathbb{N} \times \mathcal{L}^1_{ppal}$, we say that $(a', k', \varphi') \ll (a, k, \varphi)$ if and only if: $a' < a$ or ($a' = a$ and $k' < k$) or ($(a', k') = (a, k)$ and $\varphi' \in Sub(\varphi)$). It is a well-founded (partial) order.

Let us suppose that $\rho(a', k', \varphi')$ is true for all $(a', k', \varphi') \ll (a, k, \varphi)$, and let us prove that $\rho(a, k, \varphi)$ is true, by reasoning on the structure of φ. It would prove that $\rho(a, k, \varphi)$ is true for all (a, k, φ) in $\mathbb{N} \times \mathbb{N} \times \mathcal{L}^1_{ppal}$, and in particular that $(a, 0, \varphi)$ is true for all a, φ in $\mathbb{N} \times \mathcal{L}^1_{ppal}$, which implies Proposition 6.49.

Let n be in W, and $\sigma_1, \dots, \sigma_k \in \mathcal{L}^1_{ppal}$ be such that $(*)$. We note $\sigma = (\sigma_1, \dots, \sigma_k)$.

Case $\varphi = p$: First, if $k = 0$ then $\langle \varepsilon, n, p \rangle \in b$ iff $n \in V(p)$ iff $\mathcal{M}, n \models p$ and the result is proved. Hence we assume that $k \geqslant 1$. Now, by hypothesis, we have $\langle (\sigma_1, \dots, \sigma_k), n, p \rangle \in b$, then

1. $\langle (\sigma_1, \dots, \sigma_{k-1}), n, \sigma_k \rangle \in b$ (by $R\sigma$-sat), and then $\begin{cases} n \in W_{tl(\sigma)} \\ \mathcal{M}_{tl(\sigma)}, n \models \sigma_k \end{cases}$ by $\rho(a, k-1, \sigma_k)$ because $(k-1) + deg(\sigma_1) + \dots + deg(\sigma_k) \leqslant a$. Therefore $n \in W_\sigma$.
2. $\langle \varepsilon, n, p \rangle \in b$ (by RSB-sat). We obtain $\mathcal{M}, n \models p$ by definition of the valuation. Thus $\mathcal{M}_\sigma, n \models p$

Cases $\varphi = \neg p, \top, \bot$: Idem, using the fact that the branch is open.

Case $\varphi = \varphi_1 \vee \varphi_2$: By hypothesis, we have $\langle \sigma, n, \varphi_1 \vee \varphi_2 \rangle \in b$. Thus, by $R\vee$-sat, we have $\langle \sigma, n, \varphi_1 \rangle \in b$ or $\langle \sigma, n, \varphi_2 \rangle \in b$. As for both i, $deg(\varphi_i) \leqslant deg(\varphi)$, and $\varphi_i \in Subf(\varphi)$ we can apply $\rho(a, k, \varphi_i)$ and thus we obtain $\begin{cases} n \in W_\sigma \\ \mathcal{M}_\sigma, n \models \varphi_1 \vee \varphi_2 \end{cases}$

Case $\varphi = \varphi_1 \wedge \varphi_2$: Identical, using $R\wedge$-sat.

Case $\varphi = K_i\varphi_1$: By hypothesis, we have $\langle \sigma, n, K_i\varphi_1 \rangle \in b$. Therefore, on one hand we obtain $\langle \sigma, n, \varphi_1 \rangle \in b$ (by RT-sat) and thus $\begin{cases} n \in W_\sigma \\ \mathcal{M}_\sigma, n \models \varphi_1 \end{cases}$ (by $\rho(a, k, \varphi_1)$). On the other hand, let $m \in W_\sigma$ be such that $n \sim_i m$, let us show that $\mathcal{M}_\sigma, m \models \varphi_1$. But by definition of $\sim_i$, $nR_im \in b$ and thus, by RK-saturation, $\langle \sigma, m, \varphi_1 \rangle \in b$.

Case $\varphi = \neg K_i\varphi_1$: By hypothesis, we have $\langle \sigma, n, \neg K_i\varphi_1 \rangle \in b$. By $R\hat{K}$-sat, there exists a $m \in W$ s.t. $nR_im \in b$ and $\langle \sigma, m, \neg\varphi_1 \rangle \in b$. This leads to the desired result by $\rho(a, k, \neg\varphi_1)$.

Case $\varphi = [\sigma_{k+1}]\varphi_1$: We suppose $a \geqslant 2$, otherwise it is already proven by (†) because $k + deg(\varphi) > a$. We write $\sigma' = \sigma \circ \sigma_{k+1}$. Now, by hypothesis we have $\langle \sigma, n, [\sigma_{k+1}]\varphi_1 \rangle \in b$, thus by $R[.]$-saturation:

- either $\langle \sigma, n, \neg\sigma_{k+1} \rangle \in b$ and then $\mathcal{M}_\sigma, n \models \neg\sigma_{k+1}$ by $\rho(a-2, k, \neg\sigma_{k+1})$, because
  $k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(\neg\sigma_{k+1})$
  $\leqslant k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(\sigma_{k+1}) + deg(\varphi_1)$
  $= k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg([\sigma_{k+1}]\varphi_1) - 2 \leqslant a - 2$

- or $\langle \sigma, n, \sigma_{k+1} \rangle \in b$. In this case $\langle \sigma', n, \varphi_1 \rangle \in b$ and we obtain that $\begin{cases} n \in W_{\sigma'} \\ \mathcal{M}_{\sigma'}, n \models \varphi_1 \end{cases}$ by $\rho(a-1, k+1, \varphi_1)$ because
  $k + 1 + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(\sigma_{k+1}) + deg(\varphi_1)$
  $= k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg([\sigma_{k+1}]\varphi_1) - 1 \leqslant a - 1$

Case $\varphi = \neg[\sigma_{k+1}]\varphi_1$: We suppose $a \geqslant 2$, otherwise it is already proven by (†), because $k + deg(\varphi) > a$. We write $\sigma' = \sigma \circ \sigma_{k+1}$. By hypothesis we have $\langle \sigma, n, \neg[\sigma_{k+1}]\varphi_1 \rangle \in b$, thus by $R\langle . \rangle$-saturation, $\langle \sigma, n, \sigma_{k+1} \rangle \in b$ and $\langle \sigma', n, \neg\varphi_1 \rangle \in b$, which implies that $\mathcal{M}_\sigma, n \models \sigma_{k+1}$ and $\mathcal{M}_{\sigma'}, n \models \sigma_{\neg\varphi_1}$ by $\rho(a-1, k+1, \neg\varphi_1)$ because
$k + 1 + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(\sigma_{k+1}) + deg(\neg\varphi_1)$
$= k + deg(\sigma_1) + \cdots + \mathrm{deg}(\sigma_k) + deg(\neg[\sigma_{k+1}]\varphi_1) - 1 \leqslant a - 1.$

Case $\varphi = P\varphi_1$: We suppose $a \geqslant 2$, otherwise it is already proven by (†), because $k + deg(\varphi) > a$. By hypothesis we have $\langle \sigma, n, P\varphi_1 \rangle \in b$, thus by RP-saturation, there is a $\Pi(n, S_1, S_2) \in b$ such that $\{S_1 \vdash \sigma, S_1 \dashv \sigma\top, S_2 \vdash \sigma \circ \varphi_1\} \subseteq b$. Therefore, by construction of $\mathcal{P}$, $(n, g(S_1), g(S_2)) \in \mathcal{P}$. Let us show that $g(S_1) = W_\sigma$ and $g(S_2) \subseteq W_{\sigma\circ\varphi_1}$.

First let $m \in g(S_1)$, thus $[m \in S_1]$ is in b and then, by $R\vdash$-saturation, $\langle \sigma, m, \top \rangle \in b$. Thus by $\rho(a-1, k, \top)$ we obtain $m \in W_\sigma$, because $k + deg(\sigma_1) + \cdots + deg(\sigma_k) \leqslant k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(P\varphi_1) - 1 \leqslant a - 1$. Thus $g(S_1) \subseteq W_\sigma$.

Second let $m \in g(S_2)$, thus $[m \in S_2]$ is in b and then, by $R\vdash$-saturation, $\langle \sigma \circ \varphi_1, m, \top \rangle \in b$. By $R\sigma$-sat $\langle \sigma, m, \varphi_1 \rangle \in b$ and $\rho(a, k, \varphi_1)$ we obtain $\begin{cases} n \in W_\sigma \\ \mathcal{M}_\sigma, n \models \varphi_1 \end{cases}$, because $k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(\varphi_1) \leqslant k + deg(\sigma_1) + \cdots + deg(\sigma_k) + deg(P\varphi_1) \leqslant a$. Thus $g(S_1) \subseteq W_{\sigma\circ\varphi_1}$.

Third, let $m \in W_\sigma$. Towards a contradiction assume that $m \notin g(S_1)$. Thus $[m \notin S_1]$ is in b, and by $R\dashv$-saturation we have that $\langle \varepsilon, m, [\sigma]\bot \rangle \in b$. Therefore, by $R[.]$-saturation iterated,

- either $\langle \varepsilon, m, \neg\sigma_1 \rangle \in b$ and $\mathcal{M}, m \models \neg\sigma_1$ by $\rho(a, 0, \neg\sigma_1)$
- or $\langle \sigma_1, m, \neg\sigma_2 \rangle \in b$ and $\mathcal{M}_{\sigma_1}, m \models \neg\sigma_2$ by $\rho(a, 1, \neg\sigma_2)$
- ...
- or $\langle (\sigma_1, \ldots, \sigma_k), m, \bot \rangle \in b$ and $\mathcal{M}_\sigma, m \models \bot$ by $\rho(a-1, k, \top)$

In all cases, that is in contradiction with the hypothesis $m \in W_\sigma$.

Case $\varphi = \neg P\varphi_1$: We suppose $a \geqslant 2$, otherwise it is already proven by (†), because $k + deg(\varphi) > a$. By hypothesis we have $\langle \sigma, n, \neg P\varphi_1 \rangle \in b$ and we want to show that show that for all $(n, S^1, S^2) \in \mathcal{P}$, $S^1 \neq [\![\langle\sigma\rangle\top]\!]$ or $S^2 \not\subseteq [\![\langle\sigma\rangle\varphi_1]\!]$. Let $(n, S^1, S^2) \in \mathcal{P}$, by definition of $\mathcal{P}$, we have $\Pi(n, S_1, S_2) \in b$ with $g(S_1) = S^1$ and $g(S_2) = S^2$. Thus, by $R\neg P$-saturation, either $(S_1 \not\vdash \sigma) \in b$ or $(S_1 \not\dashv \sigma) \in b$ or $(S_2 \not\vdash \sigma \circ \varphi_1) \in b$.

In the first case, by $R\not\vdash$-saturation, there exists a $m \in W$ such that $(m \in S_1)$ is in b and $\langle \varepsilon, m, [\sigma]\bot \rangle \in b$. Thus $m \in S^1$ and $\langle \varepsilon, m, [\sigma]\bot \rangle \in b$. As in the previous case, this is equivalent to $m \notin W_\sigma$. Therefore $S^1 \not\subseteq [\![\langle\sigma\rangle\top]\!]$.

In the third case we prove $S^2 \not\subseteq [\![\langle\sigma\rangle\varphi_1]\!]$ in the same way.

In the second case, by $R\not\dashv$-saturation, there exists a $m \in W$ such that $(m \notin S_1)$ is in b and $\langle \sigma, m, \top \rangle \in b$. Thus $m \notin g(S_1)$ and, by $\rho(a-1, k, \top)$, $m \in W_\sigma$. Therefore $S^1 \not\supseteq [\![\langle\sigma\rangle\top]\!]$.

□

We are now able to prove completeness of the tableau method:

Theorem 6.49 (Completeness) *If every tableau for φ is open then φ is satisfiable.*

Proof

Let S_N, called the *naive strategy* be the following application of the tableau rules $S_N \quad : \Big((R\vee); (RK_i); (R[.]); (R\sigma); (R\wedge); (R\hat{K}_i); (R\langle.\rangle); (R\neg); (RT); (RS); (RSB); (R4); (R\bot);$

$(R\neg P);(R\in);(R\Pi);(RP1);(RP2);(RP);(R\vdash);(R\nvdash);(R\dashv);(R\not\dashv)\Big)^*$. It can be called *naive* because it simply applies all the rules in an arbitrarily chosen order, and then starts again. If a rule is applicable, it is applied to one of the oldest instances that match with the conditions of the rule (FIFO). Remark that S_N satisfies the fact that for all $i \geqslant 0$, if a rule R is applicable in T_i then it will be considered at some step $j \geqslant i$.

Now let $\varphi \in \mathcal{L}^1_{ppal}$ be such that for all $\mathcal{T} \in \mathcal{T}ab(\varphi)$, $\mathcal{T}$ is open. Let $(\mathcal{T}_n) = \mathcal{T}ab(\varphi)$ the (possibly infinite) list of all the tableaux for φ constructed by applying strategy S_N to the initial tableau for φ, $\mathcal{T}_0 = \mathcal{T}_0(\varphi)$.

Now there are two possibilities for a given formula φ: starting from the initial tableau for φ, either S_N ends after n steps or S_N never ends. In the first case $\mathcal{T}_n$ is saturated (because no rule can be applied anymore) and it is open by hypothesis. Then it is an open saturated tableau. By proposition 6.47 we obtain the wanted result.

In the second case, let $\mathcal{T}_\infty$ be the infinite tree representing the infinite executions of S_N: for all $n \in \mathbb{N}$, the nodes of depth n are the branches of $\mathcal{T}_n$. Therefore the root node is $\{\langle \varepsilon, 0, \varphi \rangle\}$ and every node is a set containing every one of its ancestor nodes. It is a finitely branching tree because every rule creates a finite number of branches. Therefore, as it is infinite, it has an infinite *tree-branch* $\mathcal{B}_\infty$. We prove that this is an open branch. Indeed it is an infinite union of branches that are included one in the other for all $i \in \mathbb{N}$, $\mathcal{B}_i \subseteq \mathcal{B}_{i+1}$. We can thus define their limit $\bigcup_{i\in\mathbb{N}} \mathcal{B}_i$ and obtain $\mathcal{B}_\infty$.

Now for all $i \in \mathbb{N}$, $\mathcal{B}_i$ is open. Therefore B_∞ does not contain $\langle \sigma, n, \bot \rangle$ for any σ, n. Furthermore $\mathcal{B}_\infty$ is saturated by construction. Therefore it is an open saturated (infinite) branch. By Proposition 6.47 this leads to the satisfiability of φ.

□

6.6.3 Implementation

We first have implemented the semi-algorithm corresponding to the tableau. In order to get a more efficient algorithm, we implement the tableau method using two modifications of the tableau method. First, we add two additive rules $RP1$ and $RP2$:

$$\frac{\Pi(m, S_1, S_2); \langle \sigma, n, P\varphi \rangle; S_1 \vdash \sigma; S_1 \dashv \sigma; S_2 \vdash \sigma \circ \varphi}{\Pi(n, S_1, S_2)} (RP1)$$

$$\frac{\Pi(m, S_1, S_2); \langle \sigma, n, P\varphi \rangle; S_1 \vdash \sigma; S_1 \dashv \sigma}{\Pi(n, S_1, S_3)\ ;\ S_3 \vdash \sigma \circ \varphi} (RP2)$$

Table 6.3: Two additive tableau rules

The rules $(RP1)$ and $(RP2)$ simply subsume the rule (RP) in order to avoid creation of useless extra symbols of §. Second, we construct the models following strategy *Strat* defined as the application of the first rule which is applicable in the following list $(R\vee);(RK_i);(R[.]);(R\sigma);(R\wedge);(R\hat{K}_i);(R\langle.\rangle);(R\neg);(RT);(RS);(RSB);(R4);(R\bot);$

$(R\neg P); (R \in); (R\Pi); (RP1); (RP2); (RP); (R\vdash); (R\nvdash); (R \dashv); (R\not\dashv)$ and then restart strategy $Strat$, or quit if there is no applicable rule.

