KNOWLEDGE, CHANCE, AND CHANGE

ILLC Dissertation Series DS-2003-01

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@wins.uva.nl
homepage: http://www.illc.uva.nl/

RIJKSUNIVERSITEIT GRONINGEN

KNOWLEDGE, CHANCE, AND CHANGE

PROEFSCHRIFT

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
vrijdag 5 september 2003
om 16.00 uur

door

Barteld Pieter Kooi

geboren op 28 augustus 1976
te Groningen

Promotor: prof.dr. G.R. Renardel de Lavalette
Co-promotor: dr. L.C. Verbrugge

Beoordelingscommissie: prof.dr. J.F.A.K. van Benthem
prof.dr. W. van der Hoek
prof.dr. E.C.W. Krabbe

ISBN: 90–367–1786–8

Contents

Acknowledgments

I thank Gerard Renardel for giving me the opportunity to write this thesis and his supervision. His door was always open if I wanted to discuss something. I thank Rineke Verbrugge for her supervision, which she always made time for despite the growing pressures of her busy job. I thank both of them for doing research with me on strong completeness and disharmony.

I thank Johan van Benthem for his support, for his suggestions, for his invitation to Stanford, and his interest in my work during the past four years. I thank Wiebe van der Hoek for his support, for his suggestions, for the detailed comments he wrote as a member of the reading committee, for doing research together, for his invitation to Liverpool, and for the wonderful times in Great Britain and New Zealand. I thank Erik Krabbe for his very detailed comments on the manuscript he provided as a member of the reading committee.

I thank Hans van Ditmarsch for his support during the past four years, for traveling around the world and inviting me to all the places he went, especially for inviting me to New Zealand, where I finished the first complete version of the manuscript (thanks to his encouragements), for doing research with me, and most of all for being a wonderful friend.

I thank the Netherlands Organisation for Scientific Research (NWO) and the Institute for Mathematics and Computing Science of the University of Groningen (IWI) for their financial support for my trip to New Zealand.

I thank Casper Albers and Willem Schaafsma for doing research with me on the two-envelope paradox.

I thank Robert van Rooy for doing research on questions with me and for trying to make a special issue with me.

I thank Roland Auer, Wim Hesselink, Erik Krabbe, Theo Kuipers, Wim Oudshoorn, Jeanne Peijnenburg, Jan Willem Romeyn for their comments on papers that later became parts of this thesis.

I thank Maint Berkenbosch, Edith Coenen, Dirk Jan Kort, Esther Moet, Lucian Muresan, and Simon Plantinga for being wonderful roommates.

I thank Roland Auer, Maint Berkenbosch, Henk Bruin, Robert Carls, Marc Dröge, Aletta Eikelboom, Joost le Feber, Geert Fekken, Jeroen Gerrits, Hendrik Wietze de Haan, Gert Jan van der Heiden, Theresa Helmholt-Kleefsman, Jun Hoo, Remke Kloosterman, Robert Kuik, Nico Kruithof, Erwin Loots, Theo Dirk Meijler, Arie de Niet, Martijn van Noort, Conny Pott, Marius van der Put, Lenny Taelman, Ena Tiesinga, Jaap Top, Evgeny Verbitskiy, and Peter Zwerver for lunches, tea breaks, and other breaks at the IWI. I thank Willem Labuschagne for coffee breaks, and lunches in Dunedin where we had some wonderful discussions.

I thank Joost le Feber, Hendrik Wietze de Haan, Wim Hesselink, Jan Jongejan, Jan Eppo Jonker, Conny Pott, Gerard Renardel, Jan Terlouw, for the weekly meetings of the Fundamental Computing group and for the occasional stroopwafel that was served during these meetings.

I thank Alle Meije Wink for co-organizing the AiO colloquium. I thank Jan Albert van Laar for co-organizing the grolog lectures.

I thank Gerlof Bouma, Taeda Jovicic, Erik Krabbe, Theo Kuipers, Jan Albert van Laar, Fiore Nicolai, Anna Pilatova, Gerard Renardel, Wouter Teepe, Jan Willem Romeyn, Rineke Verbrugge for the existence of grolog.

I thank Peter Arendz, Annemieke Beerenboom, Jurjen Bokma, Esmee Elshof, Desiree Hansen, Ineke Kruizinga, and Harm Paas for their technical and moral support.

I thank Fiona Douma and Theo de Vries for the inspiration the supervision of their master theses gave me.

I thank Boudewijn de Bruin, Balder ten Cate, Francien Dechesne, Paul Dekker, Sjoerd Druiven, Paul Harrestein, Rosja Mastop, Marc Pauly, and Robert van Rooy for being great company at many ILLC, OZSL, and related events.

I thank Harm Bakker, Jur van den Berg, Jasper de Boer, Meinte Boersma, Hendrik Wietze de Haan, Anton Jansen, Jan Eppo Jonker, Arend Smit for the great times organizing the IWI programming contests in 2002 and 2003.

I thank my friends for great times.

I thank my parents, and my brother for their love and support.

I thank everyone I forgot to mention above to whom I owe gratitude for not being offended by that.

Barteld P. Kooi
Groningen, July 2003

Chapter 1

Introduction

This thesis is about the logic of knowledge, chance and change. In logic one is engaged with inferences and one tries to answer the question whether an inference is correct. One does this by looking at the abstract form of inferences. Logical languages are often tailored to inferences where a certain concept is paramount. All sorts of logics have been developed, each with their own application areas. In this thesis three logics are studied: epistemic logic, dynamic logic, and probabilistic logic. A chapter is dedicated to each of them. There is a chapter on the combination of dynamic logic and epistemic logic. The central chapter of this thesis is dedicated to the combination of all three of these logics. With this logic one can study inferences about probabilistic information change.

In epistemic logic one focuses on inferences where knowledge (or more generally information) plays a crucial role. Reasoning about knowledge is particularly interesting in situation involving more than one individual. In that case someone can know whether someone else knows something or does not know something. Chapter 2 provides a short introduction to epistemic logic, without any new results. Readers familiar with epistemic logic can skip this chapter, or skim through it to get acquainted with my notation.

Dynamic logic is the logic of change. It is mostly applied to changes that occur due to the execution of computer programs. In chapter 3 some technical results about propositional dynamic logic are presented. Propositional dynamic logic is not compact, and therefore a finitary proof system can never be strongly complete. In chapter 3 a strongly complete proof system is presented and a straightforward completeness proof is provided. A curious property of the canonical model, called disharmony, is also examined. This chapter is the result of joint work with Gerard Renardel and Rineke Verbrugge, which has already resulted in a publication (see Renardel de Lavalette, Kooi, and Verbrugge (2002)).

Epistemic logic and dynamic logic can be combined to study inferences about information change. The main part of chapter 4 is dedicated to providing an overview of all dynamic epistemic logics, but a new system is also presented.

These logics differ from other approaches to information change in that higher-order information is explicitly taken into account.

In probabilistic logic inferences about probability are studied. There are many different philosophies about the nature of probability. In chapter 5 I investigate the relationship between two probabilistic logics, that have arisen from two different notions of probability.

In chapter 6 I develop a probabilistic dynamic epistemic logic, which combines the logics that were treated in the earlier chapters. It is suited to analyze inferences about the probabilistic information change, especially when higher-order probabilities play a role. The paper on which this article is based will appear later this year (see Kooi (2003)).

Once these logics are presented some problems in the area of knowledge, chance and change are explicitly discussed: the Monty Hall dilemma is analyzed in chapter 7, the game Mastermind is examined in chapter 8, and the two envelope paradox discussed in chapter 9.

Finally in chapter 10 I draw some conclusions.

Chapter 2

Epistemic logic

2.1 Introduction

In *Knowledge and Belief* Jaakko Hintikka (1962), for the first time, described knowledge in terms of possible worlds. This research area has become known as epistemic logic. It stems from the Greek word for knowledge: ἐπιστήμη. In this thesis I take the term "epistemic" broader, applying it to belief and other ways an agent might have information as well. This concurs with much of the literature in this area. In this chapter I give an outline of epistemic logic. The notions that are explained in this chapter are used throughout this thesis. This chapter is a rather brief introduction to the subject. For a more extensive introduction see Fagin, Halpern, Moses, and Vardi (1995) or Meyer and Van der Hoek (1995).

2.2 Language and semantics

Epistemic logic can be used to model the information agents have about the world, but it is especially suited to model the information agents have about each other's information. Suppose for example that there is a situation involving two agents, a and b. If the proposition 'p' means 'it is raining', then 'a knows it is raining' can be formalized as '$\Box_a p$'. The subscript indicates that we are concerned with a's knowledge. The sentence 'b knows that a knows it is raining' can be formalized as '$\Box_b \Box_a p$'.

Definition 2.1 (Language of epistemic logic $\mathscr{L}_{\mathcal{PA}}$)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of epistemic logic $\mathscr{L}_{\mathcal{PA}}$ is given by the following rule in Backus-Naur Form (BNF):

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid \Box_a \varphi$$

where $p \in \mathcal{P}$ and $a \in \mathcal{A}$. Moreover $\top$ is an abbreviation for $\neg\bot$, $(\varphi \vee \psi)$ is an abbreviation for $\neg(\neg\varphi \wedge \neg\psi)$, $(\varphi \rightarrow \psi)$ is an abbreviation for $(\neg\varphi \vee \psi)$ and $(\varphi \leftrightarrow \psi)$ is an abbreviation for $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$. $\Diamond_a\varphi$ is an abbreviation for $\neg\Box_a\neg\varphi$. I will also use the convention to omit the outermost parentheses of a sentence. □

Epistemic logic is a modal logic. The standard semantics for epistemic logic does not differ much from standard semantics for modal logic. The only difference is that there is an accessibility relation for every agent.

Definition 2.2 (Epistemic models)
An epistemic model for $\mathscr{L}_{\mathcal{PA}}$ is triple $M = (W, R, V)$ such that:

- $W \neq \emptyset$; a set of states or possible worlds;
- $R : \mathcal{A} \rightarrow 2^{W \times W}$; assigns an accessibility relation to each agent;
- $V : \mathcal{P} \rightarrow 2^{W}$; assigns a set of possible worlds to each propositional variable.

If $M = (W, R, V)$ is a model, a pair (M, w), where $w \in W$ is called a pointed model. A pair $F = (W, R)$ is called a frame and a pair (F, w), where $w \in W$, is called a pointed frame. A pair (F, V) is also a model. But I will be somewhat sloppy with the terminology. □

The class of all frames is called $K_{\mathcal{A}}$. The class of all pointed frames is called $*K_{\mathcal{A}}$. The class of all models for $\mathscr{L}_{\mathcal{PA}}$ is called $K_{\mathcal{PA}}$, and $*K_{\mathcal{PA}}$ is the class of all pointed models for that language.

Definition 2.3 (Semantics for $\mathscr{L}_{\mathcal{PA}}$)
Let an epistemic model (M, w) where $M = (W, R, V)$ be given. Let $p \in \mathcal{P}$, $a \in \mathcal{A}$, and $\varphi, \psi \in \mathscr{L}_{\mathcal{PA}}$.

$$
\begin{array}{lll}
(M, w) \not\models \bot & & \\
(M, w) \models p & \text{iff} & w \in V(p) \\
(M, w) \models \neg\varphi & \text{iff} & (M, w) \not\models \varphi \\
(M, w) \models (\varphi \wedge \psi) & \text{iff} & (M, w) \models \varphi \text{ and } (M, w) \models \psi \\
(M, w) \models \Box_a\varphi & \text{iff} & (M, v) \models \varphi \text{ for all } v \text{ such that } wR(a)v
\end{array}
$$

where $wR(a)v$ is an abbreviation for $(w, v) \in R(a)$. □

Consider the following example. Suppose two children, a and b, have been playing outside. Both of them can see whether the other child's face is muddy, yet they cannot see their own faces. This situation can be analyzed using an epistemic logic by making a model for this situation. A picture of this model is shown in figure 2.1. The states are indicated by pairs (x, y), where x and y stand for the state a's respectively b's face is in. The number 0 means the child's face is not

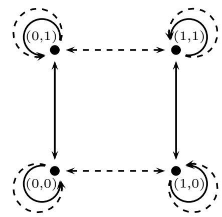


Figure 2.1: A Kripke model. The solid lines represent the accessibility relation assigned to a. The dashed lines represent the accessibility relation assigned to b.

muddy, 1 means it is muddy. The solid and dashed lines represent the accessibility relations of a and b respectively. In $(0,1)$ for example, a cannot rule out her face is muddy, but she also cannot rule out she is not muddy (which is actually the case). Therefore $(0,1)$ and $(1,1)$ are both accessible to a. She does know that b's face is muddy. Therefore in both worlds that are accessible to her b's face is muddy. This example is a specific instance of the initial situation of the muddy children puzzle, which is discussed in chapter 4 (page 35).

These so-called Kripke models have an important property. They do not only model what information the agents have about the world, they also model 'higher-order information', i.e. the information the agents have about the information that the agents have, and so on. For example, in the Kripke model above, a knows that b does not know whether she is muddy or not. But it does not stop there, because the model also gives us that b knows that a does not know that b knows that a is muddy. In this way the model allows us to stack these kinds of constructions indefinitely, and thus it models all higher-order information at once.

For the notion of validity we overload $\models$ with the following notions.

Definition 2.4 (Validities)
Let $M = (W, R, V)$ be a model, $F = (W, R)$ be a frame $w \in W$ be a world. Let $S_{\mathcal{PA}}$ and $*S_{\mathcal{PA}}$ be classes of models and pointed models respectively.

$$
\begin{array}{lll}
\models_{(M,w)} \varphi & \text{iff} & (M,w) \models \varphi \\
\models_{*S_{\mathcal{PA}}} \varphi & \text{iff} & \models_{(M,w)} \varphi \text{ for every } (M,w) \in *S_{\mathcal{PA}} \\
\models_{M} \varphi & \text{iff} & \models_{(M,w)} \varphi \text{ for every } w \in W \\
\models_{S_{\mathcal{PA}}} \varphi & \text{iff} & \models_{M} \varphi \text{ for every } M \in S_{\mathcal{PA}} \\
\models_{(F,w)} \varphi & \text{iff} & \models_{((F,V),w)} \varphi \text{ for every } V : \mathcal{P} \to 2^W \\
\models_{*S_{\mathcal{A}}} \varphi & \text{iff} & \models_{(F,w)} \varphi \text{ for every } (F,w) \in *S_{\mathcal{A}} \\
\models_{F} \varphi & \text{iff} & \models_{(F,w)} \varphi \text{ for every } w \in W \\
\models_{S_{\mathcal{A}}} \varphi & \text{iff} & \models_{F} \varphi \text{ for every } F \in S_{\mathcal{A}}
\end{array}
$$

Generally by $\models \varphi$ we mean $\models_{*S_{\mathcal{PA}}} \varphi$, and by $\Gamma \models \varphi$ we mean local logical consequence, i.e. for every pointed model $(M, w) \in *S_{\mathcal{PA}}$, if $(M, w) \models \psi$ for every $\psi \in \Gamma$, then $(M, w) \models \varphi$. $\square$

Whether these notions coincide depends on the choice of models, frames, and classes of these.

2.3 Proof systems

The simplest proof system for epistemic logic is $\mathsf{K}_{\mathcal{PA}}$.

Definition 2.5 (Proof system $\mathsf{K}_{\mathcal{PA}}$)
Let φ, ψ be sentences in $\mathscr{L}_{\mathcal{PA}}$ and let a be an agent in $\mathcal{A}$. The proof system $\mathsf{K}_{\mathcal{PA}}$ consists of the following axioms and derivation rules.

Taut	all instantiations of propositional tautologies	
Distr	$\Box_a(\varphi \to \psi) \to (\Box_a\varphi \to \Box_a\psi)$	(distribution)
MP	$\dfrac{\varphi \qquad \varphi \to \psi}{\psi}$	(modus ponens)
Nec	$\dfrac{\varphi}{\Box_a\varphi}$	(necessitation)

Let us now introduce a general notion of provability that can be used for other proof systems as well.

Definition 2.6 (Provability)
Let S be a proof system and $\mathscr{L}$ a logical language. A derivation or proof in S consists of a sequence of sentences of $\mathscr{L}$ each of which is an instance of an axiom or is the result of applying a derivation rule to sentences that occur earlier in the sequence. If φ is the last sentence in a derivation, then φ is provable, or deducible, in S, notation $\vdash_{\mathsf{S}} \varphi$. $\square$

Note that this notion of provability excludes the possibility of premises of an inference. The rules **MP** and **Nec** can only be applied when the formulas above the line have already been deduced.

The system $\mathsf{K}_{\mathcal{PA}}$ is sound and complete with respect to the class of models $K_{\mathcal{PA}}$. That is the notion of validity and deducibility coincide.

Theorem 2.1 (Soundness and completeness of $\mathsf{K}_{\mathcal{PA}}$)

$$\vdash_{\mathsf{K}_{\mathcal{PA}}} \varphi \text{ iff } \models_{*K_{\mathcal{PA}}} \varphi$$

for every sentence $\varphi \in \mathscr{L}_{\mathcal{PA}}$. The *if* above is called completeness, the *only if* is called soundness. $\square$

Soundness is often easy to prove by induction on the length of the proof. Completeness is usually shown by contraposition and involves the construction of a canonical model. The notion of completeness presented here is called *weak* completeness. For strong completeness, see chapter 3.

Epistemic logic is used for many notions involving information. The axioms and rules of $\mathsf{K}_{\mathcal{PA}}$, are minimal requirements these notions should meet. For many epistemic notions there are more requirements that should be met. For example in the case of knowledge, you want that if an agent knows something, then it is true:

$$\mathbf{T} \quad \Box_a\varphi \to \varphi \quad \text{(factivity)}$$

There are two systems that are especially relevant for epistemic logic, viz. $\mathsf{KD45}_{\mathcal{PA}}$ and $\mathsf{S5}_{\mathcal{PA}}$[1]. The system $\mathsf{KD45}_{\mathcal{PA}}$ is usually considered to be the best system to model belief, and $\mathsf{S5}_{\mathcal{PA}}$ is considered to be the best system to model knowledge.

$$\mathbf{D} \quad \Box_a\varphi \to \Diamond_a\varphi$$
$$\mathbf{4} \quad \Box_a\varphi \to \Box_a\Box_a\varphi \quad \text{(positive introspection)}$$
$$\mathbf{5} \quad \Diamond_a\varphi \to \Box_a\Diamond_a\varphi \quad \text{(negative introspection)}$$

The system $\mathsf{S5}_{\mathcal{PA}}$ consists of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}$ and the axioms **T**, **4**, and **5** for sentences of $\mathscr{L}_{\mathcal{PA}}$. The system $\mathsf{KD45}_{\mathcal{PA}}$ consists of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}$ and the axioms **D**, **4**, and **5** for sentences of $\mathscr{L}_{\mathcal{PA}}$.

By a class of pointed models $*S_{\mathcal{PA}}$ that is intended to be complete with respect to a proof system S, the following is meant. Take the subclass of frames $*S_{\mathcal{A}}$ such that all the axioms and the rules of S are valid in those models. Then take the class of pointed models $*S_{\mathcal{PA}}$ associated with $*S_{\mathcal{A}}$[2]. It turns out that the axioms of $\mathsf{KD45}_{\mathcal{PA}}$ are valid on the class of serial, transitive, and euclidean frames. The axioms of $\mathsf{S5}_{\mathcal{PA}}$ are valid on those frames where the accessibility relations are all equivalence relations.

2.4 General and common knowledge

There are two more important concepts in epistemic logic: general knowledge and common knowledge. If something is general knowledge, it means that everybody knows it. Common knowledge is a typical concept for epistemic logic. It is concerned with information about information. If something is common knowledge, then it is general knowledge, but it is also general knowledge that it is general

[1] The systematic name of $\mathsf{S5}_{\mathcal{PA}}$ would be $\mathsf{KT45}_{\mathcal{PA}}$, however this name is not commonly used. Moreover there are more systems that would have the same set of valid sentences. For example $\mathsf{KT5}_{\mathcal{PA}}$ or $\mathsf{KDB5}_{\mathcal{PA}}$ or $\mathsf{KTB4}_{\mathcal{PA}}$. So I choose for $\mathsf{S5}_{\mathcal{PA}}$.

[2] Sometimes completeness is proved with respect to a smaller class of models than those where the axioms hold. For example the logic of linear time is valid on all linear frames, but also on other frames.

knowledge, and it is also general knowledge that it is general knowledge that it is general knowledge, and so on *ad infinitum*. Common knowledge is a very useful concept, especially in the context of games.

To incorporate these two notions in the formal language two modal operators have to be added; $E_{\mathcal{B}}\varphi$, which can be read as 'every member of group $\mathcal{B}$ knows that φ' and $C_{\mathcal{B}}\varphi$, which can be read as 'it is common knowledge among members of $\mathcal{B}$ that φ.' Let $\mathscr{L}^{EC}_{\mathcal{PA}}$ be the language extended with these two operators. The models do not have to be changed or adapted to interpret sentences in which these operators occur.

Definition 2.7 (Semantics for *E* and *C*)
Let a model (M, w) where $M = (W, R, V)$ be given. Let $p \in \mathcal{P}$, $\varphi \in \mathscr{L}^{EC}_{\mathcal{PA}}$, and $\mathcal{B} \subseteq \mathcal{A}$.

$$\begin{array}{lll} (M,w) \models E_{\mathcal{B}}\varphi & \text{iff} & (M,v) \models \varphi \text{ for all } v \text{ such that } wR(\mathcal{B})v \\ (M,w) \models C_{\mathcal{B}}\varphi & \text{iff} & (M,v) \models \varphi \text{ for all } v \text{ such that } wR(\mathcal{B})^{+}v \end{array}$$

where $R(\mathcal{B}) = \bigcup_{a\in\mathcal{B}} R(a)$, and $R(\mathcal{B})^{+}$ is the transitive closure of this relation.

In this definition the accessibility relation corresponding to the common knowledge operator is interpreted as the transitive closure of the union of the accessibility relations of the agents in the group. In other approaches it is taken to be the reflexive transitive closure of this relation. However, this is not practical when other epistemic notions such as belief are studied. Nevertheless, when we are working within the class $S5_{\mathcal{PA}}$ it is the reflexive transitive closure.

For the proof system, we need two additional axioms and an additional rule.

$$\begin{array}{lll} \mathbf{E} & E_{\mathcal{B}}\varphi \leftrightarrow \bigwedge_{a\in\mathcal{B}} \Box_a\varphi & \\ \mathbf{Mix} & C_{\mathcal{B}}\varphi \rightarrow E_{\mathcal{B}}(\varphi \wedge C_{\mathcal{B}}\varphi) & \\ \mathbf{Ind} & \dfrac{\varphi \rightarrow E_{\mathcal{B}}(\psi \wedge \varphi)}{\varphi \rightarrow C_{\mathcal{B}}\psi} & \text{(induction rule)} \end{array}$$

The system $\mathsf{K}^{EC}_{\mathcal{PA}}$ consist of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}$ and the axioms **E**, **Mix**, and rule **Ind** for sentences of $\mathscr{L}^{EC}_{\mathcal{PA}}$. And in the same fashion we get $\mathsf{KD45}^{EC}_{\mathcal{PA}}$ and $\mathsf{S5}^{EC}_{\mathcal{PA}}$. Again these systems are sound and complete with respect to $*K_{\mathcal{PA}}$, $*KD45_{\mathcal{PA}}$, and $*S5_{\mathcal{PA}}$.

2.5 Bisimulation

A useful notion that stems from modal logic is bisimulation.

Definition 2.8 (Bisimulation)
Let two models $M = (W, R, V)$ and $M' = (W', R', V')$ in the class $K_{\mathcal{PA}}$ be given. A relation $\mathfrak{R} \subseteq W \times W'$ is a bisimulation iff for all $w \in W$ and $w' \in W'$ with $w\mathfrak{R}w'$:

atoms $w \in V(p)$ iff $w' \in V'(p))$ for all $p \in \mathcal{P}$

forth for all $a \in \mathcal{A}$ and all $v \in W$, if $wR(a)v$, then there is a $v' \in W'$ such that $w'R'(a)v'$ and $v\mathfrak{R}v'$

back for all $a \in \mathcal{A}$ and all $v' \in W'$, if $w'R'(a)v'$, then there is a $v \in W$ such that $wR(a)v$ and $v\mathfrak{R}v'$

We write $(M, w) \leftrightarroweq (M', w')$, iff there is a bisimulation between M and M' linking w and w'. Then we call (M, w) and (M', w') bisimilar. □

The main theorem about bisimulation is:

Theorem 2.2
Let two models (M, w) and (M', w') in $K_{\mathcal{PA}}$ be given. If $(M, w) \leftrightarroweq (M', w')$, then for every $\varphi \in \mathscr{L}^{EC}_{\mathcal{PA}}$ it holds that $(M, w) \models \varphi$ iff $(M', w') \models \varphi$. □

It generally does not hold vice versa. There are a number of cases when it does hold vice versa, for example if for every world the set of accessible worlds is finite. A generalization of bisimulation will play an important role in section 6.5.

Chapter 3

Strong completeness and disharmony

3.1 Introduction

Dynamic logic is a modal logic that was developed to reason about computer processes. The branch of logic was started by Pratt (1976). The propositional part of his logic (PDL) became an object of study in itself. Segerberg (1977) gave an axiomatization of it, that is mostly used today, but a completeness proof was not easily obtained. It took some time before several proofs were produced. The proof by Kozen and Parikh (1981) is considered to be one of the most elegant. The problem is that the canonical model method for proving completeness cannot be applied successfully. (A canonical model is a model such that every non-provable formula has a counterexample in the model, see (Blackburn, de Rijke, and Venema 2001).) The axiomatization by Segerberg is only weakly complete, because PDL is not compact (see below). The topic of this chapter is a *strongly* complete proof system for propositional dynamic logic, for which the canonical model method can be used to prove completeness.

Strong completeness (also called extended completeness) with respect to a class of frames S is the following property of a modal logical system S:

$$\Gamma \models_S \varphi \text{ implies } \Gamma \vdash_{\mathsf{S}} \varphi, \text{ for all formulas } \varphi \text{ and all sets of formulas } \Gamma.$$

This generalizes weak completeness, where Γ is empty. Observe that weak completeness implies strong completeness whenever the logic in question is semantically compact, i.e. when $\Gamma \models_S \varphi$ implies that there is a finite $\Gamma' \subseteq \Gamma$ with $\Gamma' \models_S \varphi$, hence $\models_S \bigwedge \Gamma' \to \varphi$. This is, for example, the case in modal logics such as K and $\mathsf{S5}$.

Propositional dynamic logic is a well known example of a non-compact logic: we have for the relevant class of frames S, that $\{[a^n]p \mid n \in \mathbb{N}\} \models_S [a^*]p$ but there is no natural number k with $\{[a^n]p \mid n \leq k\} \models_S [a^*]p$. As a consequence, we do not have strong completeness for any finitary axiomatization, *a fortiori* not for its usual, weakly complete proof system (see definition 3.3). So strong completeness requires an infinitary proof system.

The infinitary proof system presented in this paper is not the only strongly complete proof system for propositional dynamic logic. A detailed comparison to other work is provided in section 3.5. What is new about our system is that it is relatively simple, which makes the proof relatively simple. It can easily be extended to other modal logics. Moreover our study of the proof system has revealed a very peculiar feature of the canonical model, which we have dubbed *program disharmony*. In the standard completeness proof for modal logic, the truth lemma states that the set of formulas that is true in a world of the canonical mode is exactly the maximal consistent set of formulas, which is identified with that world. This we call *formula harmony* and its analogue for programs we call *program harmony*. The word harmony is used because when the model would have both formula and program harmony the semantics and proof theoretical aspects of the canonical model are in complete agreement. We show that the canonical model lacks program harmony while at the same time it does have formula harmony.

The rest of the chapter is structured as follows. Section 3.2 presents the infinitary proof system PDL_ω, as well as proofs of some derived rules, which are used in the central section 3.3 to prove that PDL_ω is strongly complete. Section 3.4 generalizes the method to prove strong completeness for enumerably axiomatized modal logics, in particular epistemic logic with a common knowledge operator. A comparison with related work is given in section 3.5. In section 3.6 it is shown that the canonical model for PDL_ω does not satisfy *program harmony*. Finally section 3.7 contains a conclusion and ideas for further research. This chapter is the result of joint work with Gerard Renardel and Rineke Verbrugge (see Renardel de Lavalette, Kooi, and Verbrugge (2002)).

3.2 The infinitary proof system PDL_ω

The infinitary proof system PDL_ω is an extension of the usual axiom system for PDL, with respect to the same language and the same Kripke semantics. As a reminder, we repeat the definitions of both language and semantics (for more information on PDL, see Harel, Kozen, and Tiuryn (2000)).

Definition 3.1 (Language of PDL)
Let a countable set of propositional variables $\mathcal{P}$ and a countable set of atomic programs Π be given. The language of PDL $\mathscr{L}_{\mathcal{P}\Pi}$ consists of a set of formulas φ and the set of programs α, given by the following rules in BNF:

$$\begin{array}{lcl}\varphi & ::= & \bot \mid p \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid [\alpha]\varphi \\ \alpha & ::= & a \mid \alpha_1;\alpha_2 \mid \alpha_1 \cup \alpha_2 \mid \alpha^* \mid ?\varphi\end{array}$$

where $p \in \mathcal{P}$, and $a \in \Pi$. □

Definition 3.2 (Propositional dynamic models)
A propositional dynamic model for $\mathscr{L}_{\mathcal{P},\Pi}$ is a tuple $M = (W, R, V)$ such that:

- $W \neq \emptyset$: a set of states or possible worlds;
- $R(a)$: a binary relation on W for each atomic program a;
- $V : \mathcal{P} \to 2^W$; assigns a set of states to each propositional variable.

If $M = (W, R, V)$ is a model, a pair (M, w), where $w \in W$ is is called a pointed model. A pair $F = (W, R)$ is called a frame and a pair (F, w), where $w \in W$, is called a pointed frame. A pair (F, V) is also a model. □

The truth definition is as expected for normal modal logics. As a reminder, here follows the modal clause:

$$(M, w) \models [\alpha]\varphi \text{ iff } (M, v) \models \varphi \text{ for all } v \text{ with } wR(\alpha)v$$

where R is extended in the following way

- $R(\alpha; \beta) = R(\alpha) \circ R(\beta)$;
- $R(\alpha \cup \beta) = R(\alpha) \cup R(\beta)$;
- $R(\alpha^*) = R(\alpha)^*$ = reflexive transitive closure of $R(\alpha)$;
- $R(?\varphi) = \{(w, w) \mid (M, w) \models \varphi\}$.

Let S be the class of propositional dynamic frames. We show that PDL_ω is complete with respect to this class of frames S:

$$\Gamma \models_S \varphi \text{ implies } \Gamma \vdash_{\mathsf{PDL}_\omega} \varphi, \text{ for all formulas } \varphi \text{ and all sets of formulas } \Gamma.$$

Below $\models_S$ and $\vdash_{\mathsf{PDL}_\omega}$ are abbreviated to $\models$ and $\vdash$ respectively. By $\Gamma \models \varphi$ we mean the local consequence relation, i.e. $\Gamma \models \varphi$ iff for every model M such that the corresponding frame is in S, $(M, w) \models \psi$ for every $\psi \in \Gamma$ implies that $(M, w) \models \varphi$.

Definition 3.3 (Axioms for PDL)
Here follows the usual set of axioms for PDLwithout the induction axiom.

Taut	all instantiations of propositional tautologies
Distr	$[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$
?**AX**	$[?\varphi]\psi \leftrightarrow (\varphi \to \psi)$
; **AX**	$[\alpha; \beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$
$\cup$**AX**	$[\alpha \cup \beta]\varphi \leftrightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$
$*$**AX**	$[\alpha^*]\varphi \leftrightarrow (\varphi \wedge [\alpha][\alpha^*]\varphi)$

where $\varphi, \alpha \in \mathscr{L}_{\mathcal{P}\Pi}$. □

In the following we extend the system PDL to an infinitary proof system PDL_ω by inductively defining a derivation relation $\Gamma \vdash \varphi$ (φ a formula, Γ a set of formulas). Notice that in the following definition, the language remains finitary (all formulas have finite length) and only the rule **Inf*** is non-standard.

Besides the usual shorthand notation Γ, φ for $\Gamma \cup \{\varphi\}$, Γ, Δ for $\Gamma \cup \Delta$, $\vdash \varphi$ for $\emptyset \vdash \varphi$, and $\varphi_1, \ldots, \varphi_n \vdash \psi$ for $\{\varphi_1, \ldots, \varphi_n\} \vdash \psi$, we shall also write:

$\Gamma \vdash \Delta$	for	$\Gamma \vdash \varphi$ for all $\varphi \in \Delta$
$[\alpha]\Gamma$	for	$\{[\alpha]\varphi \mid \varphi \in \Gamma\}$
$\varphi \to \Gamma$	for	$\{\varphi \to \psi \mid \psi \in \Gamma\}$

Definition 3.4 (Infinitary derivation relation for PDL_ω)
$\Gamma \vdash \varphi$ is defined as the smallest relation closed under the following rules:

AX	$\vdash \varphi$ if φ is an axiom of PDL	
MP	$\varphi, \varphi \to \psi \vdash \psi$	(modus ponens)
Inf*	$\{[\alpha; \beta^n]\varphi \mid n \in \mathbb{N}\} \vdash [\alpha; \beta^*]\varphi$	(infinitary $*$-introduction)
Nec	if $\vdash \varphi$ then $\vdash [\alpha]\varphi$	(necessitation)
W	if $\Gamma \vdash \varphi$ then $\Gamma, \Delta \vdash \varphi$	(weakening)
Cut	if $\Gamma \vdash \Delta$ and $\Gamma, \Delta \vdash \varphi$ then $\Gamma \vdash \varphi$	

where $\varphi, \psi, \alpha, \beta \in \mathscr{L}_{\mathcal{P}\Pi}$ and $\Gamma, \Delta \subseteq \mathscr{L}_{\mathcal{P}\Pi}$. □

The usual induction axiom for iteration $(\varphi \wedge [\alpha^*](\varphi \to [\alpha]\varphi) \to \varphi)$ is derivable in this system.

It is not hard to verify that these rules are sound with respect to the semantics of PDL (i.e. $\Gamma \vdash \varphi$ implies that $\Gamma \models \varphi$). We shall show in section 3.3 that the system PDL_ω is also strongly complete with respect to these semantics. For this, we shall use some derived rules that we introduce now. The most important of these are *deduction* (**Ded**) and *strong necessitation* (**SNec**), while the other rules are only used to prove **Ded** and **SNec**. **Ded** will be used in the Lindenbaum lemma and both **Ded** and **SNec** in the Truth lemma. We remark that the fact that **SNec** holds while only **Nec** is part of the proof system, is essential in our proof of strong completeness.

Lemma 3.1 (Derived rules of PDL_ω)
We can prove the following derived rules:

SCut	if $\Gamma \vdash \Delta$ and $\Gamma', \Delta \vdash \varphi$ then $\Gamma, \Gamma' \vdash \varphi$	(strong cut)
Det	if $\Gamma \vdash \varphi \to \psi$ then $\Gamma, \varphi \vdash \psi$	(detachment)
Cond	if $\Gamma \vdash \varphi$ then $(\psi \to \Gamma) \vdash \psi \to \varphi$	(conditionalization)
Ded	if $\Gamma, \varphi \vdash \psi$ then $\Gamma \vdash \varphi \to \psi$	(deduction)
SNec	if $\Gamma \vdash \varphi$ then $[\alpha]\Gamma \vdash [\alpha]\varphi$	(strong necessitation)

where $\varphi, \psi, \alpha, \beta \in \mathscr{L}_{\mathcal{P}\Pi}$ and $\Gamma, \Gamma', \Delta \subseteq \mathscr{L}_{\mathcal{P}\Pi}$. □

Proof Notice that the structure of the proof below is an infinitary induction over derivations; this is not a problem because of the well-foundedness of derivations.

SCut is easy to prove using **W** and **Cut**. **Det** is also easy, with **MP** and **SCut**.

Cond : Proof by induction over a derivation of $\Gamma \vdash \varphi$. The cases below are named by the rule applied last:

AX In this case φ is an axiom, $\Gamma = \emptyset$, and $\vdash \varphi$; also $\vdash \varphi \to (\psi \to \varphi)$ for it is a tautology. By **MP** we have $\varphi, \varphi \to (\psi \to \varphi) \vdash \psi \to \varphi$. Now apply **SCut** twice to obtain $\vdash \psi \to \varphi$.

MP Follows via tautologies and **Det**.

Inf* By ?**AX**and ;**AX**, we have $\vdash (\psi \to [\alpha;\beta^n]\varphi) \to [?\psi;\alpha;\beta^n]\varphi$ for all n, and $\vdash [?\psi;\alpha;\beta^*]\varphi \to (\psi \to [\alpha;\beta^*]\varphi)$; now the result follows via **SCut**, ;**AX**, and **Det**.

Nec If $\vdash \varphi$, then by **Nec** $\vdash [?\psi;\alpha]\varphi$, so with ?**AX**, ;**AX** and **MP** we get $\vdash \psi \to [\alpha]\varphi$.

W, Cut Easy application of the induction hypothesis.

Ded : if $\Gamma, \varphi \vdash \psi$, then (by **Cond**) $(\varphi \to \Gamma), (\varphi \to \varphi) \vdash (\varphi \to \psi)$. With **SCut** we remove the tautology $\varphi \to \varphi$ and obtain $(\varphi \to \Gamma) \vdash \varphi \to \psi$. Now the result follows via **SCut** from $\Gamma \vdash (\varphi \to \Gamma)$, i.e. $\Gamma \vdash \varphi \to \chi$ for all $\chi \in \Gamma$. But this follows via **W** from $\chi \vdash \varphi \to \chi$, which is a consequence of **Taut** and **Det**.

SNec : induction over a derivation of $\Gamma \vdash \varphi$.

AX Easy, with **Nec**.

MP Take **Distr** and apply **Det** twice.

Inf* We have $\{[\gamma;\alpha;\beta^n]\varphi \mid n \in \mathbb{N}\} \vdash [\gamma;\alpha;\beta^*]\varphi$; via ;**AX** and **SCut** we obtain

$$[\gamma]\{[\alpha;\beta^n]\varphi \mid n \in \mathbb{N}\} \vdash [\gamma][\alpha;\beta^*]\varphi$$

Nec Easy, with **Nec**.

W,Cut These cases follow directly by the induction hypothesis.

□

An immediate consequence of **Ded** is that if Γ, φ is PDL_ω-inconsistent, then $\Gamma \vdash \neg\varphi$.

3.3 Strong completeness: the canonical model for PDL_ω

In this section, we consider a fixed language $\mathscr{L}_{\mathcal{P}\Pi}$. We shall prove strong completeness of PDL_ω via the Henkin construction of a canonical model. There are two steps: first we show that every PDL_ω-consistent set can be extended to a maximal PDL_ω-consistent set, then we construct a Kripke model consisting of maximal PDL_ω-consistent sets. In section 3.6 we show that this model does not satisfy program harmony, by giving a countermodel.

We recall the obvious fact that a collection of formulas Γ is maximal PDL_ω-consistent iff it is PDL_ω-consistent (i.e. $\Gamma \nvdash \bot$) and Γ contains exactly one from $\varphi, \neg\varphi$ for every formula φ in the language $\mathscr{L}_{\mathcal{P},\Pi}$. In the remainder of this section, we will omit the prefix PDL_ω before "consistent".

Lemma 3.2 (Lindenbaum lemma for PDL_ω)
Every consistent set can be extended to a maximal consistent set. □

Proof Let Δ be a consistent set, i.e. $\Delta \nvdash \bot$. Let $\{\varphi_n \mid n \in \mathbb{N}\}$ be an enumeration of all formulas in $\mathscr{L}_{\mathcal{P},\Pi}$. We shall inductively define an increasing sequence $\{\Gamma_n \mid n \in \mathbb{N}\}$ of formula sets extending Δ, and show that $\Gamma = \bigcup\{\Gamma_n \mid n \in \mathbb{N}\}$ is maximal consistent.

$$
\begin{aligned}
\Gamma_0 &= \Delta \\
\Gamma_{n+1} &= \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if } \Gamma_n \vdash \varphi_n \\ \Gamma_n \cup \{\neg\varphi_n\} & \text{if } \Gamma_n \nvdash \varphi_n \text{ and } \varphi_n \text{ is not} \\ & \text{of the form } [\alpha;\beta^*]\psi \\ \Gamma_n \cup \{\neg\varphi_n, \neg[\alpha;\beta^k]\psi\} & \text{otherwise, where } k \text{ is the} \\ & \text{least natural number} \\ & \text{such that } \Gamma_n \nvdash [\alpha;\beta^k]\psi \\ & \text{(and } \varphi_n \text{ is of the form} \\ & [\alpha;\beta^*]\psi) \end{cases}
\end{aligned}
$$

We observe that the k in the last case always exists: for if $\Gamma_n \vdash [\alpha;\beta^k]\psi$ for all $k \in \mathbb{N}$, then (by **Inf*** and **Cut**) $\Gamma_n \vdash [\alpha;\beta^*]\psi$, contradicting $\Gamma_n \nvdash \varphi_n$. So the definition of Γ_n is correct. □

Now we claim the following for all formulas φ, ψ; from these claims, especially from (3) and (6), it follows immediately that Γ is maximal consistent:

1. every Γ_n is consistent;
2. $\vdash \varphi \Rightarrow \varphi \in \Gamma$;
3. $\varphi \notin \Gamma \Leftrightarrow \neg\varphi \in \Gamma$;

4. $(\varphi \to \psi) \in \Gamma \Leftrightarrow (\varphi \in \Gamma \Rightarrow \psi \in \Gamma)$;
5. $\Gamma \vdash \varphi \Rightarrow \varphi \in \Gamma$;
6. $\Gamma \nvdash \bot$.

The proofs of (2), (3), and (4) are as usual. We give the proofs in the three unusual cases:

1. Induction over n. For $n = 0$, consistency of Δ is given. Now assume that Γ_n is consistent. If, in the definition of Γ_{n+1}, the first or second case applies, it is clear that Γ_{n+1} is consistent. If the last case applies and Γ_{n+1} were inconsistent, then $\Gamma_n \vdash [\alpha; \beta^*]\psi \vee [\alpha; \beta^k]\psi$ via **Ded**, so, by using $*$**AX** k times, $\Gamma_n \vdash [\alpha; \beta^k]\psi$, contradicting the definition of k.

5. We shall prove a more general statement with induction over a derivation of $\Gamma' \vdash \varphi$: if $\Gamma' \subseteq \Gamma$ and $\Gamma' \vdash \varphi$, then $\varphi \in \Gamma$.

- φ is an axiom: by (2).
- **MP**: by (4).
- **Inf***: Let $\{[\alpha; \beta^n]\varphi \mid n \in \mathbb{N}\} \subseteq \Gamma$. To show that $[\alpha; \beta^*]\varphi \in \Gamma$, assume using contraposition that this is not the case: then by (3), $\neg[\alpha; \beta^*]\varphi \in \Gamma$. Let n be the index with $\varphi_n = [\alpha; \beta^*]\varphi$, then $\Gamma_n \nvdash \varphi_n$ (for otherwise, by the first case in the definition of Γ_{n+1}, $[\alpha; \beta^*]\varphi \in \Gamma_{n+1} \subseteq \Gamma$), so $\neg[\alpha; \beta^k]\varphi \in \Gamma_{n+1} \subseteq \Gamma$ for some k by the last case of the definition of Γ_{n+1}. But also, by assumption, $[\alpha; \beta^k]\varphi \in \Gamma$, so $\{\neg[\alpha; \beta^k]\varphi, [\alpha; \beta^k]\varphi\} \in \Gamma_m$ for some $m > n$, contradicting the consistency of Γ_m (1).
- **Nec**: by (2).
- **W**: direct consequence of the induction hypothesis.
- **Cut**: so $\Gamma' \vdash \Gamma''$ and $\Gamma' \cup \Gamma'' \vdash \varphi$ for some Γ''. By the induction hypothesis, we get $\Gamma'' \subseteq \Gamma$, so $\Gamma' \cup \Gamma'' \subseteq \Gamma$; by applying the induction hypothesis again we obtain $\varphi \in \Gamma$.

6. Suppose $\Gamma \vdash \bot$, then $\bot \in \Gamma$ (by **5**), but then $\bot \in \Gamma_n$ for some n, which contradixts **1**.

Now we can define the canonical model needed for strong completeness.