We have implemented the strategy $Strat$ under LotrecScheme[1] which is a software for rewriting terms close to the tool Lotrec (see [Gasquet et al., 2005] and [Said, 2010]). We have written all rules of the Table 6.2. Figure 6.2 shows the implementation of rules $(R\sigma)$, $(R4)$ and $(RP2)$.

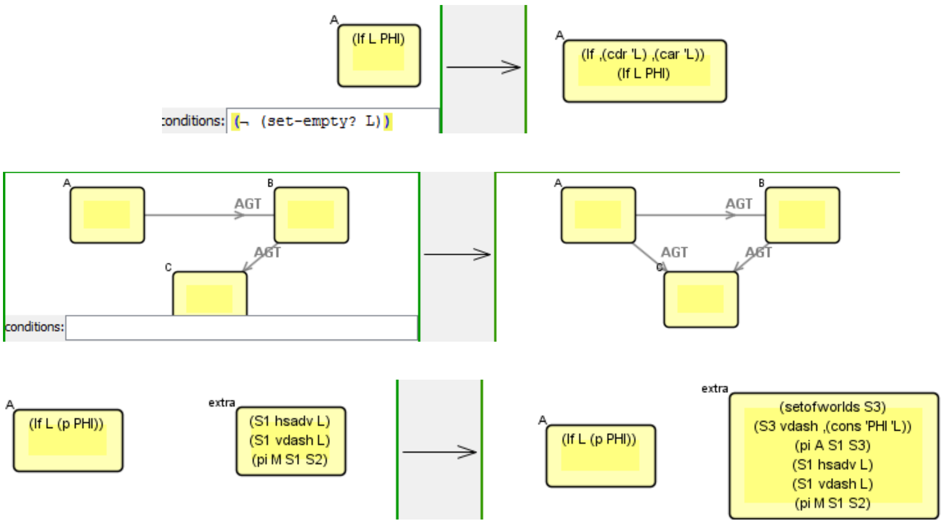


Figure 6.2: Rules $(R\sigma)$, $(R4)$ and $(RP2)$

We write (lf L PHI) in the node A for $\langle L, A, \varphi \rangle$. We use the primitive *cdr* and *car* of Scheme to get respectively the tail and the head of the list L. $(R4)$ is a graphical representation of transitivity. All terms $S \vdash \sigma$, $\Pi(n, S_1, S_2)$... are written in an extra node called "extra".

We adapt also the language $\mathcal{L}_{ppal}$ in order to implement it. Thus $K_i\varphi$ is written $\Box iPHI$ and $[\psi]\varphi$ is written $Ann\Box(PSI)PHI$. The other constructions $(\neg, \wedge, ...)$ are identical.

Let us see, as an example, the satisfiability of the following formula $\varphi = PK_1p \wedge PK_1q \wedge \neg PK_1(p \wedge q) \wedge \hat{K}_2PK_1(p \wedge q) \wedge [K_1p]PK_1(p \wedge q)$. It expresses the possibility of the following:

- Agent 1 is permitted to say (she knows) p and (she knows) q but is not permitted to say (she knows) $p \wedge q$ – as for Alex in the first example.
- Agent 2 imagines that Agent 1 is permitted to say $p \wedge q$ (she may not know the rule).
- Agent 1 has the permission to say (she knows) $p \wedge q$ after the announcement that (she knows) p.

Figure 6.3 presents the output of LotrecScheme if we ask for a model of φ. The actual state is $n24$, in the top left of the model. The node 'Pi' is not a state but represents the set

[1] http://www.irit.fr/~Francois.Schwarzentruber/lotrecscheme/.

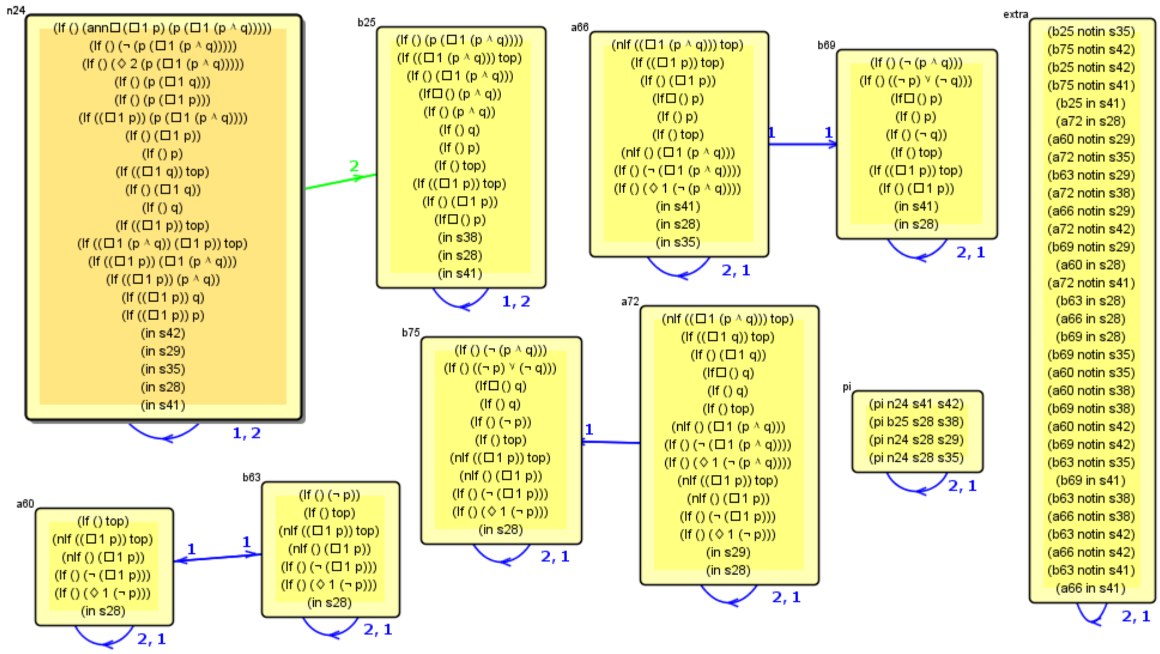


Figure 6.3: The output of the tableau method for $PK_1p \wedge PK_1q \wedge \neg PK_1(p \wedge q) \wedge \hat{K}_2 PK_1(p \wedge q) \wedge [K_1p]PK_1(p \wedge q)$

$\mathcal{P}$ of permitted transitions. As for 'extra', it is a list of the membership (or not) of the states of the models to the sets of states considered in 'Pi'. The membership of a state n to a set S is reproduced inside the state.

Let us have a look at 'Pi': $\Pi(n24, S28, S35)$ corresponds to PK_1p, $\Pi(n24, S28, S29)$ corresponds to PK_1q, $\Pi(b25, S28, S38)$ corresponds to $\hat{K}_2 PK_1(p \wedge q)$ and $\Pi(n24, S41, S42)$ corresponds to $[K_1p]PK_1(p \wedge q)$. We explain the details of this last example.

$\Pi(n24, S41, S42)$ means that in the state $n24$, the transition from the submodel based on $S41$ to the submodel based on $S42$ is a permitted transition. $S41$ is the following set of states: $S41 = \{n24, b25, a66, b69\}$ (the four states in the top of the model). Those are exactly the states that satisfy K_1p: $S41 = [\![K_1p]\!]$. Now $S42 = \{n24\}$, and $n24$ satisfies $K_1(p \wedge q)$. Therefore the restriction to $S42$ is stronger than the restriction to $[\![\langle K_1p\rangle K_1(p \wedge q)]\!]$. Therefore $[K_1p]PK_1(p \wedge q)$.

Unfortunately this tableau method does not provide a terminating algorithm. For instance, if you want to check if $P(\hat{K}_1\hat{K}_2p)$ the tableau method will not terminate. But we believe that we can tune the tableau method by adding a loop check rule in order to obtain a terminating procedure. We guess that the loop check rule may look like: "if there are two nodes $n1$ and $n2$ containing the same formulas, and such that "$n1 \in S$" $\in b$ iff "$n2 \in S$" $\in b$ and "$\Pi(n1, S1, S2)$" $\in b$ iff "$\Pi(n2, S1, S2)$" $\in b$ then we merge the two nodes $n1$ and $n2$."

Nevertheless, this tableau method for $PPAL$ opens perspectives in the purpose of creating a framework taking rules of games expressed in $POPAL$ and building automatically artificial agents able to reason about the corresponding game and play it.

Chapter 7

Private Permissions

7.1 Introduction

A medical laboratory (L) gets the results of the blood analysis of a patient called Michel (M). This confirms that Michel does not have AIDS (A). But of course, the results could have been different. To prevent that patients commit suicide when they learn that they are ill, French laboratories are not allowed to inform a patient directly of the results of a blood analysis (by email, by post, or whatever inconvenient form of impersonal or unprofessional communication). They have to inform a doctor (D), who receives the patient in his office, and then informs the patient. This protocol has to be followed when the patient has AIDS, but also when he does not have AIDS, otherwise having an appointment with the doctor could already be interpreted as confirmation of the disease, and we still get the terrible situation of lonely people in distress, that are a suicide risk.

Our aim, in this chapter, is to be able to formalize this kind of situation in which agents can communicate *with each other*, and where there are restrictions, that can be deontic, moral or hierarchical, on these announcements.

To formalize the concept of 'having the permission to say to somebody' we develop here a variant of $POPAL$ presented in the previous chapter. Indeed, the language considered here is an extension of Plaza's public announcement logic ([Plaza, 1989]), which we could call 'private announcement logic', with a modal operator P_i^G for permission, where $P_i^G\varphi$ expresses that agent i is allowed to say φ to the agents of the group G. As for $POPAL$, this logic can be seen as an adaption of the dynamic logic of permission proposed by [van der Meyden, 1996], later elaborated by [Pucella and Weissman, 2004], presented in Section 3.2. Recall that in Van der Meyden's work, $\Diamond(\alpha, \varphi)$ means "*there is a way to execute α which is permitted and after which φ is true*": we treat now the particular case where actions are announcements made by an agent to a group of agents. We also introduce an obligation operator $O_i^G\psi$, meaning that the agent is obliged to say ψ to the group G.

Once again, there is a relation between this proposal and the extension of public announcement logic with protocols by [van Benthem et al., 2009] presented in Section 3.3. In their approach, one cannot just announce anything that is true, but one can only announce a true formula that is part of the *protocol*, i.e., that is the first formula in a sequence of formulas (standing for a sequence of successive announcements) that is a member of a set of such sequences called the protocol. In other words, one can only announce *permitted* formulas. We do not have this limitation here: we can distinguish an announcement that cannot be done (because its content is false) from an announcement that is feasible but forbidden.

The permissions we model here are permissions for individual agents modelled in a multi-agent system. For example, if we have three agents a, b, c, we want to formalize that a has permission to say p to b, but not to c. We can model permission for agents using the standard method that agents only announce what they know: so agent a says $K_a p$ to b only. This would leave open what c learns from this interaction. The solution we chose is similar to the semi-public announcements ([Baltag and Moss, 2004]) where agents not involved in the communicative interaction at least are aware of the topic of conversation and of the agents involved in it: if a actually announces p to b, c considers it possible that a announces $K_a p$ to b, or that he announces $\neg K_a p$ to the same b. We also model such permissions and obligations of individual agents towards other agents in the system, or to groups of other agents.

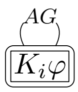
7.2 Logic of Permitted and Obligatory Private Announcements

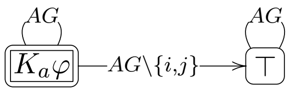
7.2.1 Introducing Agency and Individual Permission

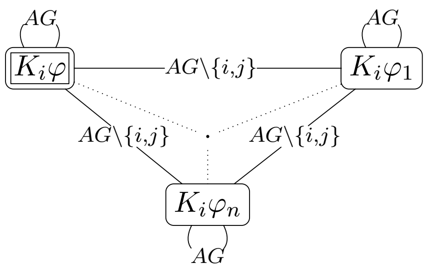
The reader may recall the previous chapter to better understand the current one. Indeed, as we will see in Proposition 7.15, the notion of permission of this logic with private announcements can be seen as an extension of the notion of permission formalized in $POPAL$ and presented in Chapter 6: if the group that 'receive' the announcement is always the whole group of agents AG, then any 'private' announcements to the group G is a public one. The analogous feature is not true for obligation, given that we have a different intuition of obligation in this chapter.

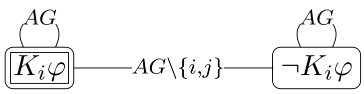
We want to consider private announcements, *i.e.* informative events in which an agent gives a piece of information that she has to another agent (or to a group of agents). Some choices have to be made. First, we consider that the content of an announcement is true, second we consider that the agent who is speaking can speak only about her own knowledge (it is the only thing she can actually know to be true), third we consider that the agents who hear the announcement believe it (and update their knowledge in consequence). This third point implies, in our understanding, that if the receiver of the announcement is a group then the information will be common to all its agents: anyone of them knows that any other one modified her knowledge. These points, except maybe the second, are classical in the field of dynamic epistemic logic ([van Ditmarsch et al., 2007]). But one characteristic of 'private announcement' still has to be fixed: what do the other agents learn? Indeed, the announcement can be hidden (and the others may believe that nothing is happening), or it may be the case that the other agents see who is communicating with who, without knowing anything about the content of the message. They may also know both the agents involved and the topic of the announcement, without knowing its truth value.

In [Baltag and Moss, 2004] the authors propose a general framework to express these different kinds of announcements. The main idea is that an announcement is represented by a graph: its states are deterministic events (and are labelled by formulas of the language that express the content of the messageà); a relation between two states, labeled by an agent i, is like in Kripke models an uncertainty for these agents about which of the two messages is given. With this formalism the previous examples of announcement can be represented as follows (in these examples i gives the information φ to j, the actual event being double-surrounded):

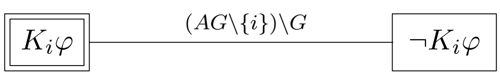
Public announcement: 

Hidden announcement : 

Visible private announcement: 

Idem with known topic: 

For both technical and practical reasons we restrict our formalism to the last kind of announcements, in which both the agents involved and the topic of the message are publicly known. In this context, we consider announcements of the type $!_i^G\varphi$ with $\varphi \in \mathcal{L}_{el}$, $i \in AG$ and $G \subseteq AG$. This formula represents the semi-private announcement by i to the group G of what she knows about φ. That can be "I know that φ is true", "I don't know if φ is true" (analogously to the treatment of questions in [Groenendijk and Stokhof, 1997]). Our formalism does not use event models, announcements are simply modelled as models restrictions, but the result of such an announcement is exactly the same as the result of the action of the corresponding event model as described previously:



You can read this in the following way: The actual action is that agent i says "I know that φ" and the agents in the group G know it, but every other agent that is not in G cannot distinguish this action from agent i saying "I don't know either φ".

It is worth noting that every state of every model satisfies only one of the two previous preconditions: in every state, either i knows φ, or she does not know φ. This implies that the action of $!_i^G\varphi$ on an epistemic model is only a copy of it with less epistemic arrows (if j learns that i knows φ, she does not consider anymore the states in which $\neg K_i\varphi$ was true).

Note that $!_i^G\varphi$ is identical to $!_i^G K_i\varphi$ and that $!_i^G\neg K_i\varphi$ is the same action model but pointed in the other state.

Such semi-public announcements can be modelled as restrictions on accessibility relations, while keeping the entire domain intact.

In the following subsection, instead of a permission relation that is the same for all agents, we define individual permission relations, one for each agent, and based on these structures we propose operators (let us take the one argument version) $P_i^G\psi$ and $O_i^G\psi$ for "agent i is permitted to announce whether she knows ψ to group G", and "agent i is obliged to announce whether she knows ψ to group G". The more general form of obligation is $O_i^{\vec{G}}\vec{\psi}$, where i has obligations ($\vec{\psi} =$) $\psi_1, ..., \psi_n$ to groups of agents ($\vec{G} =$) $G_1, ..., G_n$.

Let us see it in details.