Definition 3.5 (Canonical model)
We define the canonical Kripke model
$M = (W, R, V)$ by

- $W = \{\Gamma \mid \Gamma \text{ maximal consistent}\}$
- $R(a) = \{(\Gamma, \Delta) \in W^2 \mid \varphi \in \Delta \text{ for all } \varphi \text{ such that } [a]\varphi \in \Gamma\}$

- $V(p) = \{\Gamma \in W \mid p \in \Gamma\}$ □

The Truth lemma shows that $(M, \Gamma) \models p \Leftrightarrow p \in \Gamma$ extends to all formulas of the language:

Lemma 3.3 (Truth lemma)
For all $\Gamma \in W$ and for all formulas $\varphi \in \mathscr{L}_{\mathcal{P}\Pi}$, we have $(M, \Gamma) \models \varphi$ iff $\varphi \in \Gamma$. □

Proof Induction over φ. The atomic and propositional cases are standard. We will prove the case $\varphi = [\alpha]\psi$, by induction over α; the cases for complex programs α of the forms $?\chi$, $\beta;\gamma$ and $\beta \cup \gamma$ are easy (with the corresponding axioms), so we only give the proofs of the remaining two unusual cases. Note that the proof as a whole has the form of an induction over a well-ordering of formulas, where $[\alpha^n]\varphi$ is considered to be a subformula of $[\alpha^*]\varphi$.

1. $\alpha = a$, atomic. Using the definition of the truth relation and the induction hypothesis $(M, \Delta) \models \psi \Leftrightarrow \psi \in \Delta$ for all $\Delta \in W$, we see that $(M, \Gamma) \models [a]\psi$ is equivalent to

$$\text{for all } \Delta \in W (\text{for all } \chi([a]\chi \in \Gamma \Rightarrow \chi \in \Delta) \Rightarrow \psi \in \Delta) \qquad \text{(A)}$$

   It is evident (A) follows from $[a]\psi \in \Gamma$. To see that (A) implies $[a]\psi \in \Gamma$ as well, we argue via contraposition. So assume $[a]\psi \notin \Gamma$, i.e. (by maximal consistency) $\neg[a]\psi \in \Gamma$. We shall show that there is a maximal consistent Δ with $\theta \in \Delta$ for all θ such that $[a]\theta \in \Gamma$, and $\neg\psi \in \Delta$. By the Lindenbaum lemma, it suffices to show that $\{\chi \mid [a]\chi \in \Gamma\} \cup \{\neg\psi\}$ is consistent. Assume it is not, i.e. $\{\chi \mid [a]\chi \in \Gamma\} \cup \{\neg\psi\} \vdash \bot$, then $\{\chi \mid [a]\chi \in \Gamma\} \vdash \psi$ via **Ded**. Thus, with **SNec**: $\{[a]\chi \mid [a]\chi \in \Gamma\} \vdash [a]\psi$. Hence *a fortiori* $\Gamma \vdash [a]\psi$ and $[a]\psi \in \Gamma$, contradicting the assumption $[a]\psi \notin \Gamma$. Therefore Δ is consistent, and for all $\chi([a]\chi \in \Gamma \Rightarrow \chi \in \Delta)$, however $\psi \notin \Delta$. Therefore (A) is not the case.

2. $\alpha = \beta^*$: $(M, \Gamma) \models [\beta^*]\psi \Leftrightarrow$ for all $n \in \mathbb{N}$ $(M, \Gamma) \models [\beta^n]\psi \Leftrightarrow ([\beta^n]\psi \in \Gamma$ for all $n \in \mathbb{N}) \Leftrightarrow [\beta^*]\psi \in \Gamma$, using the induction hypothesis in the second step, and $*$**AX**, **Inf*** in the last step.

□

Note that in the Truth lemma, we do not prove the dual property for programs, namely, $\Gamma R(\alpha)\Delta$ iff $\varphi \in \Delta$ for all φ such that $[\alpha]\varphi \in \Gamma$ (it holds by definition for atomic programs a). In section 3.6 we elaborate on this lack of "full harmony" (definition 3.8). By full harmony we mean that the semantics of the canonical model agree with its proof theoretical aspects.

Theorem 3.1 (Strong completeness of PDL_ω)
Let S be the class of all Kripke frames for the language $\mathscr{L}_{\mathcal{P},\Pi}$. Then for all formulas φ and all sets of formulas Φ, $\Phi \models_S \varphi$ implies $\Phi \vdash \varphi$. □

Proof By contraposition. Suppose $\Phi \nvdash \varphi$, then $\Phi \cup \{\neg\varphi\}$ is consistent. By lemma 3.2, $\Phi \cup \{\neg\varphi\}$ is extended to a maximal consistent set Γ with $\neg\varphi \in \Gamma$ and $\Phi \subseteq \Gamma$. Now by lemma 3.3, we conclude that in the canonical Kripke model, $(M, \Gamma) \not\models \varphi$ and $(M, \Gamma) \models \Phi$, as desired. □

Note that the completeness proof immediately gives a canonical standard model, contrary to the early proofs of weak completeness for PDL as they appear in Kozen and Parikh (1981, Harel, Kozen, and Tiuryn (2000), which use nonstandard models.

3.4 Generalization to other modal logics

The approach of sections 3.2 and 3.3 may be generalized to other denumerably axiomatized modal logics. We briefly describe the general method. Take a denumerable modal language and define the derivability relation $\vdash$ generated by **Taut**, **Distr**, **MP**, **Nec**, **W**, **Cut** as well as the following denumerable set of rules:

$$\textbf{Rules} = \{(\Gamma_i, \varphi_i) \mid i \in \mathbb{N}\}$$

Here we assume without loss of generality that **Rules** is closed under **Cond** and **SNec**. For, one can close off a denumerable set of rules under **Cond** and **SNec** in denumerably many steps, so that the resulting set of rules is still denumerable. (This is similar to Segerberg (1994, Goldblatt (1993), even though we consider the *smallest* set of rules closed under **Cond** and **SNec** and not just any).

Then one can prove a lemma about derived rules analogous to lemma 3.1. For the analogue of the Lindenbaum lemma 3.2, the following adaptation is needed.

Lemma 3.4 (General Lindenbaum lemma)
Every consistent set can be extended to a maximal consistent set. □

Proof The idea of this proof is that, just as in the proof of lemma 3.2 we build a maximal consistent set step by step. We want to ensure that the set is consistent at the end of its construction. We do not want to accidentally add all the elements of a Γ_i to the set without adding φ_i to it (if $(\Gamma_i, \varphi_i) \in \textbf{Rules}$). Therefore when we add the negation of a formula that occurs as the conclusion of a rule, we also add a witness to make sure that the rule cannot be applied. Since there can be infinitely many rules which have the same formula as its conclusion we have to take an enumeration of all formulas where every formula occurs infinitely often. Every time we encounter a formula we deal with another rule.

Let Δ be a consistent set, i.e. $\Delta \nvdash \bot$. Let $\{\psi_n \mid n \in \mathbb{N}\}$ be an enumeration of all formulas in the language, such that every formula appears infinitely many times. This yields a set of numbers $J_\varphi = \{n \mid \psi_n = \varphi\}$ for every formula φ. Moreover we have a set of numbers $I_\varphi = \{i \mid (\Gamma_i, \varphi) \in \textbf{Rules}\}$ for every formula

φ. With these two sets we can define a function $f_\varphi : I_\varphi \to J_\varphi$ such that $f_\varphi(i_k) = j_k$ (the k-th number in I_φ is mapped to the k-th number in J_φ). We shall inductively define an increasing sequence $\{\Delta_n \mid n \in \mathbb{N}\}$ of formula sets extending Δ; then $\Gamma = \bigcup\{\Delta_n \mid n \in \mathbb{N}\}$ is maximal consistent.

$$\begin{array}{lcl} \Delta_0 & = & \Delta \\ \Delta_{n+1} & = & \begin{cases} \Delta_n \cup \{\psi_n\} & \text{if } \Delta_n \vdash \psi_n \\ \Delta_n \cup \{\neg\psi_n, \neg\gamma\} & \text{if } \Delta_n \nvdash \psi_n \text{ and there is an } i \text{ with } f_{\psi_n}(i) = n \\ & \text{(i.e. there exists a } (\Gamma_i, \psi_n) \in \mathbf{Rules}\text{)} \\ & \text{and } \gamma \text{ is the smallest formula in } \Gamma_i \\ & \text{such that } \Delta_n \cup \{\neg\psi_n, \neg\gamma\} \text{ is consistent.} \\ \Delta_n \cup \{\neg\psi_n\} & \text{otherwise} \end{cases} \end{array}$$

In the second case such a γ always exists. If there were no such γ then $\Delta_n \cup \{\neg\psi_n\} \vdash \Gamma_i$ and therefore $\Delta_n \cup \{\neg\psi_n\} \vdash \psi_n$ by **SCut** on all formulas in Γ_i, which would mean that $\Delta_n \vdash \psi_n$ contradicting the assumption that $\Delta_n \nvdash \psi_n$.

Now we can prove Γ is consistent analogously to the Lindenbaum lemma for PDL. □

Then, to complete the completeness proof, we define a canonical model as follows, analogously to definition 3.5:

Definition 3.6 (Canonical model)
We define the canonical Kripke model
$M = (W, \{R(\alpha) : \alpha \text{ is a modality in the language }\}, V)$ by

- $W = \{\Gamma \mid \Gamma \text{ maximal consistent}\}$
- $R(\alpha) = \{(\Gamma, \Delta) \in W^2 \mid \varphi \in \Delta \text{ for all } \varphi \text{ such that } [a]\varphi \in \Gamma\}$
- $V(p) = \{\Gamma \in W \mid p \in \Gamma\}$ □

Proving the Truth lemma for this canonical model is straightforward, cf. lemma 3.3. Now soundness and strong completeness of $\vdash$ with respect to **Rules**-models follows easily.

Note that the strong completeness result for PDL_ω does not follow straightforwardly from the above proof, and in that sense it is not a true generalization. This is because PDL obeys a structure among its modalities, so that some more work is required. The same is true for epistemic logic with common knowledge, to which we will turn next.

3.4.1 Epistemic logic

In this section, we extend the strong completeness result for PDL_ω to epistemic logic with common knowledge. For the relevant definitions of the language and semantics, see chapter 2. Here we limit general knowledge and common knowledge operators to the whole group of agents $\mathcal{A}$, and therefore we abbreviate $E_{\mathcal{A}}\varphi$ and $C_{\mathcal{A}}\varphi$ as $E\varphi$ and $C\varphi$ respectively.

The similarity between PDL and epistemic logic with common knowledge has long been noted. In fact, strong completeness of the latter immediately reduces to the former by the following embedding $'$. Suppose the set of agents $\mathcal{A} = \{a_1, \ldots, a_n\}$, and define:

$$\begin{array}{lcl} (\Box_{a_i}\varphi)' & = & [a_i]\varphi' \\ (E\varphi)' & = & [a_1 \cup \ldots \cup a_n]\varphi' \\ (C\varphi)' & = & [(a_1 \cup \ldots \cup a_n)^*]\varphi' \end{array}$$

However, we can prove something stronger if we take a direct approach. It turns out that the Henkin method can, in addition to K, also be used for systems like T, S4 and S5 with common knowledge.

We extend all four axiom systems to infinitary proof systems $\mathsf{KEC}_{\mathcal{PA}\omega}$, $\mathsf{TEC}_{\mathcal{PA}\omega}$, $\mathsf{S4EC}_{\mathcal{PA}\omega}$, and $\mathsf{S5EC}_{\mathcal{PA}\omega}$ by retaining all the axioms and adding a fixed set of derivation rules. We first introduce some notation in order to describe the infinitary introduction rule for C. We want this rule to contain all instances of the form

$$\{\varphi_1 \to \Box_a(\varphi_2 \to \Box_b(\ldots \to E^n\psi)) \mid n \in \mathbb{N}, n \geq 1\} \vdash \\ \varphi_1 \to \Box_a(\varphi_2 \to \Box_b(\ldots \to C\psi)),$$

where $E^n\psi$ is the obvious abbreviation defined inductively by $E^0\psi = \psi$ and $E^{n+1}\psi = EE^n\psi$. The neat way to formulate the infinitary rule is to introduce finite sequences $\pi = (\pi_1, \ldots, \pi_n)$ where the π_i are either formulas or modalities $\Box_a$ for $a \in \mathcal{A}$, with

$$\begin{array}{lcl} ()\varphi & = & \varphi \\ (\psi;\pi)\varphi & = & \psi \to (\pi)\varphi \\ (\Box_a;\pi)\varphi & = & \Box_a((\pi)\varphi) \end{array}$$

The infinitary rule may then be formulated as $\{(\pi)E^n\varphi \mid n \in \mathbb{N}, n \geq 1\} \vdash (\pi)C\varphi$. This formulation is needed for obtaining strong necessitation and strong conditionalization. We give the derivation rules for the infinitary systems.

Definition 3.7 (Infinitary derivation relations)
Let S be any of $\mathsf{KEC}_{\mathcal{PA}\omega}$, $\mathsf{TEC}_{\mathcal{PA}\omega}$, $\mathsf{S4EC}_{\mathcal{PA}\omega}$, and $\mathsf{S5EC}_{\mathcal{PA}\omega}$. $\Gamma \vdash_{\mathsf{S}} \varphi$ is defined as

the smallest relation closed under the following rules:

AX	$\vdash_{\mathsf{S}} \varphi$ if φ is an axiom of S	
MP	$\varphi, \varphi \to \psi \vdash_{\mathsf{S}} \psi$	(modus ponens)
InfC	$\{(\pi)E^n\varphi \mid n \in \mathbb{N}, n \geq 1\} \vdash_{\mathsf{S}} (\pi)C\varphi$	(infinitary C-introduction)
Nec	if $\vdash_{\mathsf{S}} \varphi$ then $\vdash_{\mathsf{S}} \Box_a\varphi$	(necessitation)
W	if $\Gamma \vdash_{\mathsf{S}} \varphi$ then $\Gamma, \Delta \vdash_{\mathsf{S}} \varphi$	(weakening)
Cut	if $\Gamma \vdash_{\mathsf{S}} \Delta$ and $\Gamma, \Delta \vdash_{\mathsf{S}} \varphi$ then $\Gamma \vdash_{\mathsf{S}} \varphi$	

where $\varphi, \psi \in \mathscr{L}_{\mathcal{PA}}$ and $\Gamma, \Delta \subseteq \mathscr{L}_{\mathcal{PA}}$ and π is a sequence of formulas or modalities. □

The usual rule for common knowledge **Ind** (see page 2.4) is derivable in this system.

Now the reader may check that for these systems, the derived rules **SCut**, **Det**, **Cond**, and **Ded** of lemma 3.1 can be proved, as well as the following analogue of strong necessitation **SNec**:

SNecK if $\Gamma \vdash_{\mathsf{S}} \varphi$ then $\Box_a\Gamma \vdash_{\mathsf{S}} \Box_a\varphi$ (strong necessitation for knowledge)

It is immediate that all four systems are sound with respect to the appropriate semantics: $\mathsf{KEC}_{\mathcal{PA}\omega}$ for all Kripke frames, $\mathsf{TEC}_{\mathcal{PA}\omega}$ for reflexive ones, $\mathsf{S4EC}_{\mathcal{PA}\omega}$ for reflexive transitive ones, and $\mathsf{S5EC}_{\mathcal{PA}\omega}$ for equivalence relations.

Theorem 3.2
Let S be any of the systems $\mathsf{KEC}_{\mathcal{PA}\omega}$, $\mathsf{TEC}_{\mathcal{PA}\omega}$, $\mathsf{S4EC}_{\mathcal{PA}\omega}$, and $\mathsf{S5EC}_{\mathcal{PA}\omega}$. Then S is strongly complete with respect to the appropriate set of frames. □

Proof sketch By a Henkin construction of a canonical model, analogously as in section 3.3. The presence of the appropriate axioms from **T**, **4**, and **5** in the maximal consistent sets induces the appropriate properties of the accessibility relations in the canonical model. In the analogue of the Lindenbaum lemma, the last clause for Γ_{n+1} should be "$\Gamma_{n+1} = \Gamma_n \cup \{\neg\varphi_n, \neg(\pi)E^k\psi\}$ otherwise, where k is the least natural number ≥ 1 such that $\Gamma_n \not\vdash (\pi)E^k\psi$ (and φ_n is of the form $(\pi)C\psi$)". The definition of the canonical model is as usual for epistemic logics. The main difference from the proof of the Truth lemma 3.3 is the induction step for operator C, which works very smoothly: $(M, \Gamma) \models C\psi \Leftrightarrow$ for all $n \in \mathbb{N}, n \geq 1$ $(M, \Gamma) \models E^n\psi \Leftrightarrow (E^n\psi \in \Gamma$ for all $n \in \mathbb{N}, n \geq 1) \Leftrightarrow C_\psi \in \Gamma$, using the induction hypothesis in the second step, and **Mix** (see section 2.4), **InfC** in the last step.

It is clear from the proof sketch that all four epistemic logics with common knowledge are canonical: on their canonical frames, all their axioms are valid, see Blackburn, de Rijke, and Venema (2001).

3.5 Comparison to other work

As was noted in the introduction, the main class of logics for which strong completeness is hard to obtain is the class of non-compact logics. When non-compactness is caused by a modality that is interpreted as a (reflexive) transitive closure of another modality it seems rather natural to consider infinitary proof systems due to the infinitary character of these modal operators. The proof system presented in this chapter is by no means the first attempt to devise a proof system that is strongly complete for such a non-compact logic. In this section we provide a historic overview of the results that were established in the past.

Sundholm (1977) proves strong completeness for Von Wright's temporal logic. This is a logic with two modalities, the *next-time* (or *tomorrow*) operator, and the *always* operator, where the latter is the reflexive transitive closure of the former. The intended model for this logic consists of the natural numbers as the set of possible worlds with the relations 'successor' and 'less than or equal' for the two modalities. Sundholm proved this theorem when Segerberg was studying Von Wright's tense logic and he found that the Lindenbaum Lemma could not be applied successfully, because of the infinitary rules. Von Wright's logic is a logic of linear time, and consequently the next time operator is deterministic, which makes some things easier compared to PDL.

A strongly complete infinitary proof system for *Propositional Algorithmic Logic* (PAL) is presented in Mirkowska (1981). PAL is very similar to PDL, the construction of formulas is identical, but different program constructions are used. Apart from atomic programs, sequential composition and non-deterministic choice, there are "if ... then ... else ... " statements and "while ... do ... " loops. The infinitary character of iteration is mimicked by the while loops. Note that PAL does not contain tests. The semantics of PAL are based on Kripke models. Although in the general setting there are no constraints on the semantics, strong completeness is only proved in the case where atomic programs are deterministic or when the indeterminism is bounded (in the sense that for every atomic program there is a bound on the number of different executions). This restriction is made because otherwise the infinitary proof rule is not sound. Therefore the results are not directly applicable to PDL.

Goldblatt (1982) proves strong completeness for a (nameless) logic that is also very similar to PDL, but there are important differences. The language is different from PDL in two ways. There is a distinction made between *expressions*, *commands*, and *formulas*. The commands and formulas are just like the programs and formulas of PDL. The commands differ from programs of PDL in program constructions: commands are atomic programs, sequential composition, "if ... then ... else" statements, or "while ... do ... " loops. Note that there are no nondeterministic choices, or tests. The expressions form a distinct set of formulas. The main feature of these is that the interpretation can be undefined, where undefinedness is inherited by sequentially evaluating an expression from left to right. So if the truth value of φ is undefined and the truth value of ψ is 0, then the truth value of $\varphi \wedge \psi$ is undefined, but the truth value of $\psi \wedge \varphi$ is 0. There are separate axioms and proof rules for these expressions. The general outline of the completeness proof for this logic is similar to the proof given in this chapter. The different features make the proof rather cumbersome. An extension of this logic with nondeterministic choice is also considered.

In Knijnenburg (1988), Knijnenburg gives an infinitary axiomatization of PDLİt contains the ∞-rule *from* $\{\psi \rightarrow [\alpha^i]\varphi \mid i < \omega\}$ *infer* $\psi \rightarrow [\alpha^*]\varphi$. Weak completeness for this axiomatization is claimed, but we suspect that the proof is incorrect. In Knijnenburg and van Leeuwen (1991), another proof of strong completeness is presented with respect to a specific model. Although we think that the result holds, we have trouble with some of the details of the paper. For example in the proof of theorem 5.10 of Knijnenburg and van Leeuwen (1991) in order to extend a set of formulas to a maximally consistent one, a reference is made to Lindenbaum's theorem. It seems to us that this technique cannot be applied directly in the case of infinitary proof systems. This is the same difficulty that prompted Sundholm to start his investigations.

In Goldblatt (1992), Goldblatt introduces the Omega-Iteration proof rule *from* $\{\varphi \rightarrow [\beta; \alpha^n]\psi \mid n \in \omega\}$ *deduce* $\varphi \rightarrow [\beta; \alpha^*]\psi$ in the context of first-order dynamic logic, and proves weak completeness. In Goldblatt (1993, chapter 9) a general approach for infinitary proof systems for normal modal logics containing modalities with arbitrary arity is given. Goldblatt shows that the addition of rules that satisfy certain properties to a basic proof system, yields a proof system that is strongly complete with respect to the appropriate class of models. In that framework it is not hard to make a strongly complete proof system for PDL, although Goldblatt does not discuss any applications. Goldblatt starts out by taking a fairly strong basic proof system. Strong necessitation (**SNec**), and the deduction theorem (**Ded**) are rules in his basic system. This makes his completeness proof more difficult than the one presented in this chapter.

Segerberg (1994) proves strong completeness for a whole class of modal logics that are not compact. This includes logics with (reflexive) transitive closure modalities such as Von Wright's temporal logic (see above), Goldblatt's ancestral logic (a logic with two modalities $\Box$ and $\Box^*$, where the latter is the reflexive transitive closure of the former), PDL without tests, and epistemic logic with common knowledge. Interestingly it also includes logics that are not compact due to other reasons. Segerberg discusses logics that satisfy the *bounded chain condition*, which states that any path in the model has a length bounded by some natural number n, in other words there are no paths with a length longer than n. Consequently the logic for the class of all frames that satisfy a bounded chain condition is not compact. The set of formulas $\{\Diamond^n \top \mid n \in \mathbb{N}\}$ is not satisfiable, however every finite subset is. These logics are axiomatized by adding infinitary rules to a finitary proof system for modal logic. He considers a very general case, where the only requirement is that the set of all instances of the infinitary rules are countable, and all the examples mentioned above fall into this category. The role played by maximal consistent sets in many completeness proofs, is played by *saturated sets* in his approach. The key feature is that a saturated set contains a witness formula for every infinitary rule of which the conclusion is not in the set. These sets form the possible worlds in the canonical models. In many cases it turns out that maximal consistency and saturation coincide.

In Tanaka (2001) Segerberg's results are extended to first-order modal logic. This is done by algebraic methods where representation theorems for modal algebras correspond to completeness for modal logics.

Our contribution to this ongoing research is that the base system of axioms and rules is quite simple. The strong versions of the rules are derived rules in our system, whereas they are part of the system in the systems discussed in this section. This makes some proofs easier. Another contribution is that we show a counterexample to program harmony, to which we turn now.

3.6 Program (dis)harmony

In this section we take a closer look at the canonical model defined in section 3.3. The construction of a canonical model is one of the most used techniques in completeness proofs for modal logics. Maximal consistent sets provide the bridge between syntax and semantics that facilitates the completeness theorem. Therefore one would expect the following property for the canonical model.

Definition 3.8 (Full harmony)
A canonical model M is fully harmonious iff for all maximal consistent sets Γ and Δ, and for all φ and all α:

formula harmony: $(M, \Gamma) \models \varphi$ iff $\varphi \in \Gamma$ and
program harmony: $\Gamma R(\alpha) \Delta$ iff $\{\varphi \mid [\alpha]\varphi \in \Gamma\} \subseteq \Delta$

Formula harmony generalizes a property of atomic formulas that holds in the canonical model by definition. Program harmony generalizes a property of atomic programs that holds in the canonical model by definition. We proved that the canonical model has formula harmony in the Truth Lemma, which is used in the Completeness Theorem. We only needed program harmony for atomic programs: the semantic properties of the accessibility relations for more complex programs are sufficient to prove the Truth Lemma.

Only when one focuses on program harmony as an interesting property in itself does one notice that it does not hold for the canonical model. In Kozen and Parikh (1981) program disharmony (i.e. failure of program harmony) was claimed without proof for finite canonical models of PDL. We found it quite surprising that the infinite canonical model for PDL_ω also lacks program harmony. It does

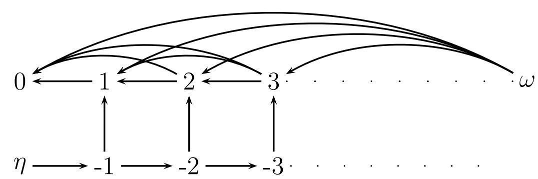


Figure 3.1: the countermodel M.

hold from left to right however. For this there is an easy proof that uses the Truth Lemma. Although it is preserved from right to left under tests, sequential composition and non-deterministic choice, it is not preserved under iteration. The remainder of this section is devoted to showing this.

We show that the canonical model is disharmonious by giving a countermodel. In principle any model which has no distinct bisimilar worlds can be seen as a part of the canonical model by taking the maximal consistent sets that are associated with the worlds in the model. Showing that program harmony fails for any two of those sets, implies that the whole canonical model is disharmonious.

The rest of this section is organised as follows. In order to prove the theorem that the canonical model is disharmonious we provide a countermodel to program harmony in definition 3.9. The proof of the theorem depends on two lemmas. The first lemma (Lemma 3.5) shows that in the countermodel there is a world, ω, which cannot be reached by the reflexive transitive closure of $R(a)$ from another world, η. The second lemma (Lemma 3.8) shows among other things that a formula is true in ω iff its extension is cofinite. This can be used to show that $\eta \models [\alpha^*]\varphi$ implies that $\omega \models \varphi$, because if $\eta \models [\alpha^*]\varphi$, then the extension of φ is cofinite. In order to prove Lemma 3.8 we need to study the structure of the countermodel in great detail. The relevant properties of the countermodel are summed up in Lemma 3.7. This is a rather involved technical Lemma which can be skipped by those readers who are not interested in the details of the proof.

Definition 3.9 (Countermodel)
Let $M = \langle \mathbb{Z} \cup \{\omega, \eta\}, R(a), V \rangle$ be a model for language $\mathscr{L}_{\emptyset,a}$, where

$$
\begin{aligned}
R(a) \quad = \quad & \{(x, y) \in \mathbb{N}^2 \mid x > y\} \cup \\
& \{(-x, x), (-x, -x-1) \mid x \in [1, \omega)\} \cup \\
& \{(\omega, x) \mid x \in \mathbb{N}\} \cup \{(\eta, -1)\}
\end{aligned}
$$

and $V = \emptyset$. □

See Figure 3.1 for a picture of this model. The idea is that program harmony fails at the worlds η and ω for the program a^*.

From the picture of the model it is quite clear that ω is not reachable from η in a finite number of steps. But we have to prove it for the maximal consistent sets associated with these worlds. Let $\Gamma = \{\varphi \mid (M,\eta) \models \varphi\}$ and $\Delta = \{\varphi \mid (M,\omega) \models \varphi\}$.

Lemma 3.5 (Unreachability of ω)
It is not the case that $\Gamma R(a^*)\Delta$. □

Proof If it would be the case that $\Gamma R(a^*)\Delta$, then there would be an n such that $\Gamma R(a^n)\Delta$. As was noted above program harmony does hold from left to right, so that would imply that $\{\varphi \mid (M,\eta) \models [a^n]\varphi\} \subseteq \Delta$. We prove that this is not the case by showing that for any n there is a φ_n such that

$$(\eta \models [a^n]\varphi_n \text{ and } \omega \models \neg\varphi_n)$$

To establish the statement take $\varphi_n = (\langle a\rangle[a]\bot \to [a^n]\bot)$, i.e. we have, for all n

$$\eta \models [a^n](\langle a\rangle[a]\bot \to [a^n]\bot)$$
$$\omega \models \neg(\langle a\rangle[a]\bot \to [a^n]\bot)$$

to see this, use $x \models \langle a\rangle[a]\bot \Leftrightarrow x \in \mathbb{N} \cup \{\omega\}$ and $\forall x \in \mathbb{N}(x \models [a^n]\bot \Leftrightarrow x < n)$. Furthermore none of the negative integers satisfy $\langle a\rangle[a]\bot$. □

Now we move to the other part of showing that the model is disharmonious. We have to show that if a formula of the form $[a^*]\varphi$ holds in η then φ holds in ω. It is rather difficult to show this in a direct way. Instead we characterize the set of all formulas that hold in ω. The idea of ω being the limit of the natural numbers gives the impression that if a formula holds from a natural number up, then it holds in ω, and vice versa. But why would the set of formulas that hold from a natural number up be a maximal consistent set? This is because any formula either holds from a certain natural number up or it does not hold from a certain natural number up, i.e. the interpretation of any formula is either finite or cofinite (a set is cofinite if its complement is finite.) In order to show this for formulas we need a similar property for programs.

Definition 3.10 (ADMS and ADMR)
The *admissible sets* $\mathsf{ADMS} \subseteq 2^{\mathbb{N}}$ are defined as follows.

$$\mathsf{ADMS} = \bigcup\{\mathsf{ADMS}(n) \mid n \in \mathbb{N}\}$$

where

$$\mathsf{ADMS}(n) = \{X \cup Y \mid X \subseteq [0,n), Y \in \{\emptyset, [n,\omega)\}\}$$

The *admissible relations* $\mathsf{ADMR} \subseteq 2^{\mathbb{N}\times\mathbb{N}}$ are defined as follows.

$$\mathsf{ADMR} = \bigcup\{\mathsf{ADMR}(n,k,p) \mid n \in \mathbb{N}, k \in [1,\omega], p \in [n,\omega)\}$$

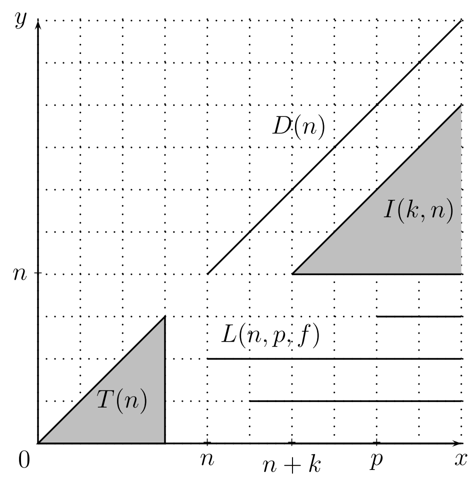


Figure 3.2: A typical element of ADMR

where $\mathsf{ADMR}(n,k,p)$ is the collection

$$\begin{array}{l}\{T \cup D \cup I(k,n) \cup L(n,p,f) \mid \\ \quad T \subseteq T(n), \\ \quad D \in \{\emptyset, D(n)\}, \\ \quad f : [0,n) \to ([n,p] \cup \{\omega\})\}\end{array}$$

and

$$\begin{array}{lcl} T(n) & = & \{(x,y) \mid n > x \geq y\} \\ D(n) & = & \{(x,x) \mid x \geq n\} \\ I(k,n) & = & \{(x,y) \mid x \geq y+k, y \geq n\} \\ L(n,p,f) & = & \{(x,y) \mid x \geq f(y), n > y\} \end{array}$$

where $x + \omega = \omega$. □

The admissible sets ADMS are simply the finite and cofinite subsets of the natural numbers. The formulation given in the definition was chosen because it suits our purposes better. See Figure 3.2 for a picture of a typical $\mathsf{ADMR}(n,k,p)$. We will show that every formula and every program is admissable. But to increase readibility we now prove the following auxiliary lemma concerning ADMR.

Lemma 3.6
ADMR is closed under composition, union, and reflexive transitive closure. □

Proof

composition To show that $R \circ R' \in \mathsf{ADMR}$ given that $R, R' \in \mathsf{ADMR}$, we argue as follows. First note that

$$\begin{array}{rcl} \mathsf{ADMR}(n,k,p) & \subseteq & \mathsf{ADMR}(n+1,k,\max(p,n+k)) \\ \mathsf{ADMR}(n,k,p) & \subseteq & \mathsf{ADMR}(n,k,p+1) \end{array}$$

So given that $R, R' \in \mathsf{ADMR}$, we may assume that there are n, p and k, k' such that $R = T \cup D \cup I(k,n) \cup L(n,p,f) \in \mathsf{ADMR}(n,k,p)$ and $R' = T' \cup D' \cup I(k',n) \cup L(n,p,f') \in \mathsf{ADMR}(n,k',p)$ with $T, T' \subseteq T(n)$ and $D, D' \in \{\emptyset, D(n)\}$. The results of the composition of the components of these R and R' are given in the following table (where $f'_k = \lambda y. f'(y) + k$, and $f_{T'} = \lambda y. \min\{f(z) \mid zT'y\}$ and $\min(\emptyset) = \omega$):

$\circ$	T'	$D(n)$	$I(k',n)$	$L(n,p,f')$
T	$T \circ T' \subseteq T(n)$	$\emptyset$	$\emptyset$	$\emptyset$
$D(n)$	$\emptyset$	$D(n)$	$I(k',n)$	$L(n,p,f')$
$I(k,n)$	$\emptyset$	$I(k,n)$	$I(k+k',n)$	$L(n,p+k,f'_k)$
$L(n,p,f)$	$L(n,p,f_{T'})$	$\emptyset$	$\emptyset$	$\emptyset$

So we have

$$R \circ R' \in \mathsf{ADMR}(n,k'',p+k) \text{ for some } k'' \in \{k, k', k+k'\}$$

and we conclude that $R \circ R' \in \mathsf{ADMR}$.

union To see that that $R \cup R' \in \mathsf{ADMR}$ we can again assume that there are n, p and k, k' such that $R \in \mathsf{ADMR}(n,k,p)$ and $R' \in \mathsf{ADMR}(n,k',p)$. So $R = T \cup D \cup I(k,n) \cup L(n,p,f), R' = T' \cup D' \cup I(k',n) \cup L(n,p,f')$ for certain $T, T' \subset T(n), D, D' \in \{\emptyset, D(n)\}, f, f' : [0,n) \to ([n,p] \cup \{\omega\})$ Now, since

$$\begin{array}{rcl} I(k,n) \cup I(k',n) & = & I(\min(k,k'),n) \\ L(n,p,f) \cup L(n,p,f') & = & L(n,p,\lambda y. \min(f(y), f'(y))) \end{array}$$

we see that $R \cup R' \in \mathsf{ADMR}(n, \min(k,k'), p) \subseteq \mathsf{ADMR}$, and we conclude that ADMR is closed under union.

reflexive transitive closure To see that $R^* \in \mathsf{ADMR}$ given that there are n, k, p such that $R \in \mathsf{ADMR}(n,k,p)$, we observe that $R^* = (R \cup D(n))^*$, because of reflexivity. So we may assume that $D = D(n)$. Because of the properties of composition given in the table above we have that $I(k,n) \subseteq R \circ R$. Observe that $I(2k,n) \subseteq I(k,n)$. Because $L(n,p,f') \cup L(n,p+k,f'_k) = L(n,p,f'_k)$ we have $R \cup (R \circ R) \in \mathsf{ADMR}(n,k,p)$. We now see that $R^m \in \mathsf{ADMR}(n,k,p)$ for all m. Since $\mathsf{ADMR}(n,k,p)$ is finite and closed under finite unions, it is closed under arbitrary unions and we have $R^* = \bigcup\{R^m \mid m \in \mathbb{N}\} \in \mathsf{ADMR}(n,k,p) \subseteq \mathsf{ADMR}$.

□

Now we prove the following.

Lemma 3.7
For every formula $\varphi \in \mathscr{L}_{\emptyset,a}$ and every program $\alpha \in \mathscr{L}_{\emptyset,a}$

$$[\![\varphi]\!] \in \mathsf{ADMS}$$
$$[\![\alpha]\!] \in \mathsf{ADMR}$$

where $[\![\varphi]\!] = \{x \in \mathbb{N} \mid x \models \varphi\}$ and $[\![\alpha]\!] = R(\alpha) \cap \mathbb{N} \times \mathbb{N}$. □

Proof We prove this by simultaneous induction on the structure of φ and α.

programs For the base case, the program a, observe that $[\![a]\!] = I(1,0)$. Moreover $T(0) = \emptyset$, therefore, we can take $D = \emptyset$, and $L(0,p,f) = \emptyset$ for every p and f. Therefore $[\![a]\!] \in \mathsf{ADMR}(0,1,2)$.

For programs the case for tests $?\varphi$ is also simple. The induction hypothesis implies that $[\![\varphi]\!] \in \mathsf{ADMS}$, therefore there is some n such that $[\![\varphi]\!] \in \mathsf{ADMS}(n)$. Note that $[\![?\varphi]\!] = \{(x,x) \mid x \in [\![\varphi]\!]\}$. Now we show that $[\![?\varphi]\!] \in \mathsf{ADMR}(n,\omega,n)$. For D we take $\{(x,x) \mid n \leq x \text{ and } x \in [\![\varphi]\!]\}$, and for T we take $\{(x,x) \mid n > x \text{ and } x \in [\![\varphi]\!]\}$. The set $I(\omega,n)$ is empty and we can take $L(n,p,f)$ to be empty by letting $f(x)$ be ω for all $x \in [0,n)$. Therefore $[\![?\varphi]\!] \in \mathsf{ADMR}$.

The cases for sequential composition, nondeterministic choice and iteration follow from the fact that ADMR is closed under composition, union, and reflexive transitive closure. See lemma 3.6.

formulas For the base case we only have to consider the formula $\bot$ and the program a. Now $[\![\bot]\!] = \emptyset$ and $\emptyset \in \mathsf{ADMS}(0)$.

The induction step for formulas is easy for negations and conjunctions. It is obvious that each $\mathsf{ADMS}(n)$ is closed under complementation and intersection. Consequently the whole set ADMS is (so it is also closed under union).

In order to finish the proof of Lemma 3.7 we have to show that $[\![\langle\alpha\rangle\varphi]\!] \in \mathsf{ADMS}$ given that $[\![\alpha]\!] \in \mathsf{ADMR}$ and $[\![\varphi]\!] \in \mathsf{ADMS}$. So we may assume by the induction hypothesis that for some n,k,p, $[\![\varphi]\!] \in \mathsf{ADMS}(n)$ and $[\![\alpha]\!] \in \mathsf{ADMR}(n,k,p)$, so $[\![\alpha]\!] = T \cup D \cup I(k,n) \cup L(n,p,f)$ with $T \subseteq T(n)$ and $D \in \{\emptyset, D(n)\}$. Now $[\![\langle\alpha\rangle\varphi]\!]$ is a subset of the domain of $[\![\alpha]\!]$. In fact $[\![\langle\alpha\rangle\varphi]\!] = (\mathsf{dom}(T) \cap [\![\langle\alpha\rangle\varphi]\!]) \cup (\mathsf{dom}(D(n)) \cap [\![\langle\alpha\rangle\varphi]\!]) \cup (\mathsf{dom}(I(k,n)) \cap [\![\langle\alpha\rangle\varphi]\!]) \cup (\mathsf{dom}(L(n,p,f)) \cap [\![\langle\alpha\rangle\varphi]\!])$. We have

$$\begin{aligned}
&\mathsf{dom}(T) \cap [\![\langle\alpha\rangle\varphi]\!] \subseteq [0,n)\\
&\mathsf{dom}(D(n)) \cap [\![\langle\alpha\rangle\varphi]\!] = [\![\varphi]\!] \cap [n,\omega)\\
&\mathsf{dom}(I(k,n)) \cap [\![\langle\alpha\rangle\varphi]\!] = \{x \mid x \geq \min([\![\varphi]\!] \cap [n,\omega)) + k\}\\
&\mathsf{dom}(L(n,p,f)) \cap [\![\langle\alpha\rangle\varphi]\!] = \{x \mid x \geq \min\{f(y) \mid y \in [\![\varphi]\!] \cap [0,n)\}\}
\end{aligned}$$

As a consequence, $[\![\langle\alpha\rangle\varphi]\!] \in \mathsf{ADMS}(\max(n+k,p)) \subseteq \mathsf{ADMS}$.

This concludes the proof of Lemma 3.7. □

Now we are ready to show that as far as formulas are concerned, ω is the limit of the natural numbers.

Lemma 3.8
For every $\varphi \in \mathscr{L}_{\emptyset,a}$ and for all programs $\alpha \in \mathscr{L}_{\emptyset,a}$:

$$\omega \models \varphi \text{ iff } [\![\varphi]\!] \text{ is cofinite} \quad (1)$$
$$\omega R(\alpha) y \text{ iff } \{x \mid (x,y) \in [\![\alpha]\!]\} \text{ is cofinite} \quad (2)$$
$$\omega R(\alpha) \omega \text{ iff } \{x \mid (x,x) \in [\![\alpha]\!]\} \text{ is cofinite} \quad (3)$$

Proof The proof is by induction on the structure of φ and α simultaneously. For formulas the atomic case and the cases for negation and conjunction follow directly from Lemma 3.7. In the case for modal formulas $[\alpha]\varphi$ we take the dual formula $\langle\alpha\rangle\varphi$. For the induction hypothesis, suppose that (1),(2), and (3) hold for φ and α. It follows from the semantics that $\omega \models \langle\alpha\rangle\varphi$ is equivalent with

there is a $y \in \mathbb{N}$ such that $\omega R(\alpha) y$ and $y \models \varphi$ or $\omega R(\alpha)\omega$ and $\omega \models \varphi$

By the induction hypotheses this is equivalent with

$$\exists y \in \mathbb{N}(\{x \mid (x,y) \in [\![\alpha]\!]\} \text{ is cofinite and } y \models \varphi, \text{ or}$$
$$\{x \mid (x,x) \in [\![\alpha]\!]\} \text{ is cofinite and } [\![\varphi]\!] \text{ is cofinite}$$

It is clear that this implies that $[\![\langle\alpha\rangle\varphi]\!]$ is cofinite. In the first case the set of worlds that can reach y is cofinite. In the second case the set of worlds where φ holds that can reach themselves is cofinite.

To see this the other way around follows from properties of ADMR. One mainly needs the property that if the domain of α is cofinite, then its intersection with $D(0)$ is cofinite (i.e. $\{x \mid (x,x) \in [\![\alpha]\!]\}$ is cofinite, or there is a world y such that $\{x \mid (x,y) \in betekenis\alpha\}$ is cofinite.

For (2) the case from right to left with $\alpha = \beta;\gamma$ is the most complicated. Assume $\{x \mid (x,y) \in [\![\beta;\gamma]\!]\}$ is cofinite, therefore the domain of $[\![\beta]\!]$ is cofinite. If $\{x \mid (x,y) \in [\![\gamma]\!]\}$ finite, then there must be a $z \in \mathbb{N}$ such that $\{x \mid (x,z) \in [\![\beta]\!]\}$ and $zR(\gamma)y$, therefore we can conclude $\omega R(\beta;\gamma)y$ using the induction hypothesis (2) for β. On the other hand if $\{x \mid (x,y) \in [\![\gamma]\!]\}$ is cofinite, then it follows from the induction hypothesis (2) for γ that $\omega R(\gamma)y$. If $\{x \mid (x,x) \in [\![\beta]\!]\}$ is cofinite, then from the induction hypothesis (3) for β it follows that $\omega R(\beta;\gamma)y$. Otherwise there is a $z \in \mathbb{N}$ such that $\{x \mid (x,z) \in [\![\beta]\!]\}$ is cofinite and we can find one such that $zR(\gamma)y$. Therefore also in this case $\omega R(\beta;\gamma)y$ using the induction hypothesis for β.

The proof of (3) is not too difficult and we do not provide details here. □

Theorem 3.3
The canonical model for PDL_ω does not have program harmony □

Proof This follows directly from Lemma 3.5 and Lemma 3.8 together with the observation that if $\eta \models [a^*]\varphi$, then $[\![\varphi]\!] = \mathbb{N}$ and therefore cofinite. □

3.7 Conclusion

In this chapter we have presented a proof system for propositional dynamic logic which is strongly complete. The method can be generalized and applied to other denumerably axiomatized modal logics as well as epistemic logic with common knowledge.

We suspect that the reason that the canonical model method works for this axiomatization, is that the infinitary *-introduction rule is much closer to the semantics than the usual induction axiom or rule. The latter links up with the idea of the Kleene star as a fixed point, whereas our rule links up with the idea that it is an infinitary conjunction. However, as we showed in section 3.6, there is no complete harmony between the proof system and the semantics. The countermodel used in the completeness proof has *formula harmony*. This is shown in the truth lemma. *Program harmony* is unattainable. To our surprise it was not needed for the completeness proof. To our astonishment it was not even true. Although we came up with a countermodel rather quickly, the subtlety of the arguments involved, was also unexpected. It would be interesting to try to construct a fully harmonious canonical model for PDL_ω. As of yet we did not find one in the literature.

There still remain some issues that need to be investigated further. Propositional dynamic logic and epistemic logic with common knowledge are examples where the introduction of an infinitary rule can be used to attain strong completeness, although the logics are not semantically compact. It should be investigated how to characterize the class of non-compact logics where the introduction of such an infinitary rule can also lead to a strong completeness result. The general approach of Goldblatt (1993) seems to be a good starting point.

Another interesting issue is whether the relation $\Gamma \vdash \varphi$ between recursively enumerable sets of formulas Γ and formulas φ is decidable.