7.2.2 Syntax of $\mathcal{L}_{popral}$

We first define the following partial language $\mathcal{L}_{pral}$:

Definition 7.1 ($\mathcal{L}_{pral}$) *The language of private announcement logic $\mathcal{L}_{pral}$ over $PROP$ and AG is defined inductively as follow:*

$$\psi ::= p|\bot|\neg\psi|\psi \vee \psi|K_i\psi|[!_i^G\psi]\psi$$

where $i \in AG$, $G \subseteq AG$, and $p \in PROP$

We are now able to introduce properly the syntax of our language:

Definition 7.2 *The language of permitted and obligatory private announcements logic $\mathcal{L}_{popral}$ is defined inductively as follows:*

$$\varphi ::= p|\bot|\neg\varphi|\varphi \vee \varphi|K_i\varphi|[!_i^G\psi]\varphi|P_i^G\psi|O_i^{\vec{G}}\vec{\psi}$$

where $i \in AG$, $G \subseteq AG$, $p \in PROP$, $\psi \in \mathcal{L}_{pral}$, $\vec{\psi} = (\psi_1, \ldots, \psi_n)$ is a tuple of $\mathcal{L}_{pral}$-formulas and $\vec{G} = (G_1, \ldots, G_n)$ a tuple of subsets of AG. We call $\mathcal{L}_{ppral}$ the fragment of the language without obligation operators (and $\mathcal{L}_{pral}$ is the fragment of the language without permission and obligation operators).

The boolean operators have the classical reading, and $K_i\varphi$ is read "agent i knows that φ". We read $[!_i^G\psi]\varphi$ as "after the announcement by agent i to the group G that (she knows) ψ, where the agents not in G also consider possible that i announces that she does not know ψ, φ becomes true", $P_i^G\psi$ by "i is allowed to say ψ to the group G" and $O_i^{\vec{G}}\vec{\psi}$ by "i is obliged to say ψ_1 to G_1 or $\ldots$ or ψ_n to G_n". The obligation is thus presented as a list of allowed announcements, and the agent satisfies her obligation by announcing one of them. This construction may seem complicated, and looks like a disjunction: is it possible to reduce

$O_i^{\vec{G}}\vec{\psi}$ to some kind of $\bigvee_i O_i^{G_i}\psi_i$? The answer is no: the following example shows that you can have the obligation to announce one thing or one other without having the obligation to announce any particular one of the two.

Example 7.3

In the Cluedo game[1]*, a murder has been committed and every player has some information about the weapon that has been used, the murderer or the room where the murder took place. A player (A) makes a proposal: "I suggest it was Professor Plum (PP), in the library (L), with the candlestick (C)." If another player (B) knows that this proposal is not correct, she has to show to A one card that invalidates it (for example showing the card that says that Professor Plum is innocent). B is thus obliged to give information to A, but she has no obligation to give one particular information. Suppose that B knows that the three propositions PP, L and C are false. Therefore we have (considering $\vec{A} = (A, A, A)$) :*

$$O_B^{\vec{A}}(\neg PP, \neg L, \neg C) \wedge \neg(O_B^A(\neg PP) \vee O_B^A(\neg L) \vee O_B^A(\neg C)).$$

The following technical notation will allow us to define the notion of *strong* obligation:

Definition 7.4 *Let $k, n \in \mathbb{N}$, let* $\begin{cases} \vec{\psi} := (\psi_1, \ldots, \psi_n) \\ \vec{G} := (G_1, \ldots, G_n) \end{cases}$ *and let* $\begin{cases} \vec{\psi}' := (\psi'_1, \ldots, \psi'_k) \\ \vec{G}' := (G'_1, \ldots, G'_k) \end{cases}$.

We note $(\vec{\psi}', \vec{G}') < (\vec{\psi}, \vec{G})$ if $(\vec{\psi}', \vec{G}') \neq (\vec{\psi}, \vec{G})$ and there exist $j_1, \ldots, j_k \in \mathbb{N}$ such that $1 \leqslant j_1 < \cdots < j_k \leqslant n$ and for all $l \in \{1, \ldots, k\}$, $\begin{cases} \psi_{j_l} = \psi'_l \\ G_{j_l} := G'_l \end{cases}$

This notation can be understood as the fact that the announcements (formula and group) of the first couple are announcements of the second. In particular, $k < n$.

We can now introduce the following useful abbreviations.

Definition 7.5

- $\langle !_i^G \psi \rangle \varphi := \neg [!_i^G \psi] \neg \varphi$
- Strong *obligation:* $\mathcal{O}_i^{\vec{G}}\vec{\psi} := O_i^{\vec{G}}\vec{\psi} \wedge \bigwedge_{(\vec{\psi}', \vec{G}') < (\vec{\psi}, \vec{G})} \neg O_i^{\vec{G}'}(\vec{\psi}')$
- $[!_i^G \psi^\sim]\varphi := [!_i^G \psi]\varphi \wedge [!_i^G \neg K_i \psi]\varphi$*: whatever i announces to G about her knowledge on ψ, φ becomes true after the announcement*
- *(Finite) sequence of announcements: an announcement $!_i^G \psi$ is a sequence of announcements, and if σ_1, σ_2 are sequences of announcements, then $\sigma_1; \sigma_2$ is a sequence of announcements.*
- *For all sequences of announcements σ_1, σ_2, we define $[\sigma_1; \sigma_2]\varphi := [\sigma_1][\sigma_2]\varphi$, $\langle \sigma_1; \sigma_2 \rangle \varphi := \langle \sigma_1 \rangle \langle \sigma_2 \rangle \varphi$ and $[(\sigma_1; \sigma_2)^\sim]\varphi := [\sigma_1^\sim][\sigma_2^\sim]\varphi$*

[1]For a complete explanation of the rules of this game: http://en.wikipedia.org/wiki/Cluedo

- *If the tuple of groups (and announcements) are made of a unique element, we abbreviate in the following way:* $O_i^G\varphi := O_i^{(G)}(\varphi)$

The first operator is the dual of $[!_i^G\psi]$. As we will see, it is equivalent to $[!_i^G\psi]\varphi$ with the supplementary condition that i can announce ψ. The second construction, $[!_i^G\psi^\sim]\varphi$, means that whatever i knows about ψ, if she says it to G then φ becomes true. The third one defines a stronger (and in our opinion more realistic) notion of obligation: not only a list of announcements one of which you have to ensure, but the smallest such list. This strong obligation will guarantee us to avoid Ross's paradox, indeed with this interpretation if you are (strongly) obliged to make an announcement you are not (strongly) obliged to make this announcement or another one. The fourth definition allows us to consider every sequence of announcements σ. The fifth abbreviates the notation in the case where the considered tuples are 1-uples.

7.2.3 Semantics for $\mathcal{L}_{pral}$

The models of our logic will be epistemic models augmented with an additional relation $\mathcal{P}$ between states and sets of relations, that represents, for each state, the announcements that are explicitly permitted to be done in this state. To define it properly, we need some preliminary notions:

Definition 7.6 (*AG*-relation) *Let AG and S be two sets. We define an AG-relation over S (or simply AG-relation) as a set $\mathcal{R} = \{R_i\}_{i\in AG}$ such that for all $i \in AG$, R_i is an equivalence relation over S.*

Definition 7.7 (Inclusion of *AG*-relations) *Let $\mathcal{R} = \{R_i\}_{i\in AG}$ and $\mathcal{R}' = \{R'_i\}_{i\in AG}$ be two AG-relations over S*

- *We say $\mathcal{R}' \subseteq \mathcal{R}$ if for all $i \in AG$, $R'_i \subseteq R_i$.*
- *For all $i \in AG$, we say $\mathcal{R}' \subseteq_i \mathcal{R}$ if $R'_i = R_i$ and for all $j \in AG\backslash\{i\}$, $R'_j \subseteq R_j$.*

Remark 7.8 *For all $i, G \in AG \times 2^{AG}$, we have that $\mathcal{R}' \subseteq_i \mathcal{R}$ only if $\mathcal{R}' \subseteq \mathcal{R}$. Note also that $\subseteq$ and $\subseteq_i$ are partial orders on AG-relations, in particular, they are transitive.*

We are now able to define the notion of restriction of an AG-relation:

Definition 7.9 *Let $\mathcal{M} = (S, \mathcal{R}, V)$ be an* epistemic *model over $PROP$ and AG (cf. Definition 2.17), $G \subseteq AG$ and $\psi \in \mathcal{L}_{el}$ an epistemic formula. We denote by $\mathcal{R}_{!_i^G\psi}$ the AG-relation $\mathcal{R}' = \{R_i\}_{i\in AG}$ such that:*

- $R'_i = R_i$
- *for all $j \notin G$, $R'_j = R_j$*

- *for all* $j \in G$, $R'_j = \{(s,t) \in R_j$ *s.t.* $\mathcal{M}, s \models K_i\psi$ *iff* $\mathcal{M}, t \models K_i\psi\}$.

Therefore we define $\mathcal{M}_{!_i^G\psi} = (S, \mathcal{R}_{!_i^G\psi}, V)$.

We underline some elements:

- We could have removed the first line $R'_i = R_i$. Indeed, in both cases ($i \in G$ and $i \notin G$) the other two lines would have imposed this condition. We do not remove it to be more explicit.
- $\mathcal{R}_{!_i^G\psi} \subseteq_i \mathcal{R}$
- $\mathcal{M}_{!_i^G\psi}$ is still an epistemic model, as $\mathcal{R}_{!_i^G\psi}$ is clearly an equivalence relation.

We remark also that we can extend this notion of restriction to any $\mathcal{L}_{pral}$-formula considering the following semantics to interpret $\mathcal{L}_{pral}$-formulas:

Definition 7.10 (Satisfiability relation for $\mathcal{L}_{pral}$) *Let* $\mathcal{M}$ *be a model and let* s *be a state of* S. *The satisfiability relation* $\mathcal{M}, s \models \varphi$ *is defined inductively on the structure of* φ*:*

- $\mathcal{M}, s \models p$ *iff* $s \in V(p)$
- $\mathcal{M}, s \not\models \bot$
- $\mathcal{M}, s \models \neg\psi$ *iff* $\mathcal{M}, s \not\models \psi$
- $\mathcal{M}, s \models \psi_1 \vee \psi_2$ *iff* ($\mathcal{M}, s \models \psi_1$ *or* $\mathcal{M}, s \models \psi_2$)
- $\mathcal{M}, s \models K_i\psi$ *iff for all* $t \sim_i s$, $\mathcal{M}, t \models \psi$
- $\mathcal{M}, s \models [!_i^G\psi]\chi$ *iff* $\mathcal{M}, s \models K_i\psi$ *implies* $\mathcal{M}_{!_i^G\psi}, s \models \chi$

where $\mathcal{M}_{!_i^G\psi} = (S, \mathcal{R}_{!_i^G\psi}, V)$ *with* $\mathcal{R}_{!_i^G\psi}$ *the AG-relation* $\mathcal{R}' = \{R_i\}_{i \in AG}$ *such that:*

- $R'_i = R_i$
- *for all* $j \notin G$, $R'_j = R_j$
- *for all* $j \in G \backslash \{i\}$, $R'_j = \{(s,t) \in R_j$ *s.t.* $\mathcal{M}, s \models K_i\psi$ *iff* $\mathcal{M}, t \models K_i\psi\}$.

7.2.4 Semantics for $\mathcal{L}_{popral}$

We can now define the *models* of our logic, that are epistemic models *augmented* in the following way.

Definition 7.11 *A model over a countable set of atomic propositions* $PROP$ *and a countable set of agents* AG *is a structure* $\mathcal{M} = (S, \mathcal{R}, V, \mathcal{P})$ *with*

- S *being a non-empty set of states*

- *$\mathcal{R}$ being a AG-relation over S.*
- *V mapping every $p \in PROP$ to a subset of S*
- *for all $i \in AG$, $\mathcal{T}_i = \{(s, \mathcal{R}', \mathcal{R}'') : \mathcal{R}', \mathcal{R}''$ are AG-relations, $s \in S$, $\mathcal{R}'' \subseteq_i \mathcal{R}' \subseteq \mathcal{R}\}$*
- *$\mathcal{P} = \{\mathcal{P}_i\}_{i \in AG}$ where for all $i \in AG$, $\mathcal{P}_i \subseteq \mathcal{T}_i$.*

The membership of $(s, \mathcal{R}', \mathcal{R}'')$ in $\mathcal{P}_i$ can be interpreted as follows: in state s, after every announcement that restricts the AG-relation to $\mathcal{R}'$, every announcement of i that restricts the AG-relation to $\mathcal{R}''$ is 'permitted'. Indeed, only $\mathcal{L}_{pral}$-formulas can be announced, so the definitions of restriction appearing in Definitions 7.9 and 7.10 are sufficient to define the update of a model $\mathcal{M}$ after the announcement $!_i^G\psi$ as the restriction $\mathcal{M}_{!_i^G\psi}$, and the interpretation of our logical language employing that model restriction.

Definition 7.12 (Satisfiability relation and restricted model) *Let $\mathcal{M}$ be a model and let s be a state of S. The satisfiability relation $\mathcal{M}, s \models \varphi$ is defined inductively on the structure of φ:*

- *$\mathcal{M}, s \models p$ iff $s \in V(p)$*
- *$\mathcal{M}, s \not\models \bot$*
- *$\mathcal{M}, s \models \neg\psi$ iff $\mathcal{M}, s \not\models \psi$*
- *$\mathcal{M}, s \models \psi_1 \vee \psi_2$ iff ($\mathcal{M}, s \models \psi_1$ or $\mathcal{M}, s \models \psi_2$)*
- *$\mathcal{M}, s \models K_i\psi$ iff for all $t \sim_i s$, $\mathcal{M}, t \models \psi$*
- *$\mathcal{M}, s \models [!_i^G\psi]\chi$ iff $\mathcal{M}, s \models K_i\psi$ implies $\mathcal{M}_{!_i^G\psi}, s \models \chi$*
- *$\mathcal{M}, s \models P_i^G\chi$ iff there exists $\psi \in \mathcal{L}_{pral}$ such that*
  1. *$\mathcal{M}, s \models K_i\psi$*
  2. *$[\![K_i\psi]\!]_{\mathcal{M}} \subseteq [\![K_i\chi]\!]_{\mathcal{M}}$ and*
  3. *$(s, \mathcal{R}, \mathcal{R}_{!_i^G\psi}) \in \mathcal{P}_i$*
- *$\mathcal{M}, s \models O_i^{\vec{G}}\vec{\varphi}$ iff*
  1. *for all $k \in \{1, \ldots, |\vec{G}|\}$, $\mathcal{M}, s \models K_i\varphi_k$ and $(s, \mathcal{R}, \mathcal{R}_{!_i^{G_k}\varphi_k}) \in \mathcal{P}_i$, and*
  2. *for all $(s, \mathcal{R}, \mathcal{R}_{!_i^H\chi}) \in \mathcal{P}_i$ there exists a $k \in \{1, \ldots, |AG|\}$ such that $\mathcal{R}_{!_i^H\chi} = \mathcal{R}_{!_i^{G_k}\chi}$ and $[\![K_i\chi]\!]_{\mathcal{M}} \subseteq [\![K_i\varphi_k]\!]_{\mathcal{M}}$*

where $\mathcal{M}_{!_i^G\psi} = (S, \mathcal{R}_{!_i^G\psi}, V, \mathcal{P}')$ with $\mathcal{P} = \{\mathcal{P}_i\}_{i \in AG}$ such that for all $j \in AG$, $\mathcal{P}'_j = \{(s, \mathcal{R}^1, \mathcal{R}^2) \in \mathcal{P}_j$ s.t. $\mathcal{R}^1 \subseteq \mathcal{R}_{!_i^G\psi}\}$

The semantics of the permission operator is thus the following: we say that i is allowed to say (that she knows) χ to G if there is something (ψ) that she knows, which announcement is more informative than the announcement of χ and gives a restriction that is in $\mathcal{P}_i$. The intuition hard encoded in the semantics for obligation is that given two different things that you are permitted to say, you should only have the obligation to announce the weaker of both. This explains the part "for all $(s, \mathcal{R}, \mathcal{R}_{!_i^G\chi}) \in \mathcal{P}_i$ there exists a k such that $\mathcal{R}_{!_i^G\chi} = \mathcal{R}_{!_i^{G_k}\chi}$ and $[\![K_i\chi]\!]_{\mathcal{M}} \subseteq [\![K_i\varphi_k]\!]_{\mathcal{M}}$" of the definition. This also intuitively entails that if you are obliged to do something then it should at least be permitted, and that intuition is indeed valid for the given semantics of obligation: for all $\varphi \in \mathcal{L}_{pral}$: $\models O_i^G\varphi \longrightarrow P_i^G\varphi$.

But this notion of obligation still does not say what you are actually forced to say: indeed, adding formulas to the tuple $\vec{\varphi}$ would maintain this obligation. This pushed us to define the notion of 'strong obligation' as the smallest such $\vec{\varphi}$ satisfying the definiendum: $\mathcal{M}, s \models \mathcal{O}_i^{\vec{G}}\vec{\varphi}$ iff $\mathcal{M}, s \models O_i^{\vec{G}}\vec{\varphi}$ and for all $\vec{\varphi'} < \vec{\varphi}$, $\mathcal{M}, s \not\models O_i^{\vec{G'}}(\vec{\varphi'})$.

Remark 7.13 *We define also $\mathcal{M}_{(!_i^G\psi;!_j^H\chi)} := (\mathcal{M}_{!_i^G\psi})_{!_j^H\chi}$, and we obtain inductively $\mathcal{M}_\sigma$ for every finite sequence of announcements σ.*

What is precisely the epistemic effect of the restriction $!_i^G\psi$? For every agent $j \in G$ after the announcement of $!_i^G\psi$ it becomes valid that j is able to distinguish the states where i knew ψ from the states where she did not know ψ. The other agents cannot distinguish any state they could not distinguish before, but they know that group G agents can.

Let us make another important remark.

Remark 7.14 *From Definition 7.5 we obtain:*

- $\mathcal{M}, s \models \langle !_i^G\psi\rangle\varphi$ *iff* $\mathcal{M}, s \models K_i\psi$ *and* $\mathcal{M}_{!_i^G\psi}, s \models \varphi$
- *By a direct induction on n and by Remark 7.13 we obtain that for any sequence of announcements $\sigma = !_{i_1}^{G_1}\psi_1; !_{i_2}^{G_2}\psi_2, \ldots; !_{i_n}^{G_n}\psi_n$, we have $\mathcal{M}, s \models \langle\sigma\rangle\varphi$ iff $(\mathcal{M}_{!_{i_1}^{G_1}\psi_1}, s \models \langle !_{i_2}^{G_2}\psi_2, \ldots; !_{i_n}^{G_n}\psi_n\rangle\varphi$ and $\mathcal{M}_{(!_{i_1}^{G_1}\psi_1; !_{i_2}^{G_2}\psi_2)}, s \models \langle !_{i_3}^{G_3}\psi_3, \ldots; !_{i_n}^{G_n}\psi_n\rangle\varphi$ and $\ldots$ and $\mathcal{M}_{\langle\sigma\rangle}, s \models \varphi)$. We abbreviate this by saying that $\mathcal{M}, s \models \langle\sigma\rangle\varphi$ iff $\mathcal{M}, s \models \langle\sigma\rangle\top$ and $\mathcal{M}_\sigma, s \models \varphi$.*

7.2.5 Comparison With the *Non-Agent* Version

As we announced in section 7.1, except for the notion of obligation that differs in the semantics, we can see this work as an extension of the previous work on permitted *public* announcement logic presented in Chapter 6. More precisely, consider the fragment of $\mathcal{L}_{popral}$ in which the group that receives any announcement is the entire group of agents AG. Let us call $\mathcal{L}_{popral}^{AG}$ this particular language, in which the announcements have the form $!_i^{AG}\psi$ (we thus abbreviate it in $!_i\psi$). The important remark is that such an announcement just divides the model into two submodels, depending on the initial valuation of $K_i\psi$: $\mathcal{R}_{!_i\psi}$ is the restriction of $\mathcal{R}$ to $[\![K_i\psi]\!]$

and $[\![\neg K_i\psi]\!]$. Now, for every model $\mathcal{M} = (S, \mathcal{R}, V, \mathcal{P})$ of $PPrAL$ we define the following model $\mathcal{M}^* = (S^*, \mathcal{R}^*, V^*, \mathcal{P}^*)$ of $PPAL$:

- $S^* = S$
- $\mathcal{R}^* = \mathcal{R}$
- $V^* = V$
- for every announcement $!_i\psi$ we define $S_{!_i\psi} = [\![K_i\psi]\!]_{\mathcal{M}}$. Therefore we can define S_σ for every succession of announcements σ by a direct induction (using Remark 7.13). Note that $\mathcal{R}_\sigma$ is
- $\mathcal{P}^* = \{(s, S_\sigma, S_{\sigma !_i\psi}) \in S \times 2^S \times 2^S \text{ s.t. } (s, \mathcal{R}_\sigma, \mathcal{R}_{\sigma !_i\psi}) \in \mathcal{P}\}$,

Therefore, an announcement $!_i\psi$ in $\mathcal{M}$ has the same meaning as the public announcement of $K_i\psi$ in $\mathcal{M}^*$: if $\varphi \in \mathcal{L}^{AG}_{popral}$ we define $\varphi^* \in \mathcal{L}_{popal}$ where φ^* is obtained from φ by replacing any occurrence of an announcement $!_i\psi$ by $K_i\psi$ and any occurrence of $P_i^{AG}\psi$ by $P(\top, K_i\psi)$. Hence the following:

Proposition 7.15 *Let φ be a $\mathcal{L}^{AG}_{ppral}$-formula. Then we have :*

$$\textit{for every model } \mathcal{M}, \textit{ and every state } s \in S,\ \mathcal{M}, s \models \varphi \textit{ iff } \mathcal{M}^*, s \models \varphi^*.$$

To prove it we need the following lemma:

Lemma 7.16 *Let $\varphi \in (\mathcal{L}^{AG}_{ppral})^*$ (i.e. the fragment of $\mathcal{L}_{popal}$ obtained by translating a $\mathcal{L}_{ppral}$-formula). Therefore for every model $\mathcal{M}$, every sequence of announcements $\sigma = !_{i_1}\psi_1; \dots; !_{i_n}\psi_n$ such that $\mathcal{M}, s \models \langle\sigma\rangle\top$ we have $(\mathcal{M}_\sigma)^*, s \models \varphi$ iff $\mathcal{M}^*|K(\sigma), s \models \varphi$ where $K(\sigma) := K_{i_1}\psi_1, \dots, K_{i_n}\psi_n$.*

In particular (if $n = 1$) for every model $\mathcal{M}$, announcements $!_i\psi$ such that $\mathcal{M}, s \models K_i\psi$ we have $(\mathcal{M}_{!_i\psi})^*, s \models \varphi$ iff $\mathcal{M}^*|K_i\psi, s \models \varphi$.

Proof We prove it by induction on the formula φ:

Base cases: for $\varphi = p$ by definition of V^*. For $\varphi = \bot$ it is trivial.

Inductive cases: Suppose that it is true for every subformula of φ^*

- $\varphi = \neg\chi, \chi_1 \vee \chi_2$: direct
- $\varphi = K_j\chi$: $(\mathcal{M}_\sigma)^*, s \models K_j\chi$
  iff for $t \in S'^*(= S)$ such that $s(R_j'^*)t$, $(\mathcal{M}_\sigma)^*, t \models \chi$
  iff for $t \in S$ such that $sR_j't$, $(\mathcal{M}_\sigma)^*, t \models \chi$ (because $\mathcal{R}'^* = \mathcal{R}'$)
  iff for $t \in S$ such that sR_jt and $\mathcal{M}, t \models \langle\sigma\rangle\top$, $(\mathcal{M}_\sigma)^*, t \models \chi$
  (by definition of R_j')
  iff for $t \in S$ such that sR_jt and $\mathcal{M}, t \models K_i\psi$, $\mathcal{M}^*|K_i\psi, t \models \chi$ (by IH)
  iff $\mathcal{M}^*|K_i\psi, s \models K_j\chi$

- $\varphi = \langle\chi\rangle\theta$: Note that φ is obtained by translating a $\mathcal{L}_{ppral}$-formula, thus $\varphi = \langle K_j\chi^*\rangle\theta^*$ where $\chi \in \mathcal{L}_{pral}, \theta \in \mathcal{L}_{ppral}$. Now:

$$\begin{array}{ll} & (\mathcal{M}_{!_i\psi})^*, s \models \langle K_j\chi^*\rangle\theta^* \\ \text{iff} & (\mathcal{M}_{!_i\psi})^*, s \models K_j\chi^* \text{ and } (\mathcal{M}_{!_i\psi})^*|(K_j\chi^*), s \models \theta^* \\ \text{iff} & \mathcal{M}^*|K_i\psi, s \models K_i\chi^* \text{ (by IH) and } (\mathcal{M}_{!_i\psi;!_j\chi})^*, s \models \theta^* \text{ (by IH)} \\ \text{iff} & \text{for } t \in S \text{ such that } sR'_jt,\ (\mathcal{M}_{!_i\psi})^*, t \models \chi \text{ (because } \mathcal{R}'^* = \mathcal{R}') \\ \text{iff} & \text{for } t \in S \text{ such that } sR_jt \text{ and } \mathcal{M}, t \models K_i\psi,\ (\mathcal{M}_{!_i\psi})^*, t \models \chi \\ \text{iff} & \text{for } t \in S \text{ such that } sR_jt \text{ and } \mathcal{M}, t \models K_i\psi,\ \mathcal{M}^*|K_i\psi, t \models \chi \text{ (by IH)} \\ \text{iff} & \mathcal{M}^*|K_i\psi, t \models K_j\chi \end{array}$$

$\square$

Proof (of Proposition 7.15) We prove it by induction on the structure of φ.

base cases ($\varphi = p, \bot$): Comes from the fact that $S = S^*$ and $V = V^*$.

induction steps Let us suppose that it is true for every subformula of φ:

- $\varphi = \neg\psi, \psi_1 \vee \psi_2$: by a simple use of IH
- $\varphi = K_i\psi$: by IH using that $\mathcal{R}^* = \mathcal{R}$
  (We can consider in particular that the property is true for every $\psi \in \mathcal{L}_{el}$)
- $\varphi = \langle !_i\psi\rangle\chi$: Therefore, $\varphi^* = \langle K_i\psi\rangle\chi^*$. Now

$$\begin{array}{ll} & \mathcal{M}, s \models \langle !_i\psi\rangle\chi \\ \text{iff} & \mathcal{M}, s \models K_i\psi \text{ and } \mathcal{M}_{!_i\psi}, s \models \chi \\ \text{iff} & \mathcal{M}^*, s \models K_i\psi \text{ and } (\mathcal{M}_{!_i\psi})^*, s \models \chi^* \\ \text{iff} & \mathcal{M}^*, s \models K_i\psi \text{ and } (\mathcal{M})^*|K_i\psi, s \models \chi^* \\ \text{iff} & \mathcal{M}^*, s \models \langle K_i\psi\rangle\chi^* \\ \text{iff} & \mathcal{M}^*, s \models (\langle !_i\psi\rangle\chi)^* \end{array}$$

- $\varphi = P_i^{AG}\chi$:

$$\begin{array}{ll} & \mathcal{M}, s \models P_i^{AG}\chi \\ \text{iff} & \text{there exists } \psi \in \mathcal{L}_{el} \text{ s.t. } \begin{cases} \mathcal{M}, s \models K_i\psi \\ [\![K_i\psi]\!]_{\mathcal{M}} \subseteq [\![K_i\chi]\!]_{\mathcal{M}} \text{ and} \\ (s, \mathcal{R}, \mathcal{R}_{!_i^G\psi}) \in \mathcal{P}_i) \end{cases} \\ \text{iff} & \text{there exists } \psi \in \mathcal{L}_{el} \text{ s.t. } \begin{cases} \mathcal{M}^*, s \models K_i\psi \\ [\![K_i\psi]\!]_{\mathcal{M}^*} \subseteq [\![K_i\chi]\!]_{\mathcal{M}^*} \text{ and} \\ (s, S, [\![K_i\psi]\!]) \in \mathcal{P}_i^*) \end{cases} \\ \text{iff} & \mathcal{M}^*, s \models P(\top, K_i\chi) \end{array}$$

$\square$

We also have a different intuition than in Chapter 6 of what 'strong obligation' means. The obligation is here seen as a disjunction (the obligation to make *one of* the announcements of $\vec{\psi}$) but as the minimal one.

7.2.6 Properties

First, here is a characterization of the link between a model and its reduction after an announcement:

Proposition 7.17 *For all formulas φ, all models $\mathcal{M}$, all states s of the model and all sequences of announcements σ we have: $\mathcal{M}_\sigma, s \models \varphi \iff \mathcal{M}, s \models [\sigma^\sim]\varphi$.*

Proof Let us first prove it for a single announcement $!_i^G\psi$:

$$\begin{array}{ll} & \mathcal{M}_{!_i^G\psi}, s \models \varphi \\ \text{iff} & \left\{ \begin{array}{l} \mathcal{M}, s \models K_i\psi \text{ implies } \mathcal{M}, s \models \langle !_i^G\psi\rangle\varphi \\ \mathcal{M}, s \models \neg K_i\psi \text{ implies } \mathcal{M}, s \models \langle !_i^G\neg K_i\psi\rangle\varphi \end{array} \right. \\ \text{iff} & \mathcal{M}, s \models (K_i\psi \longrightarrow \langle !_i^G\psi\rangle\varphi) \wedge (\neg K_i\psi \longrightarrow \langle !_i^G\neg K_i\psi\rangle\varphi) \\ \text{iff} & \mathcal{M}, s \models [!_i^G\psi]\varphi \wedge [!_i^G\neg K_i\psi]\varphi \text{ (because } \models \neg K_i\psi \longleftrightarrow K_i\neg K_i\psi) \\ \text{iff} & \mathcal{M}, s \models [!_i^G\psi^\sim]\varphi. \end{array}$$

By the definition of $[.^\sim]$ in Definition 7.5 and by Remark 7.13 the result extends to every sequence of announcements.

□

Let us see some properties of our logic, and in particular a reduced language that is expressively equivalent.

Proposition 7.18 *For all $p \in PROP$, all $i \in AG$, all $G \subseteq AG$, all $\psi \in \mathcal{L}_{pral}$, and all $\varphi, \varphi_1, \varphi_2 \in \mathcal{L}_{popral}$*

1. $\models [!_i^G\psi]p \longleftrightarrow (K_i\psi \longrightarrow p)$
2. $\models [!_i^G\psi]\bot \longleftrightarrow \neg K_i\psi$
3. $\models [!_i^G\psi]\neg\varphi \longleftrightarrow (K_i\psi \longrightarrow \neg[!_i\psi^G]\varphi)$
4. $\models [!_i^G\psi](\varphi_1 \vee \varphi_2) \longleftrightarrow ([!_i^G\psi]\varphi_1 \vee [!_i^G\psi]\varphi_2)$
5. *for all $j \in G \cup \{i\}$, $\models [!_i^G\psi]K_j\varphi \longleftrightarrow (K_i\psi \longrightarrow K_j[!_i^G\psi]\varphi)$*
6. *for all $j \notin (G \cup \{i\})$, $\models [!_i^G\psi]K_j\varphi \longleftrightarrow (K_i\psi \longrightarrow K_j([!_i^G\psi^\sim]\varphi))$*

These equivalences need some explanation, let us see the first one. It says that p is true after every possible announcement by i of ψ iff if i knows ψ (and then he *can* announce it) then p is true. This only says that an announcement cannot change the valuation.

Proof This proof is very similar to the proof of reduction of PAL (see [Plaza, 1989] for details). Let us see the proof of the last two ones, with $\mathcal{R}' := \mathcal{R}_{!_i^G\psi}$.