Chapter 4

Information change

4.1 Introduction

Information pervades so many aspects of our daily lives that the present age has been dubbed 'the information age'. During the first Gulf War in 1991, operation Desert Storm started with Apache helicopters destroying radar defense systems. China has recently denied its citizens access to the Internet search engine Google. The international forum on urban poverty has put information and communication technologies on the agenda of fighting poverty. The ability of human beings to use information and to communicate is one of the most empowering abilities that exist. Information and power are strictly related in the information age.

Just like many other aspects of our daily lives information enjoys the interest of scientists. There are many scientific theories about information: information theory, situation theory, probability theory, statistics, computer science, game theory, philosophy of science, logic. All seem to focus on a different aspect of information. It is debatable whether a unified theory of information can ever be attained or whether it would even be worthwhile. The focus on higher-order information is distinctive for (dynamic) epistemic logic. Let us take a brief look at some of the theories about information to see what the merits of each of these theories are.

The oldest scientific theory that explicitly mentions information is *information theory*. Originally developed by Shannon (1948), it deals with quantitative questions about information. How much information does a message contain? How much information can be communicated? How can information be sent efficiently? Information theory was generalized by Kolmogorov (1956). This area is now known as Kolmogorov Complexity. For a modern introduction to information theory or Kolmogorov complexity see Cover and Thomas (1991) and Li and Vitányi (1993) respectively.

Situation theory deals with fundamental issues of information and information flow. This research program, initiated by Barwise and Perry (1983), was inspired by problems in the semantics of natural language. Situation theory deals with the question what information actually is and how it is possible that information is passed at all. How can it be that one thing carries information about another thing?

A discipline that looks at information change from a somewhat higher viewpoint is *philosophy of science*. Science is by no means static; scientific theories often change. Theories can even be replaced entirely. Philosophy of science tries to answer the question how such changes occur and whether such changes are for the best.

Probability theory and *Statistics* also deal with information change. A series of observations tells you something about the world. And although you have only made a limited number of observations, there are still general statements that are justified by these observations. The question is how certain you are about these general statements. I will come back to probability theory and statistics in chapter 5 and 6.

Computer science deals with information change because computers communicate with each other. In that case information is communicated. Many of the communication protocols used on the Internet were developed by computer scientists to ensure 'good' communication.

In games information change also occurs. Moves in a game can change physical reality or they can change the information the players have. In *game theory* a distinction is made between games of perfect information (such as chess) and games of imperfect information (such as poker). When a game of imperfect information is represented by a tree, players are uncertain about which node of the tree they are in. In games of perfect information they know exactly where they are. The set of nodes a player cannot distinguish thus represents the uncertainty of a player, which can be seen as representing the (lack of) information a player has.

From a meta-perspective *logic* can also be seen as a study of information. Logic focuses on inferences and their validity. The premises represent some of the information of the person performing the inference. And the inference shows that the conclusion follows from these and that in a sense the conclusion also represents some of the information of that person.

Within logic there are fields that deal with information explicitly. Especially *belief revision* and *epistemic logic* deal with information. As a discipline of logic belief revision began with the famous paper by Alchourrón, Gärdenfors, and Makinson (1985), which was followed by many papers that tried to refine it and expand on it. Belief revision tries to model how new information is processed. It focuses especially on non-monotonic reasoning and processing information that is inconsistent with the information a system already had. See Gärdenfors and Rott (1995) for a general introduction. *Epistemic logic* is a modal logic, initially developed by Hintikka (1962). His main goal was a conceptual analysis of knowledge and belief. Epistemic logic typically deals with what an agent considers to be possible given his current information. This information also contains information about information other agents have, because epistemic logic is suited to deal with situations involving more than one agent. In this way epistemic logic also deals with *higher-order information*, i.e. information about information. The ability to deal with higher-order information is what distinguishes epistemic logic from most other approaches to information. Let us look at a number of problems that are often used as examples that help to explain the notion of higher-order information. These examples also show that it is sometimes necessary to take higher-order information into account in order to give a satisfactory analysis of a situation or a problem.

4.1.1 Muddy children

The muddy children puzzle is one of the best known puzzles having to do with higher-order information. It is known that versions of this puzzle were circulating in the fifties. The earliest source of the puzzle I could find is a puzzle book by Gamow and Stern (1958). They present the 'cheating wives' version.

> The great Sultan Ibn-al-Kuz was very much worried about the large number of unfaithful wives among the population of his capital city. There were forty women who were openly deceiving their husbands, but, as often happens, although all these cases were a matter of common knowledge, the husbands in question were ignorant of their wives' behavior. In order to punish the wretched women, the sultan issued a proclamation which permitted the husbands of unfaithful wives to kill them, provided, however, that they were quite sure of the infidelity. The proclamation did not mention either the number or the names of the wives known to be unfaithful; it merely stated that such cases were known in the city and suggested that the husbands do something about it. However, to the great surprise of the entire legislative body and the city police, no wife killings were reported on the day of the proclamation, or on the days that followed. In fact, an entire month passed without any result, and it seemed the deceived husbands just did not care to save their honor.
>
> "O Great Sultan," said the vizier to Ibn-al-Kuz, "shouldn't we announce the names of the forty unfaithful wives, if the husbands are too lazy to pursue the cases themselves?"
>
> "No," said the sultan. "Let us wait. My people may be lazy, but they are certainly very intelligent and wise. I am sure action will be taken very soon."
>
> And, indeed, on the fortieth day after the proclamation, action suddenly broke out. That single night forty women were killed, and a quick check revealed that they were the forty who were known to

> have been deceiving their husbands. (Gamow and Stern (1958, pp. 20 – 21)[1])

Apparently this version was not considered politically correct, because in the best known version of the puzzle, there are children whose forehead may be muddy or not and a father announces that at least one of them is muddy and that those children who are sure that their own face is muddy should step forward. If there are n children, then the n-th time he announces this all muddy children step forward.

To solve this puzzle a number of additional assumptions have to be made. All children are perfect logicians and this is common knowledge. Moreover every child can see all the other children and this is also common knowledge. Why the children step forward can be seen by induction. If there is only one child, then this child will step forward after the first announcement. Assume that if there are n for $n \geq 1$ muddy children, they will all step forward after the n-th announcement. Now, if there are $n + 1$ muddy children, all of these children will reason as follows: 'I can see n muddy children. If I am not muddy they will all step forward after the n-th announcement. If they do not step forward, then there are not n muddy children. Therefore there must be one more. It's me!' Therefore all $n + 1$ muddy children will step forward the $n + 1$-th time the announcement is made.

It may seem counterintuitive that the repetition of the same announcement can give information. The repetition of the announcement itself does not give any additional information. The fact that the children do not step forward gives information. By doing nothing they all say that they do not know whether they are muddy or not, this then becomes common knowledge. So even if nothing happens this still gives information. Moreover all children can see this. One may wonder how information about what information the children have can help a child to decide what the actual situation is. This is because the actual situation determines what information the children have, and therefore knowing what everyone knows gives information about the situation.

4.1.2 Byzantine generals

An example of a problem that is often used to illustrate the problems for good communication protocols is the Byzantine generals problem. It is also known as the generals paradox or the coordinated attack problem. The earliest reference to the problem I could find was in Gray (1978). But it is probably older than that.

> There are two generals on campaign. They have an objective (a hill) which they want to capture. If they *simultaneously* march on the

[1]Erik Krabbe pointed out to me that Gamow and Stern made a slight error and should have said action suddenly broke out on the thirty-ninth day instead of the fortieth.

> objective they are assured of success. If only one marches, he will be annihilated.
>
> The generals are encamped only a short distance apart, but due to technical difficulties, they can communicate only via runners. These messengers have a flaw, every time they venture out of camp they stand some chance of getting lost (they are not very smart.)
>
> The problem is to find some protocol which allows the generals to march together even though some messengers get lost. (Gray (1978, p.465))

There is no real solution to this problem, because the generals are never sure that a messenger arrives. Let us call these generals Belisarius and Narses. Suppose that one of Belisarius' messengers arrives at Narses' camp. Belisarius will not know his messenger has arrived and Narses knows this. The only way for Belisarius to obtain the information that his messenger has arrived is by getting a message from Narses that states that his messenger has arrived. But when that messenger arrives, Narses does not know that he has. Therefore the problem still remains, because it will never be common knowledge among Belisarius and Narses that the first messenger arrived. The same problem occurs in communication between computers over the Internet for instance, where the computers can be thought of as the generals and the messengers are packages of information sent over the Internet. There are protocols that ensure messages get across, but additional safety conditions (certain events occur the way they should) and liveness conditions (certain events eventually occur) have to be met.

4.1.3 Sum and product

There are two people in a room: Mr. Sum and Mr. Product, who do not know the length or width of the room. They do know that these are both natural numbers between 2 and 99. Moreover the length is larger than the width. ($2 < w < l < 99$.) The sum of these two number is given to Mr. Sum, and the product of these two numbers is given to Mr. Product. And all this is common knowledge. The following conversation takes place:

Mr. Product:	I don't know what the numbers are.
Mr. Sum:	I knew you didn't know them.
Mr. Product:	Now I know what they are!
Mr. Sum:	Now I know them too!

The length and the width of the room can be deduced from the previous dialogue by an outsider.

The original formulation and solution of the problem can be found in Freudenthal (1969) and (1970) in Dutch. See McCarthy (1990) for a formulation and solution in English.

4.1.4 Cluedo

The game cluedo served as the main inspiration for the dissertation *Knowledge Games* by van Ditmarsch (2000). Cluedo has many specific features, but most important is that it is a *knowledge game*. A knowledge game is a game where the state of the world does not change, but only what the players know about it changes. In fact in many of these games the goal of the game is to find out what the state of the world is.

There are many actions in such a game that have epistemic consequence, such as answers to questions which are audible for all players. The more interesting actions are those where one player is shown a card and the other can see that a card is being shown, but they cannot see which card it is. But one of the most interesting actions in this game is the action of not winning. At the end of a turn a player who knows the state of the world can say that she knows. If she does not say so, this means she does not know the state of the world. Just like in the muddy children example it may seem that nothing happens, but information is exchanged.

4.1.5 Lecture or Amsterdam

Suppose Anne and Bert are sitting at a table in a bar. A messenger arrives and gives a letter to Anne. She tells them that the letter either contains an invitation to give a lecture or an invitation for a night out in Amsterdam. Various scenarios could unfold. Anne could read the letter out loud, so that its contents would become common knowledge among Anne and Bert. She could read the letter to herself, so that Bert would not know what the letter said, but she would, and Bert would know that she would know. Bert could also respect Anne's privacy and order a drink at the bar so that she could read the letter alone. When he would return he would not know whether Anne read the letter or not. They could also both leave the table simultaneously and get back to the table each suspecting the other person to have read the letter. There are many more possibilities.

This it is not a problem in the sense that the muddy children puzzle is a problem. Most people have good intuitions about what the effects of these actions are. It is a problem to describe these actions within a formal framework. In the scenarios above there are many actions by which Anne could have learned what the letter says, but the settings are all different. How can we give an account of this? I will use this setting as a running example in this chapter.

What all these examples have in common is that higher-order information has to be included if a satisfactory analysis is to be given. Moreover in many of these examples information change occurs. In any formal system information change also has to be taken into account, including the way higher-order information changes.

4.2 Multi-agent systems

One of the most influential attempts to model information change in the context of epistemic logic is due to Fagin, Halpern, Moses, and Vardi (1995). They are interested in multi-agent systems. These are systems where many agents interact. They are not exactly clear on what counts as an agent, but they seem to take a very broad approach and consider computer processing units and robots as well as persons as agents. The focus is on distributed systems (systems of communicating computer processors.) Consequently much of their terminology stems from computer science. The main aim of the authors is to be able to reason about what goes on in these systems from the outside. Epistemic logic on its own cannot provide a good model for multi-agent systems, because the states the agents of the system are in, must be capable of change, as a result of interaction for example.

In multi-agent systems agents can find themselves in any of a number of states. These are also called local states, because they involve only one agent. In the case of distributed systems, local states can be thought of as states of a processor's memory. But if you want to model a game like poker, for instance, the local states can be thought of as the cards an agent holds. To give a complete picture of a system the environment has to be taken into account as well. A state of the environment in a distributed system might yield information about whether a certain communication line is working or not. In case of a poker game the state of the environment might consist of the cards that are still in the deck on the table. In general there is one set of states for the environment S_e and there is one set of local states S_{a_i} for each agent a_i where the set of agents is $\mathcal{A} = \{a_1, \ldots, a_n\}$.

$$\begin{array}{lcl} S_e & = & \{s \mid s \text{ is a state of the enviroment}\} \\ S_{a_i} & = & \{s \mid s \text{ is a local state of agent } a_i\} \end{array}$$

A global state of the system is nothing more than a state of the environment combined with the local states of the agents. A global state is a tuple $(s_e, s_1, \ldots, s_n)$. The set of global states $\mathcal{G}$ of a system is defined as:

$$\mathcal{G} = S_e \times S_{a_1} \times \cdots \times S_{a_n}$$

As was noted above, multi-agent systems are subject to change. The state of the environment changes and the local states of the agents can change. So the global state of the system can change. One of the assumptions that is made is that time is discrete. Although this may give a distorted picture of time as it is usually conceived of, it is quite suitable for computer processors, because they change in discrete steps. The natural numbers are taken to model time. The advantage of the natural numbers is that they have a clear starting point, as do many computer programs.

A real system cannot simply go from any global state to any other. The ways in which a system can develop is usually limited. Such possible developments are called runs. A run is a function from the natural number to the set of global states of the system.

$$r : \mathbb{N} \to \mathcal{G}$$

A multi-agent system $\mathcal{R}$ is defined as a set of runs.

Definition 4.1 (Multi-agent system)
Let a set of global states $\mathcal{G} = S_e \times S_{a_1} \times \cdots \times S_{a_n}$ be given. A multi-agent system is a set

$$\mathcal{R} \subseteq \{r : \mathbb{N} \to \mathcal{G}\}$$

The global state of a system in a run r at time n is $r(n)$. A pair (r, n) consisting of a run r and a time n is called a point. The global state at a point (r, n) is $r(n)$. Points will be used as the possible worlds in a Kripke model. The local state of an agent a at a point (r, n) is indicated by $r_a(n)$ (note the subscript). □

In an interpreted system the truth value of every propositional variable in a global state is defined.

Definition 4.2 (Interpreted system)
Let a multi-agent system $\mathcal{R} \subseteq \{r : \mathbb{N} \to \mathcal{G}\}$ be given. Let a function π be given which assigns a valuation function to every global state.

$$\pi : \mathcal{G} \to (\mathcal{P} \to \{0, 1\})$$

An interpreted system $\mathcal{I}$ is defined to be a pair $(\mathcal{R}, \pi)$. □

In order to introduce the concept of an agent's knowledge in an interpreted system a Kripke model has to be defined based on the interpreted system. The set of points is taken as the set of possible worlds. The set of global states is not taken as the set of possible worlds, because a system can have the same global state in different runs and even at different times in the same run. The accessibility relation $R(a)$ of an agent a is defined as follows. A point (r', m') is accessible to a from a point (r, m) iff a has the same local state in both (r, m) and (r', m'). This construction ensures that the accessibility relations are equivalence relations. Hence the models are $\mathsf{S5}$ models.

Definition 4.3 (Kripke models for multi-agent systems)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set $\mathcal{A}$ of agents be given. Given an interpreted system $\mathcal{I}$ the Kripke model $M_{\mathcal{I}}$ is a triple (W, R, V) such that

- W is the set of points of $\mathcal{I}$
- R is a set of accessibility relations that contains an accessibility relation $R(a) \subseteq W \times W$ for each agent $a \in \mathcal{A}$ such that $(r, m)R(a)(r', m')$ iff $r_a(m) = r'_a(m')$

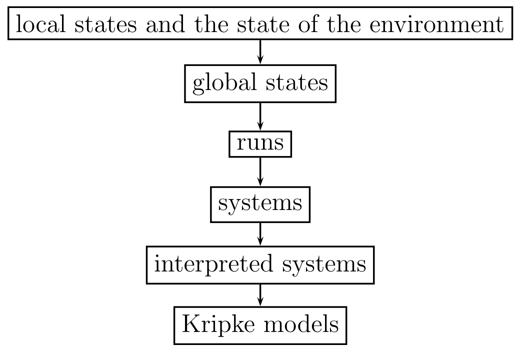


Figure 4.1: The construction of the logic for multi-agent systems

- V is a function that assigns a set of points to each proposition $p \in \mathcal{P}$ such that $V(p) = \{(r, m) \mid \pi(r(m))(p) = 1\}$. □

This construction is represented schematically in figure 4.1. The language of epistemic logic can now be extended with temporal operators in several ways to be able to reason about knowledge and time, and therefore about the change of knowledge over time along a run. We will not go into this here. For more information on temporal operators see Emerson (1990).

However, there is something that is not quite satisfactory about this way of reasoning about information and information change. It seems the focus in Fagin and Halpern (1994) is too much on the model theoretical aspects of multi-agent systems, rather than on the inferences made by the agents in such a system. The analysis of a multi-agent system in this framework often starts by defining the local states and continue from there as was indicated in figure 4.1. Then one can simply look which formulas hold in the resulting model. But one might also be interested in the question whether a certain inference is valid in a class of multi-agent systems.

Consider as an example a game such as Cluedo (see section 4.1.4). One can build a multi-agent system for this game. But it might be the case that one is not interested in the whole system, but just in what goes on in one move in the game. For example, one may wonder whether it generally holds that when one player, say a, shows one of his cards, say c, to another player, say b, that player b will know afterwards that player a has card c. It seems rather cumbersome to define the whole multi-agent system just to answer such a question. Moreover, that multi-agent system is specifically built for Cluedo. In the framework of Fagin and Halpern (1994) it seems difficult to answer the question whether such an inference is valid generally, not just for Cluedo, because one cannot refer to the act of showing a card in the logical language.

It also seems that as a model theoretical analysis the framework puts too much emphasis on the local states. The local states of the agents completely determine the accessibility relations of the agents. So the local states also determine higher-order information and the way higher-order information can change. It seems strange to put the way things can change in the states of the agents. When one wants to make a multi-agent system for Cluedo, one firstly has to define the local states of the agents. It does not suffice to let the local state of an agent simply consist of the cards that agent is holding. In Cluedo cards never change hands, therefore the local state must also contain information about the history of the game. But in order to put information about the history of the game in the states, one has to have quite a clear picture of the game beforehand.

In the next section we will be looking at logics that deal with information change by having modalities for different actions. These modalities explicitly deal with higher-order information, and also deal with the interaction between actions with higher-order information aspects and states with higher-order information aspects.

4.3 Dynamic epistemic logic

Dynamic epistemic logic is a relatively new field of research. Its aim is to provide formal means of analysis of information change. It therefore consists of two parts. One part deals with information, the other with change. The part dealing with information is epistemic logic. So it also deals with higher-order information. The part dealing with change is very much like the programs from propositional dynamic logic PDL (see chapter 3), except that in this case there are programs that explicitly deal with higher-order information change. In fact, in many dynamic epistemic logics, the focus is on higher-order information change to such an extent that changes in the world itself are no longer considered. In most of the examples mentioned in section 4.1 the only changes occur in the information of the agents, but the world itself stays the same.

The transition relations of PDL can be seen as relations *between* propositional models. The same holds for first-order dynamic logic, where the programs could be seen as relations between first-order models. In the same way the programs of dynamic epistemic logic can be seen as relations between Kripke models. The question is in which relations one is specifically interested. For PDL it is quite clear that one is interested in being able to prove the correctness of certain types of programs. The same holds for first-order dynamic logic. But what are the interesting programs that we want to reason about with dynamic epistemic logic? In this section I give four examples of dynamic epistemic logics. In section 4.3.1 DEL is dicussed, which is one of the first dynamic epistemic logics. Secondly, in section 4.3.2, the dynamic epistemic logic KAL is discussed, which was developed to deal with knowledge only. In section 4.3.3 the logic LEA is introduced, which is one of the most original and influential approaches in dynamic epistemic logic. Lastly, in section 4.3.4, the logic DML is presented, which is not so much a dynamic epistemic logic as it is a general dynamic modal logic. The various logics presented in this section provide some examples of programs that deal with higher-order information change.

4.3.1 Updating with programs

One of the systems that was first studied is DEL, developed by Gerbrandy and Groeneveld (1997) and Gerbrandy (1999), which was inspired by the work by Veltman (1996). It is a system where besides the programs of PDL, 'updates' on programs are executed. The intuition behind it is quite clear. An update in this system means that a group of agent *consciously* learns that a program has been executed. "Consciously" means that in the resulting Kripke model, it is not just the case that they know that the program has been executed, they also know that all the members of the group know that the program has been executed, and so on. In fact, it is common knowledge among the members of the group that the program has been executed. So not only the information of the agents has been updated, but also their information about the information, and so on.

The language is defined as follows:

Definition 4.4 (Language of DEL)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of DEL $\mathscr{L}^U_{\mathcal{PA}}$ consists of a set of sentences φ and a set of actions α, given by the following rules in BNF:

$$\begin{array}{lcl} \varphi & ::= & \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a \varphi \mid [\alpha]\varphi \\ \alpha & ::= & ?\varphi \mid \alpha_1 ; \alpha_2 \mid \alpha_1 \cup \alpha_2 \mid U_{\mathcal{B}}\alpha \end{array}$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$, and $\mathcal{B} \subseteq \mathcal{A}$. □

A sentence of the form $[U_{\mathcal{B}}\alpha]\varphi$ is to be read as: 'after everyone in $\mathcal{B}$ has commonly learned that α has been executed, φ holds.'

This means that for a Kripke model a lot needs to change. Intuitively one would say that for agents who are not members of the group of updated agents nothing changes. It turns out to be quite difficult to obtain this property. The idea of a multi-agent model is that the accessibility relations model the information the agents have. If the information of an agent changes, one would expect that only the accessibility relations have to be altered. The problem for a multi-agent Kripke model is that an element of an accessibility relation (an arrow) serves many purposes. Consider the model of figure 4.2. Suppose agent a learns which world is the actual world. If one models this by removing all arrows for a except the reflexive arrows, the information of agent b also changes. In the model that results, b knows that a knows which world is the actual world. This is of course

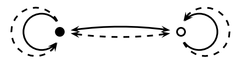


Figure 4.2: A multi-agent model for two agents. The solid node indicates that p is true and the open node indicates that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b.

due to the fact that the accessibility relations do not only model the information of the agents about the world, but also the information they have about each other. This makes Kripke models a very compact way of representation. But it makes it difficult to change it in such a way that all those levels of information in the situation you wish to model, are accurately described.

To solve this type of problem, Gerbrandy introduces a structure in which the different purposes of an element of an accessibility relation can be distinguished. This is done by using non-well-founded set theory, which was developed by Aczel (1988). In section 4.3.4 we will see a similar solution of this problem. The models for epistemic logic that Gerbrandy uses are called possibilities.

Definition 4.5 (Possibilities)
Let $\mathcal{A}$, a set of agents, and $\mathcal{P}$, a set of propositional variables, be given. The class of *possibilities* is the largest class such that:

- A possibility w is a function that assigns to each propositional variable $p \in \mathcal{P}$ a truth value $w(p) \in \{0, 1\}$ and to each agent $a \in \mathcal{A}$ an information state $w(a)$.

- An information state is a set of possibilities. (Gerbrandy (1999, p.12)) □

It is clear that this definition is circular. Possibilities are defined in terms of information states and information states are defined in terms of possibilities.
In non-well-founded set theory this is not a problem. The more modern versions of such definitions are given in terms of coalgebras. If we want to have an equivalent in ordinary set theory, the best way to think about possibilities is as trees. The relation between possibilities and Kripke models is the following. We can think of these trees as Kripke models quite easily; Kripke models are graphs, and a tree is just a special kind of graph. If we have a Kripke model with a specified world, we can unfold it into a tree, which can be viewed as a possibility. In figure 4.3 the possibility is shown, which represents the same situation as the Kripke model of figure 4.2. The model is unrolled for the agents separately.

The language of ordinary epistemic logic can be interpreted on possibilities. But because we now can distinguish the different purposes an element of an accessibility relation serves, we can define updates easier. How for example would

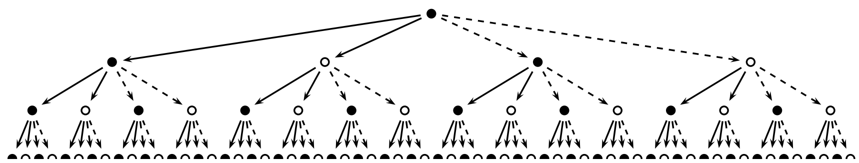


Figure 4.3: A picture of a part of the possibility bisimilar to the Kripke model of figure 4.2, where the left world of the Kripke model of figure 4.2 is taken to be the actual world. The solid nodes indicate that p is true and the open nodes indicate that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b.

one want to interpret that a group $\mathcal{B}$ commonly learns that a test on φ is successful? This means that, whereas the information states of all agents not in $\mathcal{B}$ do not change, the information states for members of $\mathcal{B}$ are changed such that those possibilities where φ does not hold are removed. Then the remaining possibilities in the information states of members of $\mathcal{B}$ are updated. This is done to make sure that the update is *conscious*, which means that the information about information is also changed.

This is an example of how an update with a test is executed, but in general an update can be executed on programs. We have the following semantics for the language of DEL:

Definition 4.6 (Semantics for $\mathscr{L}^{U}_{\mathcal{PA}}$)
Given a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$, let w be a possibility, let $p \in \mathcal{P}$ and φ, ψ be sentences of $\mathscr{L}^{U}_{\mathcal{PA}}$ and α be an action of $\mathscr{L}^{U}_{\mathcal{PA}}$

$$
\begin{array}{lll}
w \not\models \bot & & \\
w \models p & \text{iff} & w(p) = 1 \\
w \models \varphi \wedge \psi & \text{iff} & w \models \varphi \text{ and } w \models \psi \\
w \models \neg\varphi & \text{iff} & w \not\models \varphi \\
w \models \Box_a \varphi & \text{iff} & \text{for all } v \in w(a) : v \models \varphi \\
w \models [\alpha]\varphi & \text{iff} & \text{for all } v\text{: if } w[\![\alpha]\!]v \text{ then } v \models \varphi \\
 & & \\
w[\![?\varphi]\!]v & \text{iff} & w \models \varphi \text{ and } w = v \\
w[\![U_{\mathcal{B}}\alpha]\!]v & \text{iff} & w \restriction ((\mathcal{A} \setminus \mathcal{B}) \cup \mathcal{P}) = v \restriction ((\mathcal{A} \setminus \mathcal{B}) \cup \mathcal{P}) \text{ and} \\
 & & \text{for all } a \in \mathcal{B} : v(a) = \{v' \mid \exists w' \in w(a) \exists u : w'[\![\alpha]\!]u[\![U_a\alpha]\!]v'\} \\
w[\![\alpha;\alpha']\!]v & \text{iff} & \text{there is a } u \text{ such that } w[\![\alpha]\!]u[\![\alpha']\!]v \\
w[\![\alpha \cup \alpha']\!]v & \text{iff} & w[\![\alpha]\!]v \text{ or } w[\![\alpha']\!]v
\end{array}
$$

Consider the following example to see how this definition works. Let us take the possibility given in figure 4.3. Let us say that in the filled nodes the propositional

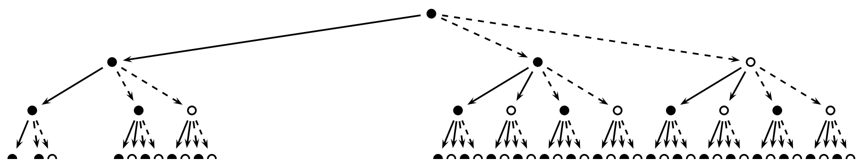


Figure 4.4: A picture of the possibility obtained by executing $U_a?p$ in the possibility of figure 4.3. The solid nodes indicate that p is true and the open nodes indicate that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b.

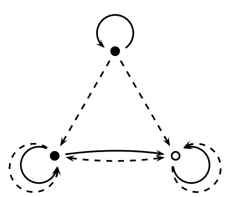


Figure 4.5: A picture of the Kripke model that can be unfolded to the possibility of figure 4.4. The solid nodes indicate that p is true and the open node indicates that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b.

variable p is true and that p is false in the open nodes. Suppose the update $U_a?p$ is executed. For agent b nothing changes, because only a is involved. For a those possibilities where p is not true are removed from her information state. Now the same happens to the possibilities remaining in her information state. See figure 4.4. See figure 4.5 for the corresponding Kripke model. In Gerbrandy (1999) a sound and complete proof system with respect to the semantics of definition 4.6 is also provided.

Definition 4.7 (Proof system DEL)
Let φ be a sentence in $\mathscr{L}^U_{\mathcal{PA}}$ and let α be an action in $\mathscr{L}^U_{\mathcal{PA}}$. The proof system DEL consists of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}$ and the following axioms

Distr	$[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$	(distribution)
?**AX**	$[?\varphi]\psi \leftrightarrow (\varphi \to \psi)$	(test)
F	$\neg[U_\mathcal{B}\alpha]\varphi \leftrightarrow [U_\mathcal{B}\alpha]\neg\varphi$	(functionality)
AP	$[U_\mathcal{B}\alpha]p \leftrightarrow p$	(atomic permanence)
RAX	$[U_\mathcal{B}\alpha]\Box_a\varphi \leftrightarrow \Box_a[\alpha][U_\mathcal{B}\alpha]\varphi$ if $a \in \mathcal{B}$	(generalized Ramsey axiom)

$$
\begin{array}{lll}
\mathbf{P} & [U_{\mathcal{B}}\alpha]\Box_a\varphi \leftrightarrow \Box_a\varphi \text{ if } a \notin \mathcal{B} & \text{(privacy)} \\
;\mathbf{AX} & [\alpha;\beta]\varphi \leftrightarrow [\alpha][\beta]\varphi & \\
\cup\mathbf{AX} & [\alpha \cup \beta]\varphi \leftrightarrow ([\alpha]\varphi \wedge [\beta]\varphi) &
\end{array}
$$

This system can be extended to DELD45 and DELS5 by adding the usual axioms (see chapter 2). Completeness can be proved using a canonical model method. But a simpler completeness proof uses a translation of $\mathscr{L}^U_{\mathcal{PA}}$ into $\mathscr{L}_{\mathcal{PA}}$, i.e. all sentences with dynamic operators can be translated into sentences without any dynamic operators is such a way that any sentence is provably equivalent to its translation. The translation simply follows the axioms above.

There are still however some problems with this system. The first problem is the semantics. The use of non-well-founded set theory is troublesome.

The main shortcoming of this system is that common knowledge is not part of the system. One of the aspects that is so interesting in the context of higher-order information is common knowledge. Especially in games such as Cluedo it plays such an important role that you want to know how common knowledge changes. The difficulty with adding common knowledge is that a completeness proof with a translation is not that easy anymore. One could add it and try to prove completeness in the same way it is proved for the system presented in section 4.3.3.

Another problem is that updates construed as in DEL can sometimes result in a non-reflexive model, whereas one started out with a reflexive one. For example when an update occurs with a sentence that is not true in the actual world, the result of the update will not be S5. This means that if you want to preserve the validity of the S5 axioms, you need to be careful. The proof system does not have a necessitation rule for $[\alpha]$ operators, because of this. In DELS5 necessitation is not sound. Nevertheless for many programs necessitation is a derived rule.

4.3.2 Learning to preserve S5

Hans van Ditmarsch provided a system of dynamic epistemic logic specifically developed for knowledge, i.e. S5 (see chapter 2). In this system, which I call KAL (knowledge action logic), the execution of actions preserves the S5 properties. Moreover KAL works with ordinary Kripke models and not with non-well-founded set-theoretic objects. The main inspiration is the game Cluedo, which is an example of a game that KAL can be applied to. This section uses the definitions as provided in van Ditmarsch (2002).

The set of programs is an extension of PDL without the Kleene star. The language is defined as follows:

Definition 4.8 (Language of KAL)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of KAL $\mathscr{L}^L_{\mathcal{PA}}$ consists of a set of sentences φ and a set of

actions α, given by the following rules in BNF:

$$\begin{array}{lll}\varphi & ::= & \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a\varphi \mid C_{\mathcal{B}}\varphi \mid [\alpha]\varphi \\ \alpha & ::= & ?\varphi \mid \alpha_1;\alpha_2 \mid \alpha_1 \cup \alpha_2 \mid \alpha_1!\alpha_2 \mid L_{\mathcal{B}}\alpha\end{array}$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$, and $\mathcal{B} \subseteq \mathcal{A}$. We call the φ expressions the sentences of $\mathscr{L}^L_{\mathcal{PA}}$ and the α expressions the actions of $\mathscr{L}^L_{\mathcal{PA}}$. □

The additional operators are the local choice operator (the exclamation mark) and the learn operator (L). The local choice operator is a bit strange when studied in isolation, but its use becomes clear when it is applied in combination with the learning operator. The meaning of the local choice operator ($\alpha!\beta$) in isolation is 'from α and β choose α locally', so the interpretation of $\alpha!\beta$ is equal to the interpretation of α. The learn operator $L_{\mathcal{B}}\alpha$ is interpreted as that it is common knowledge among all members of $\mathcal{B}$ that α is being executed. But the agents in $\mathcal{B}$ do not necessarily know exactly what α is. For example if one agent shows another agent a card and a third agent can see this, but he cannot see which card is being shown, then he does not know exactly what action is executed, but he does know what *type* of action is executed. This can be formalized as follows.

Definition 4.9 (Action type)
The action type is a function on the actions of the language of KAL: $t : \{\alpha \mid \alpha \in \mathscr{L}^L_{\mathcal{PA}}\} \to \{\alpha \mid \alpha \in \mathscr{L}^L_{\mathcal{PA}}\}\mathscr{L}^L_{\mathcal{PA}}$.

$$\begin{array}{lll} t(?\varphi) & = & ?\varphi \\ t(\alpha \cup \alpha') & = & t(\alpha) \cup t(\alpha') \\ t(\alpha;\alpha') & = & t(\alpha);t(\alpha') \\ t(\alpha!\alpha') & = & t(\alpha) \cup t(\alpha') \\ t(L_{\mathcal{B}}\alpha) & = & L_{\mathcal{B}}t(\alpha) \end{array}$$

So basically, the type of α is the expression that is obtained by replacing all occurrences of the local choice operator with nondeterministic choices, except those that occur in the scope of a test.

Let us look at the truth definition to see the interplay between the local choice operator and the learning operator.

Definition 4.10 (Semantics for $\mathscr{L}^L_{\mathcal{PA}}$)
Let a model $M = (W, R, V)$ in $S5_{\mathcal{PA}}$ and a world $w \in W$ be given. Let $p \in \mathcal{P}$, and let φ and ψ be sentences and α be an action of $\mathscr{L}^L_{\mathcal{PA}}$.

$$\begin{array}{lll} (M,w) \not\models \bot & & \\ (M,w) \models p & \text{iff} & w \in V(p) \\ (M,w) \models \neg\varphi & \text{iff} & (M,w) \not\models \varphi \\ (M,w) \models \varphi \wedge \psi & \text{iff} & (M,w) \models \varphi \text{ and } (M,w) \models \psi \\ (M,w) \models \Box_a\varphi & \text{iff} & (M,v) \models \varphi \text{ for all } v \text{ such that } wR(a)v \\ (M,w) \models C_{\mathcal{B}}\varphi & \text{iff} & (M,v) \models \varphi \text{ for all } v \text{ such that } wR(\mathcal{B})^+v \text{ see section 2.4} \\ (M,w) \models [\alpha]\varphi & \text{iff} & (M',w') \models \varphi \text{ for all } (M',w') \text{ such that } (M,w)[\![\alpha]\!](M',w') \end{array}$$

Let $W_\varphi = \{v \in W \mid (M, w) \models \varphi\}$.

$$
\begin{array}{lll}
(M,w)[\![?\varphi]\!](M',w') & \text{iff} & M' = (W_\varphi, \emptyset, V \upharpoonright W_\varphi) \text{ and } w' = w \\
(M,w)[\![\alpha \cup \alpha']\!](M',w') & \text{iff} & (M,w)[\![\alpha]\!](M',w') \text{ or } (M,w)[\![\alpha']\!])(M',w') \\
(M,w)[\![\alpha;\alpha']\!](M',w') & \text{iff} & (M,w)([\![\alpha]\!] \circ [\![\alpha']\!])(M',w') \\
(M,w)[\![\alpha!\alpha']\!](M',w') & \text{iff} & (M,w)[\![\alpha]\!](M',w') \\
(M,w)[\![L_\mathcal{B}\alpha]\!](M',w') & \text{iff} & M' \in S5_{\mathcal{PB}} \text{ and } (M,w)[\![\alpha]\!]w' \text{ and } M' = (W',R',V')
\end{array}
$$

where in the last clause for the interpretation of $L_\mathcal{B}\alpha$, the set of worlds W', the accessibility relations R', and the valuation V' are defined as follows:

- $W' = \{(M'', v'') \mid \exists v \in W : wR(\mathcal{B})^+v \text{ and } (M,v)[\![t(\alpha)]\!](M'',v'')\}$
- $(M'',v'')R'(a)(M''',u''')$ iff there is a u'' such that $v''R''(a)u''$ and $(M'',u'') \leftrightarrows (M''',u''')$ see definition 2.8, or $a \notin (gr(M'') \cup gr(M'''))$ and there are $v, u \in W$ such that $vR(a)u$ and $(M,v)[\![t(\alpha)]\!](M'',v'')$ and $(M,u)[\![t(\alpha)]\!](M'',u'')$
- $V'(p) = \{(M'',v'') \in W' \mid (M'',v'') \models p\}$

where in the definition of R' the group of a model $gr(M)$ with $M = (W, R, V)$ is the domain of R. □

Note that the relations associated with actions are from $S5_{\mathcal{PA}}$-models to $S5_{\mathcal{PB}}$-models where $\mathcal{B} \subseteq \mathcal{A}$. In the extreme case of a test we end up in an $S5_{\mathcal{P}\emptyset}$-model. It may seem strange, but by defining it in this way it is ensured that the system is *entirely* S5. One could let the accessibility relation for all the other agents be empty, but then those relations are not reflexive. By leaving the relations undefined in such cases, the models are always S5.

A learn action works as follows. When a group $\mathcal{B}$ learns that α in a model (M, w) we move to a $S5_{\mathcal{PB}}$-model (M', w'). The worlds of this model are models themselves, namely those models that can be reached by executing an action of type α in a pointed model (M, v), such that v is accessible from w to the learning group of agents $\mathcal{B}$. The valuation of a propositional variable simply consists of those world-models where the propositional variable holds. For the accessibility relations, a world model (M''', u''') is accessible from (M'', v'') to a iff there is a world in M'' that is accessible from v'' to a and that world is bisimilar to (M''', u'''), or a is not involved in any of the actions of type α (hence $a \notin (gr(M'') \cup gr(M''')$, and the 'original' worlds were in the accessibility relation of a. Consequently agents that are not involved in α will know less about α than agents that are involved in it.

Consider again the example of figure 4.2, where p is true in the world on the left and false in the one on the right. Suppose that agent a now learns whether p holds and it is common knowledge among a and b that a learns this. This action can be described as $L_{ab}(L_a?p \cup L_a?\neg p)$. This action type has only one possible execution in each of the worlds of the model. Before we look at the effect of the entire program let us see what the effects of the subprograms are. The program $?p$ can only be executed in the left world. It yields a one-world model for the empty set of agents. The program $L_a?p$ is interpreted as follows. The set of worlds are the models that result from executing $?p$, which is the model described above. The accessibility is only defined for a as he is the only agent involved in the update. The model is accessible to him from the model, because the world was accessible to him in the original model and he is not in the group of the submodel. The valuation remains the same. A similar model result from executing $L_a?\neg p$ in the other world. These two models are the worlds of the model that results from executing $L_{ab}(L_a?p \cup L_a?\neg p)$. For a only those models themselves are accessible from themselves, because every model is bisimilar to itself. Agent b cannot distinguish between the two because he could not distinguish the worlds in which the actions were executed. This is all shown in figure 4.6.

One of the problems of van Ditmarsch' approach is the non-standard interpretation of tests. The semantics requires that after an action we are in a model for the group of that action. Because a test has an empty group, executing a test will yield a borderline case of a multi-agent model: a model for no agents. The result of this is that not every sentence of the language can be interpreted. A sentence of the form $[?\psi]\Box_a\varphi$ cannot be interpreted, because $\Box_a\varphi$ cannot be interpreted in a model where a is not in the set of agents of the model. See for example the models in figure 4.6.

Until recently a sound and complete proof system is also lacking, but a lot of progress has been made by Van Ditmarsch, van der Hoek, and Kooi (2003).

4.3.3 Epistemic actions

When one thinks about *actions* that have epistemic aspects one notices that what is said about them, is quite similar to what is said about *situations* that have epistemic aspects. For example in the action $L_{ab}(L_a?p \cup L_a?\neg p)$ it is common knowledge among a and b that a knows whether a test on p is being executed or a test on $\neg p$ is being executed. In Baltag, Moss, and Solecki (1999) this idea is taken seriously. In their approach actions are seen as multi-agent Kripke models, where instead of a valuation of propositional variables there is a precondition function. The conceptual advantage of this is that actions are modeled in the same way as situations. This is very different from a PDL style dynamic logic.

Action models are defined as follows.

Definition 4.11 (Action models)
Let a finite set of agents $\mathcal{A}$ be given. An action model for $\mathscr{L}$ is a triple $\mathsf{M} = (\mathsf{W}, \mathsf{R}, \mathsf{pre})$ such that:

- $\mathsf{W} \neq \emptyset$; a set of action worlds;

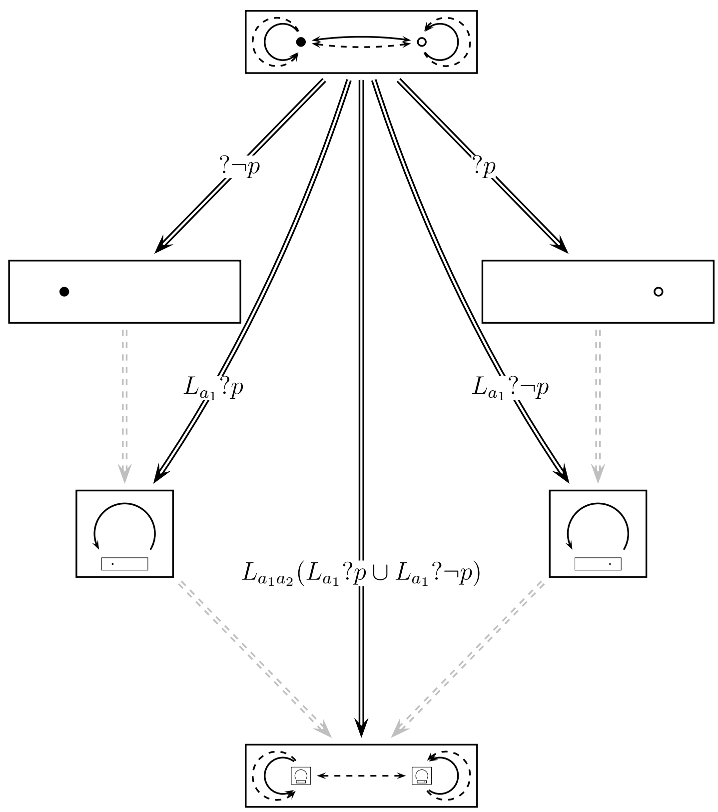


Figure 4.6: The result of executing various programs in Van Ditmarsch' approach. The two models on the left are the result of execution in the world on the left, the two models on the right are the result of executing in the world on the right. The model at the bottom is the result of execution in either world. The solid nodes indicate that p is true and the open nodes indicate that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b. The double solid lines represent the relations between the models associated with the actions on those lines. The grey, dashed, double arrows indicate that a model is a world in the model it is pointing to.

- $\mathsf{R} : \mathcal{A} \to 2^{\mathsf{W}\times\mathsf{W}}$; assigns an accessibility relation to each agent;
- $\mathsf{pre} : \mathsf{W} \to \mathscr{L}$; assigns a precondition to every action.
- $\mathsf{w} \in \mathsf{W}$

A pair (M, w) is a pointed action model. □

In this definition $\mathscr{L}$ is not specified. These models can be seen as special cases of the models of definition 2.2 (see page 4), where the set of propositional variables is $\mathscr{L}$. So we can interpret $\mathscr{L}_{\mathscr{L}\mathcal{A}}$ in these models and the class of all action models for $\mathscr{L}_{\mathscr{L}\mathcal{A}}$ is $K_{\mathscr{L}\mathcal{A}}$. For now it is only important that $\mathscr{L}$ is some language that can be interpreted in models for epistemic logic of definition 2.2 (for the rest of this section we will call those models static models to distinguish them from action models). The question is what the effect of executing an action is.