5. ($\Rightarrow$) let $j \in G$, $s \in S$ such that $\mathcal{M}, s \models [!_i^G\psi]K_j\varphi \wedge K_i\psi$ and t a state such that $(s,t) \in R_j$. We want to show that $\mathcal{M}, t \models [!_i^G\psi]\varphi$. Now either $\mathcal{M}, t \models \neg K_i\psi$ and it is

finished, or $\mathcal{M}, t \models K_i\psi$ and thus $(s,t) \in R_j$ implies that $s, t \in R'_j$ (by definition of $\mathcal{R}'$). As $\mathcal{M}_{!_i^G\psi}, s \models K_j\varphi$ we obtain that $\mathcal{M}_{!_i^G\psi}, t \models \varphi$ *Q.E.D.*

($\Leftarrow$) let $j \in G$, $s \in S$ such that $\mathcal{M}, s \models K_i\psi \wedge K_j[!_i^G\psi]\varphi$ and $t \in [\![K_i\psi]\!]$ a state such that $(s,t) \in R'_j$. We want to show that $\mathcal{M}_{!_i^G\psi}, t \models \varphi$. By definition of $\mathcal{R}'$, $(s,t) \in R'_j$ implies $(s,t) \in R_j$, and thus $\mathcal{M}, t \models [!_i^G\psi]\varphi$. As $t \in [\![K_i\psi]\!]$, $\mathcal{M}, t \models [!_i^G\psi]\varphi$ *Q.E.D.*

6. ($\Rightarrow$) let $j \notin (G \cup \{i\})$, $s \in S$ such that $\mathcal{M}, s \models [!_i^G\psi]K_j\varphi \wedge K_i\psi$ and t a state such that $(s,t) \in R_j$. We want to show that $\mathcal{M}, t \models [!_i^G\psi^\sim]\varphi$. Recall that in this case $R_j = R'_j$. Now $\mathcal{M}, s \models [!_i^G\psi]K_j\varphi \wedge K_i\psi$ implies that $\mathcal{M}_{!_i^G\psi}, s \models K_j\varphi$. Considering that $(s,t) \in R'_j$ we obtain $\mathcal{M}_{!_i^G\psi}, s \models \varphi$. By Proposition 7.17 this means that $\mathcal{M}, t \models [!_i^G\psi^\sim]\varphi$ *Q.E.D.*

   ($\Leftarrow$) let $j \notin (G \cup \{i\})$, $s \in S$ such that $\mathcal{M}, s \models K_i\psi \wedge K_j[!_i^G\psi^\sim]\varphi$ and $t \in [\![K_i\psi]\!]$ a state such that $(s,t) \in R'_j$. We want to show that $\mathcal{M}_{!_i^G\psi}, t \models \varphi$. Recalling that $R_j = R'_j$, $\mathcal{M}, s \models K_j[!_i^G\psi^\sim]\varphi$ implies $\mathcal{M}, t \models [!_i^G\psi^\sim]\varphi$. Thus $\mathcal{M}, t \models [!_i^G\psi]\varphi$ and, as $t \in [\![K_i\psi]\!]$, $\mathcal{M}_{!_i^G\psi}, t \models \varphi$ *Q.E.D.*

□

Definition 7.19 *We call $\mathcal{L}_{elPO}$ the following language:*

$$\varphi ::= p|\bot|\neg\varphi|\varphi \vee \varphi|K_i\varphi|\langle\sigma\rangle P_i^G\psi|\langle\sigma\rangle\neg P_i^G\psi|\langle\sigma\rangle O_i^{\vec{G}}\vec{\psi}|\langle\sigma\rangle\neg O_i^{\vec{G}}\vec{\psi}$$

where $\psi \in \mathcal{L}_{el}$, $i \in AG$, $G \subseteq AG$, σ is a sequence of announcements and for all ψ_j, G_j in the tuples $\vec{\psi}, \vec{G}$, we have $\psi_j \in \mathcal{L}_{el}$ and $G_j \subseteq AG$. It is the restriction of $\mathcal{L}_{popral}$ to the fragment without announcements except a sequence before permission and obligation operators. We call $\mathcal{L}_{elP}$ the restriction of $\mathcal{L}_{elPO}$ to the fragment without obligation operators.

Note that we cannot make the announcement disappear completely. This is due to the unary nature of the permission and obligation operator (cf. discussion on page 133). But it is possible to consider only formulas with announcements preceding exclusively permission or obligation operators. We could also have chosen the following equivalent language:

$$\varphi ::= p|\bot|\neg\varphi|\varphi \vee \varphi|K_i\varphi|\langle\sigma\rangle P_i^G\psi|[\sigma]P_i^G\psi|\langle\sigma\rangle O_i^{\vec{G}}\vec{\psi}|[\sigma]O_i^{\vec{G}}\vec{\psi}.$$

Corollary 7.20 (of Proposition 7.18) *$\mathcal{L}_{elPO}$ is expressively equivalent to $\mathcal{L}_{popral}$. $\mathcal{L}_{elP}$ is expressively equivalent to $\mathcal{L}_{ppral}$. $\mathcal{L}_{el}$ is expressively equivalent to $\mathcal{L}_{pral}$.*

This last language, $\mathcal{L}_{elP}$, will be used in Section 7.2.7 to prove the completeness of the given axiomatization. To prove Corollary 7.20, we use the following translation

Definition 7.21 *We define $tr : \mathcal{L}_{popral} \longrightarrow \mathcal{L}_{elPO}$ inductively on the structure of the formula:*

- *$tr(\bot) = \bot$ and for all atom $p \in PROP$, $tr(p) = p$*

- $tr(\neg\varphi) = \neg tr(\varphi)$; $tr(\varphi \vee \psi) = tr(\varphi) \vee tr(\psi)$; $tr(K_i\varphi) = K_i tr(\varphi)$
- $tr([!_i^G\psi]p) = K_i tr(\psi) \longrightarrow p$; $tr([!_i^G\psi]\bot) = \neg K_i tr(\psi)$
- $tr([!_i^G\psi]\neg\varphi) = K_i tr(\psi) \longrightarrow \neg tr([!_i^G\psi]\varphi)$
- $tr([!_i^G\psi](\varphi_1 \vee \varphi_2)) = tr([!_i^G\psi]\varphi_1) \vee tr([!_i^G\psi]\varphi_2)$
- *for all* $j \in G$, $tr([!_i^G\psi]K_j\varphi) = K_i tr(\psi) \longrightarrow K_j tr([!_i^G\psi]\varphi)$
- *for all* $k \notin G$, $tr([!_i^G\psi]K_k\varphi) = K_i tr(\psi) \longrightarrow K_k tr([!_i^G\psi^\sim]\varphi)$
- $tr(P_i^G\psi) = P_i^G tr(\psi)$; $tr(O_i^{\vec{G}}\vec{\psi}) = O_i^{\vec{G}}(tr(\psi_1), \dots, tr(\psi_n))$.

Note that it is actually true that for all $\varphi \in \mathcal{L}_{popral}$, $tr(\varphi) \in \mathcal{L}_{elPO}$. In fact, after any sequence of announcements, anything else than a P or an O is reduced by the translation. Note also that $tr(\mathcal{L}_{ppral}) = \mathcal{L}_{elP}$ and $tr(\mathcal{L}_{pral}) = \mathcal{L}_{el}$.

Proof (of Proposition 7.20) We prove the first property, we can prove the other two properties in the same way. Clearly $\mathcal{L}_{popral}$ is at least as expressive as $\mathcal{L}_{elPO}$ (because the second is included in the first). We use the translation tr defined in the previous definition. We obtain the wanted result by Proposition 7.18.

□

Another interesting property of our semantics is that, without any additive assumption, the following proposition is true:

Proposition 7.22 *For all models* $\mathcal{M}$ *and all formulas* $\psi_1, \psi_2 \in \mathcal{L}_{pral}$ *all* $i, j \in AG$, *we have that if* $\mathcal{M} \models K_i\psi_1 \longrightarrow K_i\psi_2$ *then* $\mathcal{M} \models P_i^G\psi_1 \longrightarrow P_i^G\psi_2$.

Note that this translates our intuition that: if an agent is allowed to give some information to some group of agents, then she is also allowed to give less information to the same group.

Corollary 7.23 *If we have* $\mathcal{M} \models K_i(\psi_1 \longrightarrow \psi_2)$ *then* $\mathcal{M} \models P_i^G\psi_1 \longrightarrow P_i^G\psi_2$.

This corollary comes directly from the Kripke nature of the models, that implies that $\models K_i(\psi_1 \longrightarrow \psi_2) \longrightarrow (K_i\psi_1 \longrightarrow K_i\psi_2)$.

Proof (of Proposition 7.22) By definition of the semantics of P and by transitivity of the implication. More precisely: let $s \in S$, suppose that $\mathcal{M}, s \models P_i^G\psi_1$, we want to show that $\mathcal{M}, s \models P_i^G\psi_2$. Then let $\psi_0 \in \mathcal{L}_{pral}$ be such that $\mathcal{M}, s \models K_i\psi_0$, $\mathcal{M} \models K_i\psi_0 \longrightarrow K_i\psi_1$ and $(s, \mathcal{R}, \mathcal{R}_{!_i^G\psi_0}) \in \mathcal{P}_i$, the three conditions of the semantics of P. Then we can keep the first and the third one, and replace the second, by transitivity of the implication, by $\mathcal{M} \models K_i\psi_0 \longrightarrow K_i\psi_2$. We then obtain $\mathcal{M}, s \models P_i^G\psi_2$.

□

Let us see now what are the consequences of the composition of different obligations or permissions:

Proposition 7.24 *For every $\psi \in \mathcal{L}_{popral}$, every agent i, every group G and every n-tuple $\vec{\psi}$ and $\vec{G}$ of $\mathcal{L}_{popral}$-formulas and groups,*

1. $\models P_i^G(\psi) \longrightarrow K_i\psi$
2. $\models O_i^{\vec{G}}\vec{\psi} \longrightarrow \bigwedge_{k\in\{1,\dots,n\}} P_i^{G_k}\psi_k$
3. $\models \mathcal{O}_i^{\vec{G}}\vec{\psi} \longrightarrow O_i^{\vec{G}}\vec{\psi}$
4. *for all permutation η,* $\models (\mathcal{O}_i^{\eta(\vec{G})}\eta(\vec{\psi}) \longleftrightarrow \mathcal{O}_i^{\vec{G}}\vec{\psi}) \wedge (O_i^{\eta(\vec{G})}\eta(\vec{\psi}) \longleftrightarrow O_i^{\vec{G}}\vec{\psi})$
5. $\models P_i^G\psi \wedge O_i^G\varphi \longrightarrow P_i^G(\psi \wedge \varphi_1)$

Proof The first proposition comes directly from $s \in [\![K_i\psi]\!]_{\mathcal{M}}$ in the semantics of $\mathcal{P}$, the second one from $(s, \mathcal{R}, \mathcal{R}_{!_i^{G_k}\varphi_k}) \in \mathcal{P}_i$ in the semantics of O, the third is induced by the syntax of $\mathcal{O}$ and the fourth one by the semantics of $\mathcal{O}$ and O.

Now the fifth one. Suppose that $\mathcal{M}, s \models P_i^{G_1}\psi \wedge O_i^G\varphi$. Therefore there exists $\chi \in \mathcal{L}_{pral}$ such that $\mathcal{M}, s \models K_i\chi$, $[\![K_i\chi]\!]_{\mathcal{M}} \subseteq [\![K_i\psi]\!]_{\mathcal{M}}$ and $(s, \mathcal{R}, \mathcal{R}_{!_i^{G_1}\chi}) \in \mathcal{P}_i$ (because $\mathcal{M}, s \models P_i^{G_1}\psi$). But as $\mathcal{M}, s \models O_i^G\varphi$ necessarily $[\![K_i\chi]\!]_{\mathcal{M}} \subseteq [\![K_i\varphi]\!]_{\mathcal{M}}$. Recalling that $[\![K_i\varphi]\!]_{\mathcal{M}} \cap [\![K_i\psi]\!]_{\mathcal{M}} = [\![K_i(\varphi \wedge \psi)]\!]_{\mathcal{M}}$ we get $[\![K_i\chi]\!]_{\mathcal{M}} \subseteq [\![K_i(\varphi \wedge \psi)]\!]_{\mathcal{M}}$ and thus $\mathcal{M}, s \models P_i^G(\psi \wedge \varphi_1)$

□

7.2.7 Soundness and Completeness

In this section, we give a sound and complete axiomatization of our logic. For technical reasons, we restrict this proposal to the language without obligation operators $\mathcal{L}_{ppral}$. We conjecture that a complete axiomatization for the whole logic exists, and plan to prove it in a further work.

Let $PPrAL$ be the axiomatization presented in table 7.1

Proposition 7.25 *$PPrAL$ is sound in all the models.*

Proof The soundness of the tautologies of propositional logic, of modus ponens, of the first four axioms of table 7.1, and of the necessitation rule for every K_i comes from the fact that for every model $\mathcal{M} = (S, V, \mathcal{R}, \mathcal{P})$, $(S, V, \mathcal{R})$ is a Kripke model where every $R_i \in \mathcal{R}$ is an equivalence relation (see [Fagin et al., 1995] for details). From the fifth to the tenth axiom, the soundness is proven by Propositions 7.18. Soundness of the last axiom is direct.

Let us prove soundness of the eleventh axiom. Let $\mathcal{M}$ be a model and s be a state of $\mathcal{M}$. Suppose that $\mathcal{M}, s \models \langle\sigma\rangle P_i^G\varphi$, then $\mathcal{M}, s \models \langle\sigma\rangle\top$ and $\mathcal{M}_\sigma, s \models P_i^G\varphi$ (with the abbreviation proposed in Remark 7.14). By Proposition 7.24 we get $\mathcal{M}_\sigma, s \models K_i\varphi$. Finally, $\mathcal{M}, s \models \langle\sigma\rangle K_i\varphi$.

Now note that the soundness of the permission inference rule has been proved in Proposition 7.22 using the same Remark 7.17.

□

all instantiations of propositional tautologies	
$K_i(\psi \longrightarrow \varphi) \longrightarrow (K_i\psi \longrightarrow K_i\varphi)$	distribution
$K_i\varphi \longrightarrow \varphi$	truth
$K_i\varphi \longrightarrow K_iK_i\varphi$	positive introspection
$\neg K_i\varphi \longrightarrow K_i\neg K_i\varphi$	negative introspection
$[!_i^G\psi]p \longleftrightarrow (K_i\psi \longrightarrow p)$	atomic permanence
$[!_i^G\psi]\bot \longleftrightarrow \neg K_i\psi$	ann. and false
$[!_i^G\psi]\neg\varphi \longleftrightarrow (K_i\psi \longrightarrow \neg[!_i^G\psi]\varphi)$	ann. and negation
$[!_i^G\psi](\varphi_1 \vee \varphi_2) \longleftrightarrow ([!_i^G\psi]\varphi_1 \vee [!_i^G\psi]\varphi_2)$	ann. and disjunction
if $j \in G \cup \{i\}$, $[!_i^G\psi]K_j\varphi \longleftrightarrow (K_i\psi \longrightarrow K_j[!_i^G\psi]\varphi)$	ann. and knowledge (1)
if $k \notin G \cup \{i\}$, $[!_i^G\psi]K_k\varphi \longleftrightarrow (K_i\psi \longrightarrow K_k[!_i^G\psi^\sim]\varphi)$	ann. and knowledge (2)
$\langle\sigma\rangle P_i^G\varphi \longrightarrow \langle\sigma\rangle K_i\varphi$	rationality of permission
$\langle\sigma\rangle\neg P_i^G\varphi \longrightarrow \langle\sigma\rangle\top$	ann reduction
From φ and $\varphi \longrightarrow \psi$ infer ψ	modus ponens
From φ infer $K_i\varphi$	necessitation of K_i
From φ infer $[!_i^G\psi]\varphi$	necessitation of announcement
From $[\sigma^\sim](K_i\varphi \longrightarrow K_i\varphi')$ infer $[\sigma^\sim](P_i^G\varphi \longrightarrow P_i^G\varphi')$	Permission rule

Table 7.1: The axiomatization $PPrAL$

Remark 7.26 *Note that we have in particular that for all $\varphi \in \mathcal{L}_{ppral}$, $\vdash_{PPrAL} \varphi \longleftrightarrow tr(\varphi)$. We often use this property in the following proofs, in particular to use φ instead of $tr(\varphi)$, for the sake of simplicity, when we need an $\mathcal{L}_{elP}$-formula.*

To prove the completeness result, we define the canonical model for $PPrAL$ in two phases:

Definition 7.27 (Epistemic part of the canonical model) *We define the following tuple $\mathcal{EM}^c = (S^c, V^c, \mathcal{R}^c)$ where $\mathcal{R}^c = \{R_i^c\}$:*

- *S^c is the set of all $\vdash_{PPrAL}$-maximal consistent sets*
- *for every $p \in PROP$, $V^c(p) = \{x \in S^c \mid p \in x\}$*
- *for every $i \in AG$, xR_i^cy iff $K_ix = K_iy$, where $K_ix = \{\varphi | K_i\varphi \in x\}$.*

Therefore, $\mathcal{EM}^c$ is an epistemic canonical model, and the truth lemma for $\mathcal{L}_{el}$ applies here: for all $\varphi \in \mathcal{L}_{el}$, $\mathcal{EM}^c, x \models \varphi$ iff $\varphi \in x$. Furthermore, this extends to $\mathcal{L}_{pral}$: for all $\varphi \in \mathcal{L}_{pral}$,

$\mathcal{EM}^c, x \models \varphi$
iff $\mathcal{EM}^c, x \models tr(\varphi)$ by Proposition 7.18
iff $tr(\varphi) \in x$ by the truth lemma for $\mathcal{L}_{el}$
iff $\varphi \in x$ by Remark 7.26

Given a formula $\psi \in \mathcal{L}_{pral}$ we can thus define the relation $\mathcal{R}_{!_i^G\psi}$ as in Definition 7.9. We are now able to define properly the canonical model for $\mathcal{L}_{ppral}$:

Definition 7.28 (Canonical model) *The canonical model $\mathcal{M}^c = (S^c, \mathcal{R}_i^c, V^c, \mathcal{P}^c)$ is defined as follows:*

- *S^c is the set of all $\vdash_{PPrAL}$-maximal consistent sets*
- *for every $p \in PROP$, $V^c(p) = \{x \in S^c \mid p \in x\}$*
- *for every $i \in AG$, $xR_i^c y$ iff $K_i x = K_i y$, where $K_i x = \{\varphi | K_i\varphi \in x\}$*
- *$\mathcal{T}_i^c = \{(x, \mathcal{R}_\sigma^c, \mathcal{R}_{\sigma!_i^G\chi}) : \chi \in \mathcal{L}_{pral}, \sigma$ a sequence of announcements, $G \subseteq AG\}$, and $\mathcal{P}_i^c = \mathcal{T}_i^c \setminus \{(x, \mathcal{R}_\sigma^c, \mathcal{R}_{\sigma!_i^G\chi}^c) :$ for some $\psi \in \mathcal{L}_{pral}, \vdash [\sigma^\sim](K_i\chi \longrightarrow K_i\psi)$ and $\langle\sigma\rangle\neg P_i^G\psi \in x\}$*

Remark 7.29 *Note that to define $\mathcal{P}_i^c$ in the canonical model we need to define $\mathcal{R}_\sigma$ and $\mathcal{R}_{\sigma!_i^G\chi}$ which suppose that we are able to define properly what $\mathcal{M}^c, x \models \langle\sigma\rangle\top$ and $\mathcal{M}^c, x \models \langle\sigma\rangle K_i\chi$ mean . But note that for all $\psi \in \mathcal{L}_{pral}$, $\mathcal{M}^c, x \models \psi$ iff $\mathcal{E}\mathcal{M}^c, x \models \psi$. Thus it has been done already.*

Proposition 7.30 *The canonical model is a model.*

Proof Indeed, S^c is a set, V^c, $\mathcal{P}_i^c$ and $\mathcal{R}_i^c$ for all i have the desired form. The only property we have to show is that if $(x, \mathcal{R}_1, \mathcal{R}_2) \in \mathcal{P}_i^c$ then $\mathcal{R}_2 \subseteq_i \mathcal{R}_1$. Thus, let us suppose that $(x, \mathcal{R}_1, \mathcal{R}_2) \in \mathcal{P}_i^c$, we have σ, χ, G such that $\mathcal{R}_1 = \mathcal{R}_\sigma^c$ and $\mathcal{R}_2 = \mathcal{R}_{\sigma|!_i^G\chi}^c$. By definition of $\subseteq_i$ we obtain the wanted result.

□

In the canonical model, a state is thus a set of formulas. The link between the fact that a formula φ is in a set x and the fact that $\mathcal{M}^c, x \models \varphi$ is given by the following proposition:

Proposition 7.31 (Truth Lemma for $\mathcal{L}_{elP}$) *For all $\varphi \in \mathcal{L}_{elP}$ we have:*

$$\Pi(\varphi) : for\ all\ x \in S^c, \mathcal{M}^c, x \models \varphi\ iff\ \varphi \in x$$

Proof We prove it by induction on the number of occurrences of a P operator.

base case: If φ is a formula without permission, $\Pi(\varphi)$ is a known result, the canonical model considered here being an extension of the canonical model for $S5$ (see [Blackburn et al., 2001] or [Fagin et al., 1995] for details).

Main induction step: Let us then suppose that $\Pi(\varphi)$ is true for every formula φ with at most n occurrences of a permission operator. Note that by Remarks 7.20 and 7.26 we can suppose the result for every formula of $\mathcal{L}_{ppral}$ containing at most n occurrences of a permission operator.

Let us now prove the wanted result for every formula with at most $n+1$ occurrences of a permission operator by induction on the structure of the formula φ:

- $\varphi = \neg\psi$: $\mathcal{M}_c, x \models \neg\psi$ iff $\mathcal{M}_c, x \not\models \psi$ iff $\psi \notin x$ (by IH) iff $\neg\psi \in x$ (by maximality of x)
- $\varphi = \varphi_1 \vee \varphi_2$: $\mathcal{M}_c, x \models \varphi_1 \vee \varphi_2$ iff $\mathcal{M}_c, x \models \varphi_1$ or $\mathcal{M}_c, x \models \varphi_2$ iff $\varphi_1 \in x$ or $\varphi_2 \in x$ (by IH) iff $\varphi_1 \vee \varphi_2 \in x$
- $\varphi = K_i\psi$: Let us first suppose that $K_i\psi \in x$ and let y be such that $x\mathcal{R}_i^c y$, we want to show that $\mathcal{M}_c, y \models \psi$. Indeed we have $K_i\psi \in y$ and then $\psi \in y$, which implies (by IH) that $\mathcal{M}_c, y \models \psi$.

  Reciprocally, let us suppose that $\mathcal{M}_c, x \models K_i\psi$ and that $K_i\psi \notin x$. Then $K_i x \cup \{\neg\psi\}$ is consistent which means that there exists a y such that $x\mathcal{R}_i y$ and $\neg\psi \in y$. By IH we obtain $\mathcal{M}^c, y \not\models \psi$ and thus $\mathcal{M}^c, x \not\models K_i\psi$ which is a contradiction. Thus the hypothesis $K_i\psi \notin x$ was wrong and $K_i\psi \in x$.
- $\varphi = \langle\sigma\rangle P_i^G \chi$:

  ($\Rightarrow$) By the main base case , we have that for every $\mathcal{L}_{pral}$-formula $\theta, \langle\sigma\rangle\theta \in x$ iff $\mathcal{M}^c, x \models \langle\sigma\rangle\theta$ ($*$). Let us suppose that $\langle\sigma\rangle P_i^G\chi \notin x$, *i.e.* $\neg\langle\sigma\rangle P_i^G\chi \in x$ by maximality, we want to show that $\mathcal{M}^c, x \not\models \langle\sigma\rangle P_i^G\chi$.

  Now, either $\langle\sigma\rangle\top \notin x$ and thus $\mathcal{M}^c, x \not\models \langle\sigma\rangle\top$ by IH, and then it is finished.

  Or $\langle\sigma\rangle\top \in x$, and then $\langle\sigma\rangle\neg P_i^G\chi \in x$. Let us suppose it. To show that $\mathcal{M}^c, x \not\models \langle\sigma\rangle P_i^G\chi$ let us take $\psi \in \mathcal{L}_{pral}$-formula such that $\mathcal{M}^c \models [\sigma^\sim](K_i\psi \longrightarrow K_i\chi)$ and let us prove that $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma|!_i^G\psi}) \notin \mathcal{P}_i^c$. Indeed if it is true for all such ψ we would have $\mathcal{M}^c, x \not\models \langle\sigma\rangle P_i^G\chi$. Now, by IH we have that $\vdash [\sigma^\sim](K_i\psi \longrightarrow K_i\chi)$, which means with $\langle\sigma\rangle\neg P_i^G\chi \in x$ that $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma|!_i^G\psi}) \notin \mathcal{P}_i^c$ *Q.E.D.*

  ($\Leftarrow$) If $\langle\sigma\rangle P_i^G\chi \in x$, then in particular $\langle\sigma\rangle K_i^G\chi \in x$ (and for all $\varphi \in \mathcal{L}_{ppral}$, $[\sigma^\sim]\varphi \in x$ iff $\langle\sigma\rangle\varphi \in x$, using the axioms 5–10). Now for all $\psi \in \mathcal{L}_{pral}$ such that $\vdash [\sigma^\sim](K_i\chi \longrightarrow K_i\psi)$ we have, by the permission inference rule, that $[\sigma^\sim]P_i^G\psi \in x$ and thus $\langle\sigma\rangle P_i^G\psi \in x$. Therefore, by definition of $\mathcal{P}^c$, $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma!_i^G\chi}) \in \mathcal{P}_i^c$. This proves that $\mathcal{M}^c, x \models \langle\sigma\rangle P_i^G\chi$.
- $\varphi = \langle\sigma\rangle\neg P_i^G\chi$:

  ($\Rightarrow$) Suppose that $\mathcal{M}^c, x \models \langle\sigma\rangle\neg P_i^G\chi$. Then $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma|!_i^G\chi}) \notin \mathcal{P}_i^c$, *i.e.* there exists a $\psi \in \mathcal{L}_{pral}$ such that $\vdash [\sigma^\sim](K_i\chi \longrightarrow K_i\psi)$ and $\langle\sigma\rangle\neg P_i^G\psi \in x$. Thus, by the permission inference rule , we obtain that $\langle\sigma\rangle\neg P_i^G\chi \in x$.

  ($\Leftarrow$) If $\langle\sigma\rangle\neg P_i^G\chi \in x$ then, by definition of $\mathcal{P}_i^c$, for all $\psi \in \mathcal{L}_{pral}$ such that $\vdash [\sigma^\sim](K_i\psi \longrightarrow K_i\chi)$ we have $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma|!_i^G\psi}) \notin \mathcal{P}_i^c$. This is equivalent, by the main base case, to the fact that for all $\psi \in \mathcal{L}_{pral}$ such that $\mathcal{M}^c \models [\sigma^\sim](K_i\psi \longrightarrow K_i\chi)$ we have $(x, \mathcal{R}_\sigma^c, \mathcal{R}^c_{\sigma|!_i^G\psi}) \notin \mathcal{P}_i^c$. That means exactly that $\mathcal{M}^c, x \not\models \langle\sigma\rangle P_i^G\chi$. We also have that $\langle\sigma\rangle\top \in x$ by the last axiom, which implies that $\mathcal{M}^c, x \models \langle\sigma\rangle\top$ by the main base case. Therefore from $\mathcal{M}^c, x \models \neg\langle\sigma\rangle P_i^G\chi$ we obtain $\mathcal{M}^c, x \models \langle\sigma\rangle\neg P_i^G\chi$ by using many times axiom ann. and negation.

□

Theorem 7.32 *PPrAL is sound and complete with respect to the class of all models.*

Proof Soundness has been proved in Proposition 7.25. For the completeness part, let $\varphi \in \mathcal{L}_{ppral}$ be a valid formula. Thus we have: $\models \varphi$ only if $\models tr(\varphi)$ only if $\mathcal{M}^c \models tr(\varphi)$ only if $\vdash tr(\varphi)$ (by Proposition 7.31) only if $\vdash \varphi$ (by Remark 7.26).

□

7.3 Case Study: AIDS

Let us recall the case of Michel, a patient who took an AIDS test, presented in the introduction. His case can be represented as in Figure 7.1 with individual permissions and transgressions. Let us explain the visual primitives.

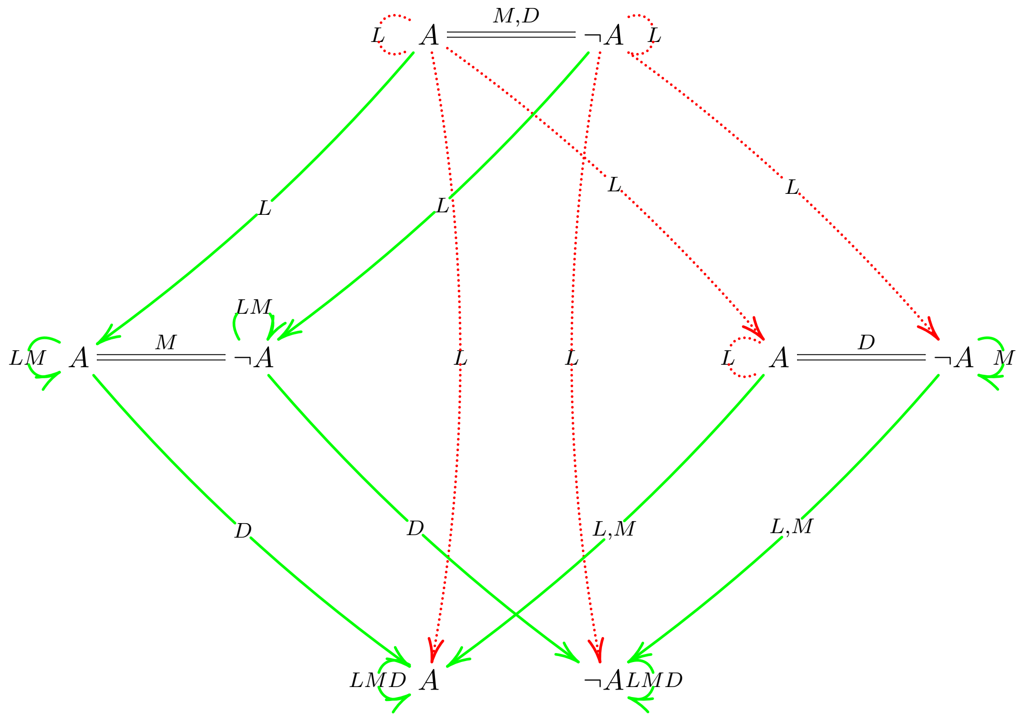


Figure 7.1: AIDS example

At the top, a two-state epistemic model where neither Michel nor the doctor can distinguish between a state where Michel has Aids (A) and a state where he has not ($\neg A$). Instead, the laboratory knows (there is no label L on the double links between A and $\neg A$). We see three more of such epistemic models in the figure, on the left is the situation where the uncertainty has been removed for the doctor but where Michel still is uncertain, on the right is the dual where the doctor is still uncertain, but Michel knows. For the record: this is the suicide-risk situation that we want to avoid! So getting there should not be permitted. At the bottom, everybody knows.

The pointed (and colored) arrows stand for the results of announcements. If they are dotted (and red), they are not permitted, if they are plain (and green), they are permitted.

The reader may note that such a representation *is not* a model of $POPrAL$: the transitions representing announcements are transitions between states and not triplets $(s, \mathcal{R}_1, \mathcal{R}_2)$. This is true, but we claim that such a graphical representation (much more readable) is analogous to the formal one. Indeed, the epistemic models appearing in the figure have the same domain (set of states) as the initial one, and their relation $\mathcal{R}'$ is a subrelation of the initial one $\mathcal{R}$. Furthermore, the pointed arrows (that represent announcements) start from a state and end in the same state of a reduced model. Finally, if such a pointed arrow is indexed by an agent a then the relation $\mathcal{R}'$ in the resulting model satisfy $\mathcal{R}' \subseteq_a \mathcal{R}$. For all these reasons, this graphic representation is identical to the formal one, that should be written in the following (complex) way: 1 $\mathcal{M} = (S, \mathcal{R}, V, \mathcal{P})$ where $PROP = \{A\}$, $AG = \{M, D, L\}$ and

- $S = \{s_1, s_2\}$
- $\mathcal{R} = \{R_M, R_D, R_L\}$ with
  - $R_M = R_D = \{(s_1, s_1), (s_1, s_2), (s_2, s_2)\}$
  - $R_L = \{(s_1, s_1), (s_2, s_2)\}$
- $V(A) = \{s_1\}$

It remains to define $\mathcal{P}$ properly. To do so, let us call $\mathcal{R}^1$ the left epistemic model's relation, $\mathcal{R}^2$ the right epistemic model's one, and $\mathcal{R}^3$ the bottom epistemic model's one. Formally for all $i \in \{1, 2, 3\}$, $\mathcal{R}^i = \{R^i_M, R^i_D, R^i_L\}$ with:

- $R^1_M = R_M$; $R^1_D = R^1_L = R_L$
- $R^2_D = R_D$; $R^2_M = R^2_L = R_L$
- $R^3_M = R^3_D = R^3_L = R_L$

We are now able to define $\mathcal{P} = \{\mathcal{P}_M, \mathcal{P}_D, \mathcal{P}_L\}$ with

- $\mathcal{P}_M = \{(s_1, \mathcal{R}^1, \mathcal{R}^1), (s_2, \mathcal{R}^1, \mathcal{R}^1), (s_2, \mathcal{R}^2, \mathcal{R}^2), (s_1, \mathcal{R}^2, \mathcal{R}^3), (s_2, \mathcal{R}^2, \mathcal{R}^3), (s_1, \mathcal{R}^3, \mathcal{R}^3), (s_2, \mathcal{R}^3, \mathcal{R}^3)\}$
- $\mathcal{P}_D = \{(s_1, \mathcal{R}^1, \mathcal{R}^3), (s_2, \mathcal{R}^1, \mathcal{R}^3), (s_1, \mathcal{R}^3, \mathcal{R}^3), (s_2, \mathcal{R}^3, \mathcal{R}^3)\}$
- $\mathcal{P}_L = \{(s_1, \mathcal{R}, \mathcal{R}^1), (s_2, \mathcal{R}, \mathcal{R}^1), (s_1, \mathcal{R}^1, \mathcal{R}^1), (s_2, \mathcal{R}^1, \mathcal{R}^1), (s_1, \mathcal{R}^2, \mathcal{R}^3), (s_2, \mathcal{R}^2, \mathcal{R}^3), (s_1, \mathcal{R}^3, \mathcal{R}^3), (s_2, \mathcal{R}^3, \mathcal{R}^3)\}$

We admit that these models are quite difficult to be read by a human being. But this is not the case for a computer, and the proposed equivalent graphical representation is useful to do direct comments, as we do here. The reflexive arrows labelled with L in the top

Kripke model (information state) for the laboratory show that the empty announcement, *i.e.* the announcement of 'true', after which the same structure results, is not permitted for the laboratory: the laboratory is obliged to say something informative. In the top information state, it is permitted for the laboratory to announce the outcome of the test to the doctor. This action (whether A is true, or the different action for when A is false) brings us to the left state (plain arrows). Also, these are the *only* plain arrows from those states: the laboratory is obliged to inform the doctor: $\mathcal{M}, s_1 \models O_L^D A$ and $\mathcal{M}, s_2 \models O_L^D \neg A$

Apart from the reflexive arrows, two more non-permitted actions on top are: informing Michel (go to right): $\mathcal{M}, s_1 \models \neg P_L^M A$, and informing the doctor and Michel at the same time (straight to the bottom, where everybody knows): $\mathcal{M}, s_1 \models \neg P_L^{\{D,M\}} A$. The other connections can be similarly explained. Finally, after the violation of the laboratory informing Michel, the laboratory is still obliged to inform the doctor, and also Michel is obliged to contact the doctor: $\mathcal{M}, s_1 \models \langle !_L^M A \rangle (O_L^D A \wedge O_M^D A)$ — which we could now interpret that action will be undertaken if Michel has not contacted the doctor after the laboratory has improperly informed him directly of the outcome of the AIDS-test. Therefore, the plain arrows from the right to the bottom are labelled both with L and with M. Further intricacies in the reflexive arrows on the right-hand side are left to the imagination of the reader.

Chapter 8

Conclusion

Situations involving norms and communication are frequent: communicative games (such as card games), medical databases, protocols of communication, *etc.* However, there exists no general framework to handle such situations. Yet such a general framework would be of great interest from a theoretical point of view to analyse and explain such situations, and from a practical one, to create artificial agents able to reason in terms of permission to communicate.

The field of logic may be appropriate to create such frameworks, and many logicians are trying to make progress in this direction (see for example [Aucher et al., 2010, van Benthem et al., 2009]). This *mémoire* is an attempt to use dynamic epistemic logics in order to understand the notion of 'right to say'. We may first recall what has been presented so far.

After having introduced our work, we presented in a first part the basic notions of modal logic for the study of knowledge. Some would disagree with the fact that such a basic presentation appears in a memoir of Ph. D. But can we pretend that scientists may not be concerned about making readable their work? To be readable was the aim of this chapter.

We then presented the works in progress on dynamic epistemic logic (*DEL*), essentially since Plaza introduced the Public Announcement Logic [Plaza, 1989], and the basic notions of deontic logic. We saw that in the field of DEL some notions could be developed, and we added our contribution to the building in the Chapter 4 and 5. We saw that the notion of objectivity and group capacity (in the field of *DEL*) introduced in these chapters may be useful to understand and solve problems about protocols of communication. We presented several technical results (decidability, complexity) of the logics $LAUOB$ and GAL introduced.

That led us to the last two chapters, which gave proposals to understand the notion of right to say in a dynamic context, namely *POPAL* and *POPrAL*. The first of them treats the case where the announcements are public and can be thought as external events (made by an omniscient agent). The second generalizes the first treating a case of private announcement (that includes public ones) and considering agency of announcements. Technical results have also been proved.

However, much remains to be done. First of all, there are technical results we would like to obtain in each chapter proposed here. For example, chapters 5,6 and 7 have few results of complexity, results we consider interesting. What is the class of complexity of the SAT-problem for $POPAL$ (resp GAL, $POPrAL$)? As another example, we would also like to extend the axiomatization of $PPrAL$ to get one for the whole language of $POPrAL$.

The tableau method for $PPAL$ proposed in Section 6.6 could also be generalized to the whole language $\mathcal{L}_{popal}$, integrating in it the deontic operator of permission O. Another easy extension would be to add a tableau rule $(RP^{=})$ as proposed in Table 8.1 to consider operator $P^{=}$ presented in section 6.5.2. This operator may be useful if a different intuition of permission is required. A more important extension would be to develop an analogous method

$$\frac{\langle \sigma, n, P\varphi \rangle}{\Pi(n, S_1, S_2); S_1 \vdash \sigma; S_1 \dashv \sigma; S_2 \vdash \sigma \circ \varphi; S_2 \dashv \sigma \circ \varphi} \; (RP^{=})$$

Table 8.1: Tableau rule for $P^{=}$

for $POPrAL$ or even for a more general logic for dynamic epistemic logic and permission (cf. Section 8.2).

Moreover, we could imagine an attempt to mix the different languages proposed in this thesis, and consider together notions of permission, group announcements, knowledge and update of objective beliefs.

Besides, other recent works could open ways to expand our formalism. We could consider for example to expand the framework with changing permissions, as in Pucella et al. [Pucella and Weissman, 2004]. In this context, as in the example of Section 7.3, this would mean to define an operation that defines or modifies the permission relation $\mathcal{P}$ of a given model, which could be a $POPAL$ or a $POPrAL$ model.

As another example, following [Cuppens et al., 2005b], it would be interesting to integrate the notion of role in the attribution of permissions and obligations, or to consider the notion of deadline (cf [Cuppens et al., 2005a]) that gives sense to the notion of obligation imposing a limited time to fulfill the obligation.

Besides it would be possible to use this formalism. Therefore, there is no doubt that this thesis opens to further researches, and an exhaustive list of all such possible technical results or conceivable extensions seems impossible – in fact the reader may imagine a lot of them. Yet some possible extensions of our work appears of greater interest, and we present in the two next sections the two most important in our opinion.

8.1 Dealing with Privacy Policies

In [Aucher et al., 2010] the authors propose a formalism close to ours with the difference that in their proposal the 'right to say' can be derived from the 'right to know'. In other words, they assume that there is a list of permissions or obligations to know that have to be satisfied. This list defines whether an announcement is permitted or not: it is permitted if and only if it leads to a situation that satisfies these obligations/permissions. A basic presentation of this framework has been done in Section 3.3.

This condition ('the permission to say is derivable from the permission to know') does not allow to model every situation: two different announcements that lead to the same epistemic situations (then that satisfy the same permissions to know) can be one permitted the other forbidden, as far as they are announced in different situations. The situation presented in section 7.3 is such a counter-example: we have $\mathcal{M}, A \models \neg P_L^{\{D,M\}} A \wedge \langle !_L^M A \rangle P_L^D A$: the laboratory is not allowed to announce A to the doctor and Michel at the same time, but after having informed Michel it is allowed to announce it to the doctor. These two announcements are not permitted in the same way, and yet they lead to the same situation (the A-situation at the bottom of Figure 7.1) — but they did not come from the same one!

But in some situations, the restrictions on announcements are derivable from a Privacy Policy which says what each agent is allowed to know. Therefore, we would try to adapt Aucher *et al.*'s notion of Privacy Policy, to model those multi-agent situations in which the right to know is the relevant notion, deriving the permission relation $\mathcal{P}$ (as presented in Chapter 7) from it. We make here a draft of such a proposal, that would avoid, in our opinion, another limit of their work, namely that it is limited to a single 2-agents situation, in which a *sender* gives information to a *receiver*, the latter having a perfect knowledge of the epistemic state of the former.

Let us see a compelling example, cited from [Aucher et al., 2010]:

> Consider the information about websites contacted by a user (U), which are available on a server logfile. The list of websites for each user is clearly a sensitive information which he would not like to disclose. However, knowing which websites have been visited is a valuable information, for example, for the configuration of a firewall, or to make statistics. Thus it has become anonym by replacing the names of the users with numbers by means of a hashcode (h). So even if one knows the list of users one cannot understand who contacted which website. However, from the association between users and numbers and between numbers and websites the original information can be reconstructed. Therefore the mappings from the users to the numbers (c) and from the numbers to the websites (e) can be distributed individually but not altogether since their association would allow to reconstruct the mapping from the users to the websites they visited (v): $c \wedge e \longrightarrow v$.

The last sentence says that the user u is permitted to know c and to know e but not to know v. The privacy policy being the set of what is (not) permitted to be known by the agent, it would be in this case $\{\neg P(K_U v)\}$: it is not permitted that U knows v. To construct our model, presented in Figure 8.1, we start from an initial epistemic model and we define $\mathcal{P}_S$ as the set of transitions such that $K_u v$ is wrong in the resulting state.

In Figure 8.1, as in Figure 7.1, double arrows represent knowledge, dotted arrows (red) represent non-permitted announcements and plain (green) arrows permitted ones. At the top a three-state Kripke model where the user does not know neither c nor e, nor v. We see three more of such models in the figure. On the left is the situation where the uncertainty on c has been removed, on the right is the situation in which the uncertainty on e has been

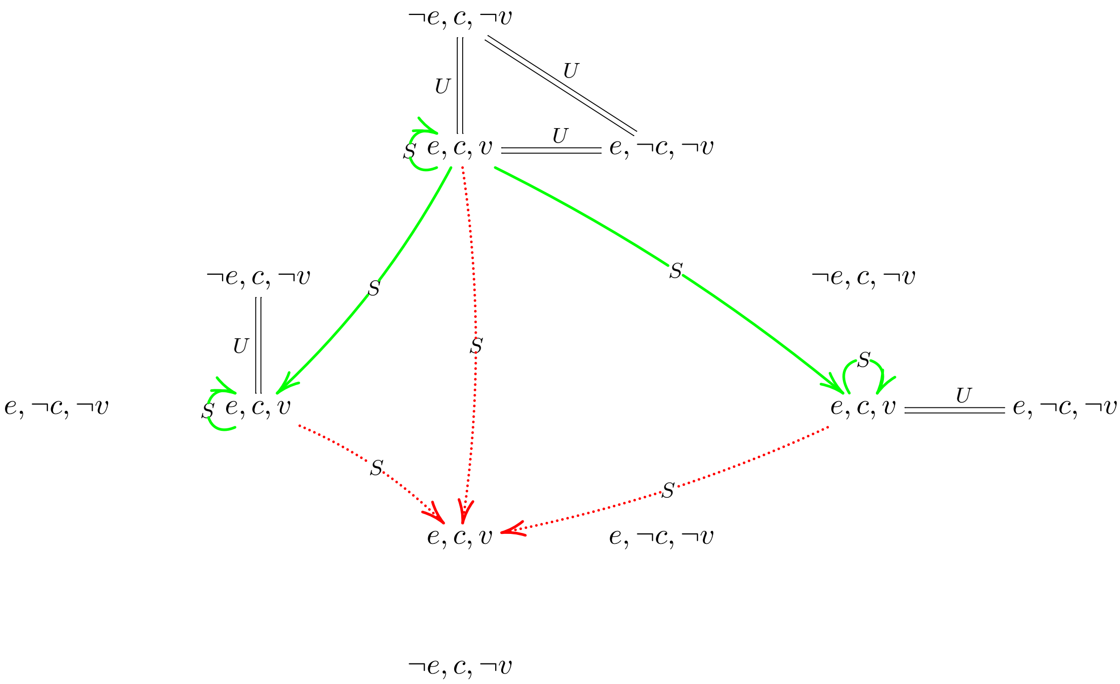


Figure 8.1: server example

removed. At the bottom, v is known (epistemic situation we want to forbid). Once again, the pointed arrows stand for the results of announcements. If they are dotted, they are not permitted, if they are plain, they are permitted. To make the figure readable, we put only the pointed arrows for the e, c, v-states. Thus this model says that the server is allowed to say c or to say e to the user, but not to say v to her: $\mathcal{M}, (e, c, v) \models P_s^u c \wedge P_s^u e \wedge \neg P_s^u v$. After the announcement of one of the two pieces of information e or c, the server is not allowed to say to the user the other one: $\mathcal{M}, (e, c, v) \models \langle !_s^u e\rangle \neg P_s^u c \wedge \langle !_s^u c\rangle \neg P_s^u e$.

We can thus define the notion of privacy policy in the following way: An *epistemic norm* is a construction of the form $pre \longrightarrow P_i\psi$ or $pre \longrightarrow \neg P_i\psi$ with $pre, \psi \in \mathcal{L}_{el}$ and $i \in AG$. A privacy policy is a finite set of epistemic norms.

We interpret $pre \longrightarrow P_i\psi$ (resp. $pre \longrightarrow \neg P_i\psi$) by "if pre is true then i is allowed (resp. not allowed) to get to a situation where ψ is true". We note $pre \longrightarrow F_i\psi := pre \longrightarrow O_i\neg\psi$.

We can thus construct deterministically a $POPrAL$ model starting from an epistemic model and a privacy policy. Following [Aucher et al., 2010] we consider two situations: the liberal situation considers that every situation that is not explicitly forbidden is permitted, the dictatorial one considers that every situation that is not explicitly permitted is forbidden.

Let $\mathcal{M} = (S, V, \mathcal{R})$ be an epistemic model and PP be a privacy policy, we construct the liberal model $\mathcal{M}^l_{PP} = (S, V, \mathcal{R}, \mathcal{P}^l)$ and the dictatorial one $\mathcal{M}^d_{PP} = (S, V, \mathcal{R}, \mathcal{P}^d)$ in the following way: for all agent $i \in AG$,

$$\mathcal{P}^l_i = \{(s, \mathcal{R}_1, \mathcal{R}_2) \mid \forall (pre \longrightarrow \neg P_i\psi) \in PP, \text{ if } \mathcal{M}_{\mathcal{R}_1}, s \models pre \text{ then } \mathcal{M}_{\mathcal{R}_2}, s \not\models \psi\}$$

$$\mathcal{P}_i^d = \{(s, \mathcal{R}_1, \mathcal{R}_2) \mid \exists (pre \longrightarrow P_i\psi) \in PP, \mathcal{M}_{\mathcal{R}_1}, s \models pre \text{ and } \mathcal{M}_{\mathcal{R}_2}, s \models \psi\}$$

8.2 Dynamic Epistemic Logic with Permission

Another possible further work would be to make a step forward in the way developed in chapters 6 and 7. Indeed, we first considered only public announcements, and then a kind of private announcements – that include public ones – in which the topic of the message and the agents involved in the communication are publicly known. The first further step in this work would thus be to avoid this limit by considering permission over other kinds of private communication. In this sense, having a general framework for permission and obligation over every epistemic action (cf. Section 3.1.3) would be a final step. How could we obtain such a framework?