Definition 4.12 (Execution)
Given a static model (M, w) and and action model (M, w) such that $(M, w) \models \mathsf{pre}(\mathsf{w})$, we say that the result of executing (M, w) in (M, w) is $(M \cdot \mathsf{M}, (w, \mathsf{w})) = ((W', R', V'), (w, \mathsf{w}))$ where

- $W' = \{(v, \mathsf{v}) \mid (M, v) \models \mathsf{pre}(\mathsf{v})\}$
- $R'(a) = \{((v, \mathsf{v}), (u, \mathsf{u})) \mid vR(a)u \wedge \mathsf{v}\mathsf{R}(a)\mathsf{u}\}$
- $V'(u, \mathsf{v}) = V(u)$ □

It is a very attractive way to view the definitions above as the semantics of some logic. The question is which language is appropriate. In the logic of epistemic action (LEA), presented in Baltag, Moss, and Solecki (1999), these models are part of the dynamical logical language.

Definition 4.13 (Language of LEA)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of LEA $\mathscr{L}^{\mathsf{M}}_{\mathcal{P}\mathcal{A}}$ is given by the following rule:

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a\varphi \mid E_{\mathcal{B}}\varphi \mid C_{\mathcal{B}}\varphi \mid [(\mathsf{M}, \mathsf{w})]\varphi$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$, $\mathcal{B} \subseteq \mathcal{A}$ and (M, w) is a *finite* action model for $\mathscr{L}^{\mathsf{M}}_{\mathcal{P}\mathcal{A}}$. □

The way to read sentences of the form $[(\mathsf{M}, \mathsf{w})]\varphi$ is 'every execution of (M, w) yields a model where φ holds.' One of the features of this logic which is a problem for many other approaches is that common knowledge is also included in the language (see section 2.4). If one is interested in higher-order information, it is quite natural to incorporate this in the language.

Definition 4.14 (Semantics for $\mathscr{L}_{\mathcal{PA}}^{\mathsf{M}}$)
Let a static model (M, w) where $M = (W, R, V)$ be given.

$(M, w) \not\models \bot$		
$(M, w) \models p$	iff	$w \in V(p)$
$(M, w) \models \neg\varphi$	iff	$(M, w) \not\models \varphi$
$(M, w) \models (\varphi \wedge \psi)$	iff	$(M, w) \models \varphi$ and $(M, w) \models \psi$
$(M, w) \models \Box_a\varphi$	iff	$(M, v) \models \varphi$ for all v such that $wR(a)v$
$(M, w) \models E_{\mathcal{B}}\varphi$	iff	$(M, v) \models \varphi$ for all v such that $wR(\mathcal{B})v$see section 2.4
$(M, w) \models C_{\mathcal{B}}\varphi$	iff	$(M, v) \models \varphi$ for all v such that $wR(\mathcal{B})^*v$see section 2.4
$(M, w) \models [(\mathsf{M}, \mathsf{w})]\varphi$	iff	$(M, w) \models \mathsf{pre}(\mathsf{w})$ implies that $(M \cdot \mathsf{M}, (w, \mathsf{w})) \models \varphi$

where $R(\mathcal{B})^*$ is the reflexive transitive closure of $R(\mathcal{B})$.

Note that the interpretation of common knowledge is the *reflexive* transitive closure contrary to the definition of common knowledge in chapter 2. As an example consider the model given in figure 4.2 again. Suppose now that a learns whether p holds, by being whispered something in his ear. Agent b can see this but he does not know whether a was given any information regarding p. Let us assume that any information given is true and that this is common knowledge among a and b. The way this action is modeled is as an action model with three worlds, one where a learns that p, one where he learns that $\neg p$, and one where he does not learn anything about p. Agent a can distinguish all these actions, but b cannot. This action model can be multiplied with the static model given in figure 4.2. This is shown in figure 4.7. Executing this action works as follows. The resulting model contains four worlds. Those actions where a learns something about p can only be executed in the worlds where that information is true. The action where a learns nothing about p can be executed in both worlds. These are the four worlds in the resulting model. As agent b considered every world possible in the static model as well as in the action model, this also holds in the resulting model. Agent a knows exactly which action was executed. The valuation is copied from the static model.

In Baltag, Moss, and Solecki (1999) a proof system for LEA is also provided. In the next definition the notation has been adapted to the notation used in this chapter.

Definition 4.15 (Proof system LEA)
Let φ be a sentence in $\mathscr{L}_{\mathcal{PA}}^{\mathsf{M}}$ and let (M, w) be an action in $\mathscr{L}_{\mathcal{PA}}^{\mathsf{M}}$. The proof system LEA consists of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}^{EC}$ (see section 2.4) and the following axioms and rule:

Distr	$[(\mathsf{M}, \mathsf{w})](\varphi \to \psi) \to ([(\mathsf{M}, \mathsf{w})]\varphi \to [(\mathsf{M}, \mathsf{w})]\psi)$	(distribution)
AP	$[(\mathsf{M}, \mathsf{w})]p \leftrightarrow (\mathsf{pre}(\mathsf{w}) \to p)$	(atomic permanence)

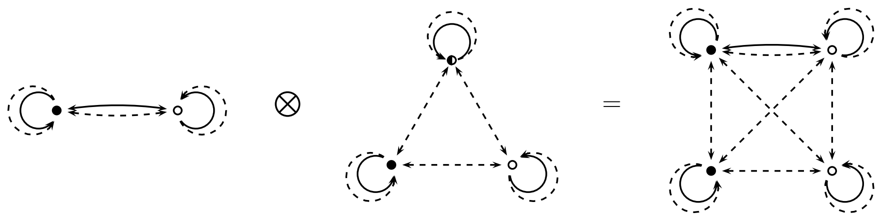


Figure 4.7: An illustration of how multiplication of a static model with an action model works in the approach of Baltag, Moss and Solecki. The model on the left is the static model. The model in the middle is the action model. The model on the right is the result of the multiplication of these two models. In the action model, the preconditions are as follows: the top world has precondition $\top$, the left world has precondition p, the right world has precondition $\neg p$. The solid nodes in the left and right model indicate that p is true and the open nodes in the left and right model indicate that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b.

PF	$[(\mathsf{M},\mathsf{w})]\neg\varphi \leftrightarrow (\mathsf{pre}(\mathsf{w}) \rightarrow \neg[(\mathsf{M},\mathsf{w})]\varphi)$	(partial functionality)
AK	$[(\mathsf{M},\mathsf{w})]\Box_a\varphi \leftrightarrow (\mathsf{pre}(\mathsf{w}) \rightarrow \bigwedge\{\Box_a[(\mathsf{M},\mathsf{v})]\varphi \mid \mathsf{wR}(a)\mathsf{v}\})$	(action-knowledge)
Nec	$\dfrac{\varphi}{[(\mathsf{M},\mathsf{w})]\varphi}$	(necessitation)

ACK (action-common-knowledge)

Let an action model (M,w) be given. Let χ_v be sentences for all v such that $\mathsf{wR}(\mathcal{B})^*\mathsf{v}$

$$\frac{\{\chi_\mathsf{v} \rightarrow [(\mathsf{M},\mathsf{v})]\varphi \mid \mathsf{wR}(\mathcal{B})^*\mathsf{v}\} \qquad \{(\chi_\mathsf{v} \wedge \mathsf{pre}(\mathsf{v})) \rightarrow \Box_a\chi_\mathsf{u} \mid \mathsf{wR}(\mathcal{B})^*\mathsf{v}, a \in \mathcal{B}, \mathsf{vR}(a)\mathsf{u}\}}{\chi_\mathsf{w} \rightarrow [(\mathsf{M},\mathsf{w})]C_\mathcal{B}\varphi}$$

This proof system is sound and complete with respect to the semantics given in definition 4.14. It is difficult to see that the action-common-knowledge rule is sound. To see this I will first show that it is sound in a very simple case. Let us take the very simple action where nothing really happens, i.e. the one-world model with precondition $\top$ accessible to all agents. The execution of this action does not change the static model. So if it can be deduced that a sentence is common knowledge after the execution, it must have been common knowledge already. So we should get an instantiation of the common knowledge induction rule. There is only one world, and its precondition is $\top$. So we get:

$$\frac{\chi \rightarrow \varphi \qquad \chi \rightarrow E_\mathcal{B}\chi}{\chi \rightarrow C_\mathcal{B}\varphi}$$

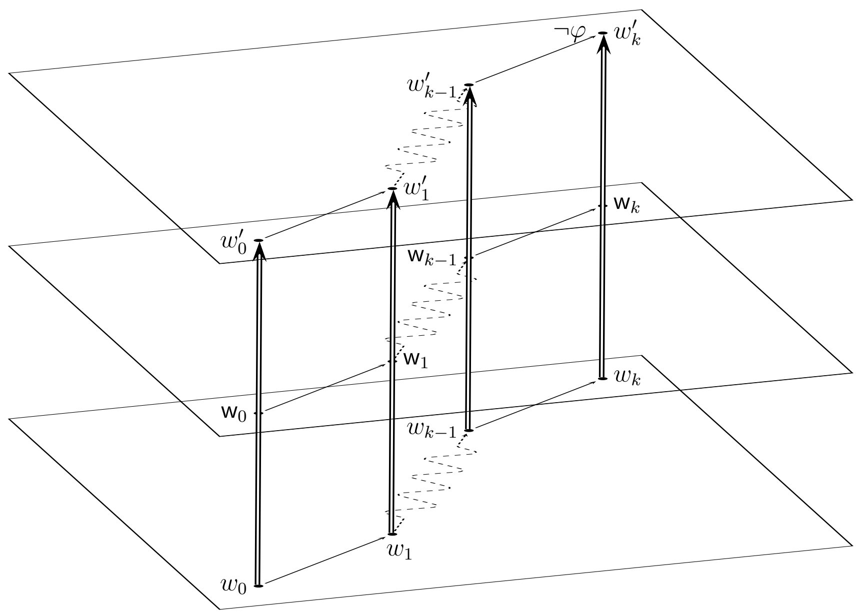


Figure 4.8: A sketch of the situation when a sentence φ is not common knowledge after action execution, by which one can see the soundness of the action-common-knowledge rule.

Clearly this is sound for ordinary common knowledge. To see the soundness of the whole rule suppose that the premises of the action-common-knowledge rule hold. Suppose moreover that φ is not common knowledge for group $\mathcal{B}$ after the execution of the action $(\mathsf{M}, \mathsf{w}_0)$, while χ_{w_0} is true. That means that in the updated model there is a path of worlds $w'_0, \ldots, w'_k$ linked by accessibility relations of members of group $\mathcal{B}$, such that $\neg\varphi$ holds in w'_k. That means that in the original model there is a sequence of worlds $w_0, \ldots, w_k$ and in the action model there are sequences of actions worlds $\mathsf{w}'_0, \ldots, \mathsf{w}'_k$ such that every action in the action sequence is executable in the corresponding world, i.e. $(M, w_i) \models \mathsf{pre}(\mathsf{w}_i)$. This situation is sketched in figure 4.8. It is clear that $\mathsf{pre}(\mathsf{w}_0)$ holds in w_0, therefore by the premises of the action-common-knowledge rule it also holds that in w_0 that $\Box_a \chi_{\mathsf{w}_1}$. We can continue this line of reasoning until we reach w_k. There χ_{w_k} must hold, and therefore by one of the premises $[(\mathsf{M}, \mathsf{w}_k)]\varphi$. Therefore it cannot be the case that $\neg\varphi$ holds in w'_k.

The main objection to this approach to dynamic epistemic logic is that the distinction between syntax and semantics is blurred. The action models are syntactic objects and semantic objects at the same time. I do not find this very elegant.

Suppose for example that you are interested in knowledge, for which the proof system $\mathsf{S5}_{\mathcal{PA}}$ is usually taken. The extra axioms of $\mathsf{S5}_{\mathcal{PA}}$ could also be added to the logic of epistemic actions. Given that we are now working with $\mathsf{S5}_{\mathcal{PA}}$, a natural constraint would be, that the kind of actions that you are interested in should also be $S5$ epistemic actions. That would mean that the non-$S5$ epistemic actions should be taken out of the language. To me this is not very appealing.

4.3.4 Changing modalities

A recent contribution by Renardel de Lavalette (2002) is based on the observation also made by Gerbrandy (see section 4.3.1) that in a Kripke model an element of an accessibility relation (an arrow) serves many purposes. Instead of using the non-well-founded possibilities, Renardel uses trees to overcome the difficulty of an arrow serving many purposes. For the use of these structures he was inspired by Fagin and Vardi (1985). The trees are called *lean modal structures* by Renardel.

Definition 4.16 (Lean modal structures)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. First lean modal structures for finite depth are given.

$$\begin{array}{lcl} MS_0 & = & \emptyset \\ MS_{n+1} & = & (\mathcal{P} \to \{0,1\}) \oplus (\mathcal{A} \to 2^{MS_n}) \end{array}$$

where $(X \to Y) \oplus (A \to B)$ is defined as $\{f \cup g \mid f \in (X \to Y) \text{ and } g \in (A \to B)\}$. The class of lean modal structures is the union of all these.

$$MS \;=\; \bigcup_n MS_n$$

A good graphical representation of a lean modal structure is the same as the graphical representation of Gerbrandy's possibilities. The main difference is that modal structures correspond to finite trees whereas possibilities correspond to infinite trees.

Definition 4.17 (Semantics for $\mathscr{L}_{\mathcal{PA}}$ using lean modal structures)
Let a lean modal structure f be given. Let $p \in \mathcal{P}$, $\varphi, \psi \in \mathscr{L}_{\mathcal{PA}}$, and $a \in \mathcal{A}$.

$$\begin{array}{lll} f \models p & \text{iff} & f(p) = 1 \\ f \models \neg\varphi & \text{iff} & f \not\models \varphi \\ f \models \varphi \wedge \psi & \text{iff} & f \models \varphi \text{ and } f \models \psi \\ f \models \Box_a\varphi & \text{iff} & g \models \varphi \text{ for all } g \text{ such that } g \in f(a) \end{array}$$

Note that the whole language $\mathscr{L}_{\mathcal{PA}}$ can be interpreted in such a structure, because if $f(a)$ is empty then any sentence of the form $\Box_a\varphi$ is true.

Just as in possibilities we can now clearly distinguish the different roles of the elements of accessibility relations. Renardel is interested in changing some modalities (for example $\Box_a$) while leaving other modalities intact (for example $\Box_b\Box_a$). This yields a general dynamic modal logic DML.

Definition 4.18 (Language of DML)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of changing modalities $\mathscr{L}^{:=}_{\mathcal{PA}}$ consists of a set of sentences φ and a set of actions α, given by the following rules

$$\begin{array}{lcl}\varphi & ::= & \bot \mid p \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid \Box_a\varphi \mid [\alpha]\varphi \\ \alpha & ::= & a \mid ?\varphi \mid p := \varphi \mid a := \alpha \mid (\alpha_1;\alpha_2) \mid (\alpha_1 \cup \alpha_2)\end{array}$$

where $p \in \mathcal{P}$ and $a \in \mathcal{A}$. □

It is easily seen that agents are special cases of actions. As far as the semantics is concerned we have the following addition to definition 4.17.

Definition 4.19 (Semantics for actions of $\mathscr{L}^{:=}_{\mathcal{PA}}$)
Let a lean modal structure f be given. Let $p \in \mathcal{P}$, let φ, ψ be sentences of $\mathscr{L}^{:=}_{\mathcal{PA}}$, and α be an action of $\mathscr{L}^{:=}_{\mathcal{PA}}$ and $a \in \mathcal{A}$.

$$f \models [\alpha]\varphi \quad \text{iff} \quad g \models \varphi \text{ for every } g \text{ such that } f[\![\alpha]\!]g$$

where

$$\begin{array}{lll} f[\![a]\!]g & \text{iff} & g \in f(a) \\ f[\![?\varphi]\!]g & \text{iff} & f \models \varphi \text{ and } f = g \\ f[\![p := \varphi]\!]g & \text{iff} & g = f[p \mapsto (f \models \varphi)] \\ f[\![a := \alpha]\!]g & \text{iff} & g = f[a \mapsto f[\![\alpha]\!]] \\ f[\![\alpha;\beta]\!]g & \text{iff} & \text{there is an } h \text{ such that } f[\![\alpha]\!]h[\![\beta]\!]g \\ f[\![\alpha \cup \beta]\!]g & \text{iff} & f[\![\alpha]\!]g \text{ or } f[\![\beta]\!]g \end{array}$$

where $f[p \mapsto (f \models \varphi)] = (f \setminus \{(p, f(p))\}) \cup \{(p,x) \mid (x = 0 \text{ and } f \not\models \varphi) \text{ or } (x = 1 \text{ and } f \models \varphi)\}$ and $f[a \mapsto f[\![\alpha]\!]] = (f \setminus \{(a, f(a))\}) \cup \{(a, f[\![\alpha]\!])\}$ and $f[\![\alpha]\!] = \{g \mid f[\![\alpha]\!]g\}$. □

Definition 4.20 (Proof system DML)
Let φ, ψ be sentences in $\mathscr{L}^{:=}_{\mathcal{PA}}$ and let α, β be actions in $\mathscr{L}^{:=}_{\mathcal{PA}}$. Let $a \in \mathcal{A}$ and $p \in \mathcal{P}$. The proof system DML consists of the following axioms and rules

Taut	all instantiations of propositional tautologies	
Distr	$[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$	
?**AX**	$[?\varphi]\psi \leftrightarrow (\varphi \to \psi)$	
; **AX**	$[\alpha;\beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$	
∪**AX**	$[\alpha \cup \beta]\varphi \leftrightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$	
SUB1	$[p := \varphi]p \leftrightarrow \varphi$	(shallow substitution)
AP1	$[p := \varphi]q \leftrightarrow q$	(atomic permanence)
F1	$[p := \varphi]\neg\psi \leftrightarrow \neg[p := \varphi]\psi$	(functionality)
I1	$[p := \varphi][a]\varphi \leftrightarrow [a]\varphi$	(independence)
SUB2	$[a := \alpha][a]\varphi \leftrightarrow [\alpha]\varphi$	(shallow substitution)

AP2	$[a := \alpha]p \leftrightarrow p$	(atomic permanence)
F2	$[a := \alpha]\neg\varphi \leftrightarrow \neg[a := \alpha]\varphi$	(functionality)
I2	$[a := \alpha][b]\varphi \leftrightarrow [b]\varphi$	(independence)
MP	$\dfrac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$	(modus ponens)
Nec	$\dfrac{\varphi}{[\alpha]\varphi}$	(necessitation)

This proof system is sound and complete with respect to the semantics provided in definition 4.17.

The main problem with shallow substitutions is that higher-order information can only be captured to a certain depth (i.e. the depth of the substitution). Higher-order information has an arbitrary depth. To overcome this problem Renardel extends his system with fixed points. This extension, μDML, can handle higher-order information change satisfactory. Adding common knowledge to μDML presumably makes it quite hard to axiomatize.

4.4 From action terms to action sentences

What is striking about all these dynamic epistemic logics is that there seems to be a tension between syntax and semantics. A clear syntax (clearly distinct from semantics), such as with the systems of Gerbrandy, Van Ditmarsch and Renardel, seems to lead to complicated semantics. And clear semantics, such as with Baltag, Moss and Solecki seems to lead to unclear syntax.

As was indicated earlier, one of the semantical problems has to do with the way Kripke models represent information. The accessibility relations represent information about the world but also higher-order information. To incorporate new information sometimes requires the addition of new worlds and also accessibility relations for those new worlds. The question is where those worlds come from. In Gerbrandy and Renardel's approaches this is solved essentially by viewing Kripke models as trees. In that way every role a possible world or an arrow plays in the interpretation of sentences can be distinguished. Then the execution of an action that affects certain information that certain agents have can focus on that part of the tree that models that information for those agents, while leaving the rest of the tree undisturbed. So the problem is solved by unraveling a Kripke model, and thus obtaining more worlds and more arrows.

The idea of the solution for this problem provided by Van Ditmarsch is rather elegant, although it makes the semantics rather cumbersome. If the agents in a group $\mathcal{B}$ learn that α takes place, the models that result from executing α are taken and those *models* are taken to be the *worlds* in the updated model. So new worlds can be created by having nondeterministic actions. The accessibility relation is then determined by the internal structure of the model-worlds.

The solution that Baltag et al. provide is that you take the worlds that are in the product of the worlds of the static model and the action model. In that way new worlds are created and the new accessibility relation is based on the accessibility relations in the static model and the action model. The main distinctive feature of this approach is that semantics is taken a the starting point, whereas a language was the starting point of the other approaches. The question is what language is appropriate for these semantics. The solution of Baltag et al. is that all the action models are inserted in the language.

Whether this is appropriate depends on the level at which you want to reason about actions. A logical language is usually developed to formalize inferences from a particular domain. Which features are considered important in that domain determine which features are in the language. If quantification is not considered important, there is no need to have quantifiers in the language. Or if arithmetic is the domain one is interested in, it seems that natural numbers must be expressible in the language. Modal logic is applied to many domains, such as knowledge, time, and programs. Yet temporal logic can be developed without reference to particular points in time. What about epistemic actions? If these actions are inserted into the language it means that every detail is considered important. In modal logic Kripke models are not inserted into the language. One cannot even refer to particular models in the language. This is not a desideratum. A sentence is true in a whole class of models. That makes the notion of validity a useful notion. If it is just the case that a certain inference is correct in one model, this is not very useful. If that inference is valid, i.e. correct in all models, then it is useful to know that it is. A sentence φ can only refer to classes of models (i.e. those based on frames that satisfy φ.) Why should we not take a similar point of view with epistemic actions? Since epistemic actions are Kripke models, why would we not treat them in the same way modal logic treats them?

4.4.1 Action language

We take this idea seriously in this section by studying a new logic ALL (action language logic). The first observation is that if epistemic actions can be construed as Kripke models, then surely, there is a modal language that can be interpreted in these models.

Definition 4.21 (Action language)
The action language $\mathscr{L}^{A}_{\mathcal{PA}}$ is given by the following rule:

$$\alpha ::= \bot \mid ?\varphi \mid \neg\alpha \mid (\alpha_1 \wedge \alpha_2) \mid \Box_a\alpha \mid C_{\mathcal{B}}\alpha$$

where $\varphi \in \mathscr{L}^{EC}_{\mathcal{PA}}$, $a \in \mathcal{A}$, and $\mathcal{B} \subseteq \mathcal{A}$. □

This idea is introduced in Baltag, Moss, and Solecki (1999), where this language is introduced as an auxiliary language. It is auxiliary in the sense that it can help to find the appropriate action model. It is just like the language of epistemic logic, but instead of a set of propositional variables, we now have a set of tests from $\mathscr{L}_{\mathcal{PA}}$, so it is very much like the language $\mathscr{L}_{\mathscr{L}_{\mathcal{PA}}\mathcal{A}}$. The language $\mathscr{L}_{\mathcal{PA}}$ can be seen as a parameter of the action language. Given another language that can be interpreted on Kripke models, we could let the basic action sentences be tests on sentences of that language.