First, given an event model $\mathcal{A}$, the language we may propose would extend action model logic with additive operators $P_i a$ and $O_i a$ standing for 'i is allowed to make the announcement a' and 'i is obliged to make the announcement a', where $a \in \mathcal{A}$. The models would be of the form $\mathcal{M} = (S, \mathcal{R}, V, \mathcal{P})$ where $\mathcal{P}$ is constituted by triples (s, E_1, E_2). In such triples, E_1 would represent the epistemic situation after a first announcement and E_2 the situation after a second one. Note that in this case, epistemic situations are not necessarily submodels of the initial model as in chapter 6, and neither are they represented by subrelations as in chapter 7. They can be state models of any kind. Therefore we could imagine the following semantics for the new operators:

$\mathcal{M}, s \models Pa$ iff $\mathcal{M}, s \models pre(a)$ implies there exists N_1, N_2 such that $(s, N_1, N_2) \in \mathcal{P}$, $N_1 \leftrightarrows \mathcal{M} \otimes \mathcal{A}, (s, a)$ and $N_2 \leftrightarrows \mathcal{M} \otimes \mathcal{A}, (s, a; b)$

It would be quite a complex semantics. Indeed, to verify if an announcement is permitted we would have to test the bisimilarity of the resulting state model with the ones appearing in $\mathcal{P}$. But avoiding every restriction on the structure of epistemic events, it is impossible to fix the structure of the resulting state model. Bisimilarity is in this context the good notion to test the epistemic result of an announcement.

Note that we would need to define the $\otimes$-closure of an event model $\mathcal{A}$, to ensure that if a and b are in $\mathcal{A}$, then $(a; b) \in \mathcal{A}$. But such a definition (and other useful notions) are feasible. We do not go further in this presentation, which is still a work in progress, but we think that it could be a good starting point for a generalization of the frameworks presented in chapters 6 and 7.

List of Figures

1 Un modèle booléen des enfants lunatiques 2
2 Enfants lunatiques: tous les mondes possibles 3
3 Modèle épistémique 7
4 Un autre modèle épistémique 7
5 Quelques exemples de modèles de Kripke 8
6 Une instance du problème du pavage 20

1.1 A boolean model for the moody children 24
1.2 Moody children: all the possible worlds 25

2.1 Epistemic model 29
2.2 Epistemic model 29
2.3 Some examples of Kripke models 30
2.4 An instance of the domino problem 41

3.1 Texas Hold'em 43
3.2 Deal of three cards between three agents 45
3.3 Alex announces: "*I don't have the 1*" 46
3.4 Cha shows her card to Brune 50
3.5 Result of the event 50
3.6 Brune actions while receiving a letter from Cha 58
3.7 Considering states of violation for Brune's actions 61
3.8 Considering labelled transition for Brune's actions 62

4.1 Possible models $\mathcal{M}_{\psi_A}$ for $A = \exists p_1 \forall p_2 \forall p_3 A'$ 70
4.2 Action model equivalent to the update $\mathcal{U}_{\varphi,G}(\mathcal{M})$ for objective beliefs 72
4.3 Two counter-examples in one model 83
4.4 Distinguishing $\mathcal{L}_{lauob}$ from $\mathcal{L}_{luob}$ 89
4.5 Brune wins with a "poker" 90
4.6 Two deals: d_1 and d_2 91

5.1 Counter-example of the validity of MacKinsey formula in GAL 101
5.2 Distinguishing $\mathcal{L}_{gal}$ from $\mathcal{L}_{el}$ 109
5.3 Distinguishing $\mathcal{L}_{apal}$ from $\mathcal{L}_{gal}$ 110
5.4 Trying to distinguish $\mathcal{L}_{gal}$ from $\mathcal{L}_{apal}$ 112
5.5 A model to prove the hardness of $MC(GAL)$ 116
5.6 Three moody children playing in the *Russian cards* problem 124

6.1 The moody children playing *la belote* 152
6.2 Rules $(R\sigma)$, $(R4)$ and $(RP2)$ 167
6.3 The output of the tableau method for $PK_1p \wedge PK_1q \wedge \neg PK_1(p \wedge q) \wedge \hat{K}_2 PK_1(p \wedge q) \wedge [K_1p]PK_1(p \wedge q)$ 168

7.1 AIDS example 189

8.1 server example 196

Bibliography

[Ågotnes, 2006] Ågotnes, T. (2006). Action and knowledge in alternating-time temporal logic. *Synthese (Special Section on Knowledge, Rationality and Action)*, 149(2):377–409.

[Ågotnes et al., 2010] Ågotnes, T., Balbiani, P., van Ditmarsch, H., and Seban, P. (2010). Group announcement logic. *Journal of Applied Logic*, 8:62–81.

[Alur et al., 2002] Alur, R., Henzinger, T. A., and Kupferman, O. (2002). Alternating-time temporal logic. *Journal of the ACM*, 49:672–713.

[Aucher et al., 2010] Aucher, G., Boella, G., and van der Torre, L. (2010). Privacy policies with modal logic: The dynamic turn. In *DEON*, pages 196–213.

[Balbiani et al., 2007] Balbiani, P., Baltag, A., van Ditmarsch, H., Herzig, A., Hoshi, T., and Lima, T. D. (2007). What can we achieve by arbitrary announcements? A dynamic take on Fitch's knowability. In *Proceedings of TARK XI*, pages 42–51. Presses Universitaires de Louvain.

[Balbiani et al., 2008] Balbiani, P., Baltag, A., van Ditmarsch, H., Herzig, A., Hoshi, T., and Lima, T. D. (2008). 'Knowable' as 'known after an announcement'. *Review of Symbolic Logic*, 1(3):305–334.

[Balbiani et al., 2010] Balbiani, P., Ditmarsch, H. v., Herzig, A., and Lima, T. d. (2010). Tableaux for public announcement logics. *Journal of Logic and Computation*, 20(1):55–76.

[Balbiani and Seban, 2008] Balbiani, P. and Seban, P. (2008). Logique de la mise à jour des croyances objectives. GDR I3 - JIAF'08.

[Balbiani and Seban, 2011] Balbiani, P. and Seban, P. (2011). Reasoning about permitted announcements. *Journal of Philosophical Logic*, Special issue about Logical Methods for Social Concepts. Published Online First: http://www.springerlink.com/content/g215033h036441v5/.

[Baltag and Moss, 2004] Baltag, A. and Moss, L. (2004). Logics for epistemic programs. *Synthese*, 139:165–224. Knowledge, Rationality & Action 1–60.

[Baltag et al., 1998] Baltag, A., Moss, L., and Solecki, S. (1998). The logic of public announcements, common knowledge, and private suspicions. In Gilboa, I., editor, *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 98)*, pages 43–56.

[Blackburn et al., 2001] Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge University Press, Cambridge. Cambridge Tracts in Theoretical Computer Science 53.

[Boh, 1993] Boh, I. (1993). *Epistemic Logic in the Later Middle Ages*. Routledge, UK.

[Broersen, 2008] Broersen, J. (2008). A complete stit logic for knowledge and action, and some of its applications. In *Proceedings of DALT@AAMAS08*, Lecture Notes in Artificial Intelligence. Springer.

[Brogaard and Salerno, 2004] Brogaard, B. and Salerno, J. (2004). Fitch's paradox of knowability. `http://plato.stanford.edu/archives/sum2004/entries/fitch-paradox/`.

[Burnyeat and Barnes, 1980] Burnyeat, M. and Barnes, J. (1980). Socrates and the jury: Paradoxes in plato's distinction between knowledge and true belief. Blackwell Publishing.

[Castaneda, 1981] Castaneda, H.-N. (1981). The paradoxes of deontic logic: the simplest solution to all of them in one fell swoop. *Synthese library*, pages 37–86.

[Chandra et al., 1981] Chandra, A., Kozen, D., and Stockmeyer, L. (1981). Alternation. *Journal of the ACM*, 28:114–33.

[Chellas, 1980] Chellas, B. (1980). *Modal Logic, an Introduction*. Cambridge, London.

[Chisholm, 1963] Chisholm, R. (1963). Contrary-to-duty imperatives and deontic logic. *Analysis*, (24):33 – 36.

[Clarke et al., 1999] Clarke, E., Grumberg, O., and Peled, D. (1999). *Model checking*. MIT Press, Cambridge, MA, USA.

[Cuppens et al., 2005a] Cuppens, F., Cuppens-Boulahia, N., and Sans, T. (2005a). Nomad: A security model with non atomic actions and deadlines. In *Proceedings of the 18th IEEE workshop on Computer Security Foundations*, pages 186–196. IEEE Computer Society, Washington, DC, USA.

[Cuppens et al., 2005b] Cuppens, F., Cuppens-Boulahia, N., Sans, T., and Miège, A. (2005b). A formal approach to specify and deploy a network security policy. In *Formal Aspects in Security and Trust*. Springer Boston.

[Dubois and Prade, 1988] Dubois, D. and Prade, H. (1988). *Possibility Theory : an Approach to Computerized Processing of Uncertainty*. Plenum Press.

[Fagin et al., 1995] Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (1995). *Reasoning about Knowledge*. MIT Press, Cambridge MA.

[Fischer and Ladner, 1979] Fischer, M. and Ladner, R. (1979). Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18(2):194–211.

[Fitch, 1963] Fitch, F. (1963). A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28(2):135–142.

[French and van Ditmarsch, 2008] French, T. and van Ditmarsch, H. (2008). Undecidability for arbitrary public announcement logic. In *Advances in Modal Logic 7*, pages 23–42. College Publications, London.

[Gasquet et al., 2005] Gasquet, O., Herzig, A., Longin, D., and Saade, M. (2005). LoTREC: Logical tableaux research engineering companion. In *International Conference TABLEAUX 2005*, pages 318–322. Springer Verlag, Berlin.

[Goldblatt, 1982] Goldblatt, R. (1982). *Axiomatising the Logic of Computer Programming*. Springer-Verlag, Berlin.

[Gradel and Otto, 1999] Gradel, E. and Otto, M. (1999). On logics with two variables. *Theoretical Computer Science*, 224:73–113.

[Groenendijk and Stokhof, 1997] Groenendijk, J. and Stokhof, M. (1997). Questions. In *Handbook of Logic and Language*, pages 1055–1124. Elsevier.

[Gödel, 1951] Gödel, K. (1951). Some basic theorems on the foundations of mathematics and their implications (1951). In *Collected works / Kurt Gödel*, volume III, pages 304–323. Clarendon Press ; New York.

[Halpern and Moses, 1992] Halpern, J. Y. and Moses, Y. (1992). A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54:319–379.

[Hintikka, 1962] Hintikka, J. (1962). *Knowledge and Belief*. Cornell University Press, Ithaca, NY.

[Hommersom et al., 2004] Hommersom, A., Meyer, J.-J., and Vink, E. D. (2004). Update semantics of security protocols. *Synthese*, 142(2):229–267.

[Horty, 2001] Horty, J. (2001). *Agency and Deontic Logic*. Oxford University Press, NewYork.

[Hoshi, 2008] Hoshi, T. (2008). Logics of public announcements with announcement protocols. Manuscript. Philosophy Department, Stanford University.

[Jamroga, 2003] Jamroga, W. (2003). Some remarks on alternating temporal epistemic logic. In *FAMAS'03 – Formal Approaches to Multi-Agent Systems, Proceedings*, pages 133–140, Warsaw, Poland.

[Jamroga and Ågotnes, 2007] Jamroga, W. and Ågotnes, T. (2007). Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non-Classical Logics*, 17(4):423–475.

[Jamroga and van der Hoek, 2004] Jamroga, W. and van der Hoek, W. (2004). Agents that know how to play. *Fundamenta Informaticae*, 63:185–219.

[Lutz, 2006] Lutz, C. (2006). Complexity and succinctness of public announcement logic. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 06)*, pages 137–144.

[Mally, 1926] Mally, E. (1926). Grundgesetze des sollens: Elemente der logik des willens. *Leuschner and Lubensky.*

[Meyer, 1988] Meyer, J.-J. (1988). A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136.

[Meyer et al., 1994] Meyer, J.-J., Dignum, F., and Wieringa, R. (1994). The paradoxes of deontic logic revisited: A computer science perspective. Technical report, Department of Computer Science, Utrecht University.

[Papadimitriou, 1994] Papadimitriou, C. (1994). *Computational Complexity*. Reading-MA.

[Parikh and Ramanujam, 2003] Parikh, R. and Ramanujam, R. (2003). A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12:453–467.

[Pauly, 2002] Pauly, M. (2002). A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166.

[Plato, BC] Plato (BC). *Theaetetus.*

[Plaza, 1989] Plaza, J. (1989). Logics of public communications. In *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems: Poster Session Program*, pages 201–216. Oak Ridge National Laboratory.

[Pucella and Weissman, 2004] Pucella, R. and Weissman, V. (2004). Reasoning about dynamic policies. In *FOSSACS 2004*, pages 453–467. Springer, Berlin.

[Ross, 1941] Ross, A. (1941). Imperatives and logic. *Theoria*, 11(1):30–46.

[Said, 2010] Said, B. (2010). *Réécriture de graphes pour la construction de modèles en logique modale.* Thèse de doctorat, Université de Toulouse.

[Spade, 1982] Spade, P. (1982). Three theories of obligationes: Burley, kilvington and swyneshed on counterfactual reasoning. *History and Philosophy of Logic, 3*, pages 1–32.

[Spade, 1992] Spade, P. (1992). If obligationes were counterfactuals. *Philosophical Topics*, 20:171–188.

[Stockmeyer and Meyer, 1973] Stockmeyer, L. and Meyer, A. (1973). Word problems requiring exponential time(preliminary report). In *STOC '73: Proceedings of the Fifth Annual ACM Symposium on Theory of computing*, pages 1–9. ACM, New York.

[van Benthem, 2004] van Benthem, J. (2004). What one may come to know. *Analysis*, 64(2):95–105.

[van Benthem et al., 2009] van Benthem, J., Gerbrandy, J., Hoshi, T., and Pacuit, E. (2009). Merging frameworks for interaction. *Journal of Philosophical Logic*, 38:491–526.

[van der Meyden, 1996] van der Meyden, R. (1996). The dynamic logic of permission. *Journal of Logic and Computation*, 6(3):465–479.

[van Ditmarsch, 2003] van Ditmarsch, H. (2003). The Russian cards problem. *Studia Logica*, 75:31–62.

[van Ditmarsch and Kooi, 2006] van Ditmarsch, H. and Kooi, B. (2006). The secret of my success. *Synthese*, 151:201–232.

[van Ditmarsch et al., 2007] van Ditmarsch, H., van der Hoek, W., and Kooi, B. (2007). *Dynamic Epistemic Logic*, volume 337 of *Synthese Library*. Springer.

[van Ditmarsch et al., 2009] van Ditmarsch, H., van Eijck, J., and Verbrugge, R. (2009). Common knowledge and common belief. In van Eijck, J. and Verbrugge, R., editors, *Discourses on Social Software*, Texts in Games and Logic (Vol. 5). Amsterdam University Press.

[von Wright, 1951] von Wright, G. (1951). *An Essay in Modal Logic*. North Holland, Amsterdam.

[Wang, 1961] Wang, H. (1961). Proving theorems by pattern recognition ii. *Bell System Technical Journal*, 40:1–41.

[Wang et al., 2009] Wang, Y., Kuppusamy, L., and van Eijck, J. (2009). Verifying epistemic protocols under common knowledge. In *TARK '09: Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 257–266, New York, NY, USA. ACM.