Definition 4.22 (Semantics for $\mathscr{L}^{\mathcal{A}}_{\mathcal{PA}}$)
Let an action model $\mathsf{M} = (\mathsf{W}, \mathsf{R}, \mathsf{pre})$ for $\mathscr{L}_{\mathcal{PA}}$ and an action $\mathsf{w} \in \mathsf{W}$ be given.

$(\mathsf{M}, \mathsf{w}) \models_{act} ?\varphi$	iff	$\mathsf{pre}(\mathsf{w}) \models_{stat} \varphi$
$(\mathsf{M}, \mathsf{w}) \models_{act} \neg\alpha$	iff	$(\mathsf{M}, \mathsf{w}) \not\models_{act} \alpha$
$(\mathsf{M}, \mathsf{w}) \models_{act} (\alpha \wedge \beta)$	iff	$(\mathsf{M}, \mathsf{w}) \models_{act} \alpha$ and $(\mathsf{M}, \mathsf{w}) \models_{act} \beta$
$(\mathsf{M}, \mathsf{w}) \models_{act} \Box_a\alpha$	iff	$(\mathsf{M}, \mathsf{v}) \models_{act} \alpha$ for all v such that $\mathsf{w}\mathsf{R}(a)\mathsf{v}$
$(\mathsf{M}, \mathsf{w}) \models_{act} C_{\mathcal{B}}\alpha$	iff	$(\mathsf{M}, \mathsf{v}) \models_{act} \alpha$ for all v such that $\mathsf{w}\mathsf{R}(\mathcal{B})^+\mathsf{v}$

where in the second clause $\mathsf{pre}(\mathsf{w}) \models \varphi$ is interpreted as local logical consequence (see definition 2.4). □

The question is how such a language could be used to reason about actions. We are interested in the effect that actions have. We now have a way of describing a whole class of actions with just one action sentence. So we can now ask ourselves what the effect is of executing an action of type α. That means that we could somehow view the sentences of ALL as dynamic operators, thus we will look at sentences of the form $[\alpha]\varphi$[2]. So we get the following definition.

Definition 4.23 (Language of ALL)
The dynamic language $\mathscr{L}^{D}_{\mathcal{PA}}$ is given by the following rule in BNF:

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid \Box_a\varphi \mid C_{\mathcal{B}}\varphi \mid [\alpha]\varphi$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$, $\mathcal{B} \subseteq \mathcal{A}$, and $\alpha \in \mathscr{L}^{A}_{\mathcal{PA}}$. □

A sentence of the form $[\alpha]\varphi$ is to read as 'φ after an action that satisfies α is executed.' The language is interpreted according to the following definition.

Definition 4.24 (Semantics for $\mathscr{L}^{D}_{\mathcal{PA}}$)
Let a static model $M = (W, R, V)$, a state $w \in W$, an action model $\mathsf{M} = (\mathsf{W}, \mathsf{R}, \mathsf{pre})$ and an action $\mathsf{w} \in \mathsf{W}$ be given.

$(M, w) \not\models_{stat} \bot$		
$(M, w) \models_{stat} p$	iff	$w \in V(p)$
$(M, w) \models_{stat} \neg\varphi$	iff	$(M, w) \not\models_{stat} \varphi$
$(M, w) \models_{stat} (\varphi \wedge \psi)$	iff	$(M, w) \models_{stat} \varphi$ and $(M, w) \models_{stat} \psi$

[2]First I considered a two-sorted language where tests on the static sentences were allowedd, but it proved very difficult to give good semantics for this language. One could consider a sequence of static and action languages, where the nesting of test operators and dynamic operators is allowed to get deeper and deeper. But I do not do so here.

$$
\begin{array}{lll}
(M,w) \models_{stat} \Box_a \varphi & \text{iff} & (M,v) \models_{stat} \varphi \text{ for all } v \text{ such that } wR(a)v \\
(M,w) \models_{stat} C_{\mathcal{B}} \varphi & \text{iff} & (M,v) \models_{stat} \varphi \text{ for all } v \text{ such that } wR(\mathcal{B})^+ v \\
(M,w) \models_{stat} [\alpha]\varphi & \text{iff} & (M \cdot \mathsf{M}', (w,\mathsf{w}')) \models_{stat} \varphi \text{ for all } (\mathsf{M}',\mathsf{w}') \text{ such that} \\
 & & (M,w) \models \mathsf{pre}(\mathsf{w}') \text{ and } (\mathsf{M}',\mathsf{w}') \models_{act} \alpha
\end{array}
$$

Where $\mathsf{pre}(\mathsf{w}) \models_{stat} \varphi$ is again interpreted as local logical consequence. □

The following theorem about bisimilarity holds. For the definition of bisimulation see definition 2.8 on page 8.

Theorem 4.1 (Preservation of bisimulation)
Let two static models (M,w) and (M',w'), and an action model (M,w) be given. If $(M,w) \leftrightarrows (M',w')$, then for every sentence $\alpha \in \mathscr{L}^A_{\mathcal{PA}}$ it holds that $(M,w) \models \mathsf{pre}(\mathsf{w})$ implies that $(M \cdot \mathsf{M}, (w,\mathsf{w})) \leftrightarrows (M' \cdot \mathsf{M}, (w',\mathsf{w}))$. □

Proof Suppose that $(M,w) \models \mathsf{pre}(\mathsf{w})$. From theorem 2.2 it follows that $(M',w') \models \mathsf{pre}(\mathsf{w})$, therefore both $(M \cdot \mathsf{M}, (w,\mathsf{w}))$ and $(M' \cdot \mathsf{M}, (w',\mathsf{w}))$ exist. Now we have to prove that $(M \cdot \mathsf{M}, (w,\mathsf{w})) \leftrightarrows (M' \cdot \mathsf{M}, (w',\mathsf{w}))$.

Let $\mathfrak{R}$ be a bisimulation that establishes $(M,w) \leftrightarrows (M',w')$. Now define

$$(v,\mathsf{v})\mathfrak{R}'(v',\mathsf{v}) \text{ iff } v\mathfrak{R}v'$$

We have to show that $\mathfrak{R}'$ is a bisimulation. Suppose $(v,\mathsf{v})\mathfrak{R}'(v',\mathsf{v})$, therefore $v\mathfrak{R}v'$

atoms trivial

forth Suppose $(v,\mathsf{v})R \cdot \mathsf{R}(a)(u,\mathsf{u})$, where $R \cdot \mathsf{R}(a)$ is the accessibility relation assigned to a in $(M \cdot \mathsf{M})$. Therefore $vR(a)u$ and $\mathsf{v}\mathsf{R}(a)\mathsf{u}$ and $(M,u) \models \mathsf{pre}(\mathsf{u})$. From the definition of bisimulation it follows that there is a world u' such that $v'R'(a)u'$ and $u\mathfrak{R}u'$. From $(M,u) \models \mathsf{pre}(\mathsf{u})$ and theorem 2.2 it follows that $(M',u') \models \mathsf{pre}(\mathsf{u})$. Therefore (u',u) exists in $(M' \cdot \mathsf{M}, (w',\mathsf{w}))$. And therefore $(v',\mathsf{v})R' \cdot \mathsf{R}(a)(u',\mathsf{u})$ and $(u,\mathsf{u})\mathfrak{R}'(u',\mathsf{u})$.

back Analogous to forth.

Therefore $\mathfrak{R}'$ establishes that $(M \cdot \mathsf{M}, (w,\mathsf{w})) \leftrightarrows (M' \cdot \mathsf{M}, (w',\mathsf{w}))$. □

Theorem 4.2
Let two static models (M,w) and (M',w'), be given. If $(M,w) \leftrightarrows (M',w')$, then for every sentence $\varphi \in \mathscr{L}^D_{\mathcal{PA}}$ it holds that $(M,w) \models \varphi$ iff $(M',w') \models \varphi$ □

Proof By induction on φ. The cases for when φ is a propositional variable, negation, conjunction, individual epistemic operator, or a common knowledge operator, are fairly straightforward.

Suppose φ is of the form $[\alpha]\psi$ and $(M,w) \models [\alpha]\psi$. Take an arbitrary action model (M,w) such that $(\mathsf{M},\mathsf{w}) \models \alpha$ and $(M,w) \not\models \mathsf{pre}(\mathsf{w})$. If such an action model does not exist, then neither does such a model exist where $(M',w') \models \mathsf{pre}(\mathsf{w})$

(by theorem 2.2). In that case we would be done. Otherwise it follows from theorem 4.1 that $(M \cdot \mathsf{M}, (w, \mathsf{w})) \leftrightarrows (M' \cdot \mathsf{M}, (w', \mathsf{w}))$. From the semantics and our assumptions it follows that $(M \cdot \mathsf{M}, (w, \mathsf{w})) \models \psi$, and therefore $(M' \cdot \mathsf{M}, (w', \mathsf{w})) \models \psi$ as well. Therefore $(M', w') \models [\alpha]\psi$. □

Now we can attempt to construct a proof system for ALL. The following are two lists of obvious validities. The question is whether these lists are enough to ensure completeness.

Definition 4.25 (Static axioms and rules of ALL)
The static axioms and rules of ALL consist of all axioms and rules of $\mathsf{K}^{EC}_{\mathcal{PA}}$ and the following axioms and rule:

Distr$_{[\alpha]}$	$[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$	($[\alpha]$-distribution)
EX	$\varphi \leftrightarrow [?\neg\varphi]\bot$	(executability)
AF	$\langle\alpha\rangle p \to [\alpha]p$	(atomic functionality)
AP	$[\alpha]p \leftrightarrow (\langle\alpha\rangle\top \to p)$	(atomic permanence)
C	$([\alpha]\varphi \wedge [\beta]\varphi) \to [\alpha \vee \beta]\varphi$	(choice)
KA	$\Box_a[\alpha]\varphi \leftrightarrow [\Box_a\alpha]\Box_a\varphi$	(knowledge-action)
CKA	$C_\mathcal{B}[\alpha]\varphi \leftrightarrow [C_\mathcal{B}\alpha]C_\mathcal{B}\varphi$	(common-knowledge-action)
Nec$_{[\alpha]}$	$\dfrac{\varphi}{[\alpha]\varphi}$	($[\alpha]$-necessitation)
AS	$\dfrac{\alpha \to \beta}{[\beta]\varphi \to [\alpha]\varphi}$	(action-static rule)

Definition 4.26 (Action axioms and rules of ALL)
The action axioms and rules of ALL consists of all axioms and rules of $\mathsf{K}^{EC}_{\mathcal{PA}}$ and the following axiom and rule:

Distr$_?$	$?(\varphi \to \psi) \to (?\varphi \to ?\psi)$	(?-distribution)
Nec$_?$	$\dfrac{\varphi}{?\varphi}$	(?-necessitation)

A derivation in this system consists of a sequence of sentences of $\mathscr{L}^A_{\mathcal{PA}}$ and $\mathscr{L}^D_{\mathcal{PA}}$ each of which is an instance of an axiom or is the result of applying a derivation rule to sentences that occur earlier in the sequence. Moreover static axioms and rules are only applied to sentences of $\mathscr{L}^D_{\mathcal{PA}}$ and action axioms and rules are only applies to sentences of $\mathscr{L}^A_{\mathcal{PA}}$. If φ is the last sentence in a derivation, then φ is provable, or deducible, notation $\vdash_{stat} \varphi$. If α is the last sentence in a derivation, then α is provable, or deducible, notation $\vdash_{act} \alpha$.

Here is a proof of the soundness of the Knowledge-Action axiom.
$(M, w) \not\models \Box_a[\alpha]\varphi$

$\equiv$ {Definition 4.24 }
$\exists v(wR(a)v$ and $(M, v) \not\models [\alpha]\varphi)$
$\equiv$ {Definition 4.24 }
$\exists v(wR(a)v$ and $\exists(\mathsf{M}_1, \mathsf{v})((M, v) \models \mathsf{pre}(\mathsf{v})$ and $(\mathsf{M}_1, \mathsf{v}) \models \alpha$ and $(M \cdot \mathsf{M}_1, (v, \mathsf{v})) \not\models \varphi))$
$\equiv$ {
$\Rightarrow$ Construct a model $(\mathsf{M}_2, \mathsf{w})$ by adding a fresh world w to M_1 such that v is the only world accessible from w to a and $\mathsf{pre}(\mathsf{w}) = \top$
$\Leftarrow$ Let $(\mathsf{M}_1, \mathsf{v})$ be $(\mathsf{M}_2, \mathsf{v})$.
}
$\equiv$
$\exists v(wR(a)v$ and $\exists(\mathsf{M}_2, \mathsf{w})((M, w) \models \mathsf{pre}(\mathsf{w})$ and $\forall \mathsf{v}(\mathsf{wR}(a)\mathsf{v} \rightarrow (\mathsf{M}_2, \mathsf{v}) \models \alpha)$ and $\exists \mathsf{v}(\mathsf{wR}(a)\mathsf{v}$ and $(M, v) \models \mathsf{pre}(\mathsf{v})$ and $(M \cdot \mathsf{M}_2, (v, \mathsf{v})) \not\models \varphi)$
$\equiv$ { Predicate Logic }
$\exists(\mathsf{M}_2, \mathsf{w})((M, w) \models \mathsf{pre}(\mathsf{w})$ and $\forall \mathsf{v}(\mathsf{wR}(a)\mathsf{v} \rightarrow (\mathsf{M}_2, \mathsf{v}) \models \alpha)$ and $\exists v \exists \mathsf{v}(wR(a)v$ and $\mathsf{wR}(a)\mathsf{v}$ and $(M, v) \models \mathsf{pre}(\mathsf{v})$ and $(M \cdot \mathsf{M}_2, (v, \mathsf{v})) \not\models \varphi)$
$\equiv$ { Definition 4.24 }
$\exists(\mathsf{M}_2, \mathsf{w})((M, w) \models \mathsf{pre}(\mathsf{w})$ and $\forall \mathsf{v}(\mathsf{wR}(a)\mathsf{v} \rightarrow (\mathsf{M}_2, \mathsf{v}) \models \alpha)$ and $(M \cdot \mathsf{M}_2, (w, \mathsf{w})) \not\models \Box_a \varphi)$
$\equiv$ { Definition 4.24 }
$\exists(\mathsf{M}_2, \mathsf{w})((M, w) \models \mathsf{pre}(\mathsf{w})$ and $(\mathsf{M}_2, \mathsf{w}) \models \Box_a \alpha$ and $(M \cdot \mathsf{M}_2, (w, \mathsf{w})) \not\models \Box_a \varphi)$
$\equiv$ { Definition 4.24 }
$(M, w) \not\models [\Box_a \alpha]\Box_a \varphi$

From these lists of axioms we can deduce some very nice other formulas and rules.

Deduced theorems

- $[\alpha \vee \beta]\varphi \rightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$.

  1. $\alpha \rightarrow \alpha \vee \beta$ **Taut**
  2. $\beta \rightarrow \alpha \vee \beta$ **Taut**
  3. $[\alpha \vee \beta]\varphi \rightarrow [\alpha]\varphi$ **AS**
  4. $[\alpha \vee \beta]\varphi \rightarrow [\beta]\varphi$ **AS**
  5. $[\alpha \vee \beta]\varphi \rightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$ Propositional logic 3,4

- for all propositional φ it holds that $\vdash_{stat} [\alpha]\varphi \leftrightarrow (\langle\alpha\rangle\top \rightarrow \varphi)$ and $\vdash_{stat} \langle\alpha\rangle\varphi \rightarrow [\alpha]\varphi$

  Proof By induction on φ

  base For atoms we have the axiom of propositional functionality and the axiom of atomic permanence.

induction hypothesis Suppose $\vdash_{stat} [\alpha]\varphi \leftrightarrow (\langle\alpha\rangle\top \to \varphi)$ and $\vdash_{stat} \langle\alpha\rangle\varphi \to [\alpha]\varphi$ and $\vdash_{stat} [\alpha]\psi \leftrightarrow (\langle\alpha\rangle\top \to \psi)$ and $\vdash_{stat} \langle\alpha\rangle\psi \to [\alpha]\psi$

induction step negation: functionality and permanence

1. $\langle\alpha\rangle\varphi \to [\alpha]\varphi$ Induction hypothesis
2. $\neg[\alpha]\varphi \to \neg\langle\alpha\rangle\varphi$ Contraposition 1
3. $\langle\alpha\rangle\neg\varphi \to [\alpha]\neg\varphi$ def. $[\alpha] \leftrightarrow \neg\langle\alpha\rangle\neg$
4. $[\alpha]\varphi \leftrightarrow (\langle\alpha\rangle\top \to \varphi)$ Induction hypothesis
5. $\neg\varphi \to (\varphi \to \bot)$ Propositional tautology
6. $[\alpha]\neg\varphi \to ([\alpha]\varphi \to [\alpha]\bot)$ $[\alpha]$-necessitation and -distribution
7. $[\alpha]\neg\varphi \to ((\langle\alpha\rangle\top \to \varphi) \to [\alpha]\bot)$ Replacement 4,6
8. $[\alpha]\neg\varphi \to ((\neg\varphi \to [\alpha]\bot) \to [\alpha]\bot)$ Contraposition
9. $[\alpha]\neg\varphi \to (\langle\alpha\rangle\top \to \neg\varphi)$ Propositional logic
10. $[\alpha]\varphi \to (\langle\alpha\rangle\top \to \varphi)$ Propositional logic 4
11. $\neg(\langle\alpha\rangle\top \to \varphi) \to \neg[\alpha]\varphi$ Contraposition 10
12. $\neg(\langle\alpha\rangle\top \to \varphi) \to [\alpha]\neg\varphi$ Replacement 1, 11
13. $\langle\alpha\rangle\top \vee [\alpha]\bot$ Propositional tautology
14. $\bot \to \neg\varphi$ Propositional tautology
15. $[\alpha]\bot \to [\alpha]\neg\varphi$ $[\alpha]$-necessitation and -distribution
16. $\langle\alpha\rangle\top \vee [\alpha]\neg\varphi$ Propositional logic 13, 15
17. $(\langle\alpha\rangle\top \to \neg\varphi) \to [\alpha]\neg\varphi$ Propositional logic 12, 16
18. $[\alpha]\neg\varphi \leftrightarrow (\langle\alpha\rangle\top \to \neg\varphi)$ Propositional logic 10, 17

conjunction: functionality

1. $\langle\alpha\rangle\varphi \to [\alpha]\varphi$ Induction hypothesis
2. $\langle\alpha\rangle\psi \to [\alpha]\psi$ Induction hypothesis
3. $\langle\alpha\rangle(\varphi \wedge \psi) \to \langle\alpha\rangle\varphi \wedge \langle\alpha\rangle\psi$ Modal logic
4. $\langle\alpha\rangle\varphi \wedge \langle\alpha\rangle\psi \to [\alpha]\varphi \wedge [\alpha]\psi$ Propositional logic 1,2
5. $[\alpha]\varphi \wedge [\alpha]\psi \to [\alpha](\varphi \wedge \psi)$ Modal logic
6. $\langle\alpha\rangle(\varphi \wedge \psi) \to [\alpha](\varphi \wedge \psi)$ Propositional logic 3,4,5

conjunction: permanence

1. $[\alpha]\varphi \leftrightarrow (\langle\alpha\rangle\top \to \varphi)$ Induction hypothesis
2. $[\alpha]\psi \leftrightarrow (\langle\alpha\rangle\top \to \psi)$ Induction hypothesis
3. $[\alpha]\varphi \wedge [\alpha]\psi \leftrightarrow (\langle\alpha\rangle\top \to \varphi) \wedge (\langle\alpha\rangle\top \to \psi)$ Propositional logic 1,2
4. $[\alpha]\varphi \wedge [\alpha]\psi \leftrightarrow [\alpha](\varphi \wedge \psi)$ Modal logic
5. $[\alpha](\varphi \wedge \psi) \leftrightarrow (\langle\alpha\rangle\top \to \varphi) \wedge (\langle\alpha\rangle\top \to \psi)$ Propositional logic 3,4
6. $((\langle\alpha\rangle\top \to \varphi) \wedge (\langle\alpha\rangle\top \to \psi)) \leftrightarrow \langle\alpha\rangle\top \to (\varphi \wedge \psi)$ Propositional tautology

7. $[\alpha](\varphi \wedge \psi) \leftrightarrow (\langle\alpha\rangle\top \rightarrow (\varphi \wedge \psi))$ Propositional logic 5,6

□

- for all propositional sentences φ it holds that $\vdash_{stat} [?\varphi]\varphi$

1. $\varphi \rightarrow (\langle?\varphi\rangle\top \rightarrow \varphi)$	**Taut**
2. $[?\varphi]\varphi \leftrightarrow (\langle?\varphi\rangle\top \rightarrow \varphi)$	see deduced theorem above
3. $\varphi \rightarrow [?\varphi]\varphi$	Replacement 1,2
4. $\neg\varphi \leftrightarrow [?\varphi]\bot$	**EX**
5. $[?\varphi]\bot \rightarrow [?\varphi]\varphi$	Modal logic
6. $\neg\varphi \rightarrow [?\varphi]\varphi$	Propositional logic 4,5
7. $[?\varphi]\varphi$	Propositional logic 3,6

Completeness for $\vdash_{act}$

The completeness proof is based on Dunin-Kęplicz and Verbrugge (2002), which in turn is based on Fagin, Halpern, Moses, and Vardi (1995).

Definition 4.27 (Action closure)
The *action closure* of α is the minimal set $\Pi \subseteq \mathscr{L}^D_{\mathcal{PA}}$ such that

1. $\alpha \in \Pi$.
2. If $\beta \in \Pi$ and γ is a subaction of β (outside the scope of a test operator), then $\gamma \in \Pi$.
3. If $\beta \in \Pi$ and β itself is not a negation, then $\neg\beta \in \Pi$
4. If $C_{\mathcal{B}}\beta \in \Pi$, then $\Box_a C_{\mathcal{B}}\beta \in \Pi$ for all $a \in \mathcal{B}$ □

Note that the closure of any action α yields a finite set of sentences Π.

Definition 4.28 (Maximal consistent in Π)
Let Π be an action closure. A finite set of sentences Γ such that $\Gamma \subseteq \Pi$ is maximal consistent in Π iff:

1. Γ is consistent, i.e. $\nvdash_{act} \neg(\bigwedge_{\beta\in\Gamma} \beta)$.
2. There is no $\Gamma' \subseteq \Pi$, such that $\Gamma \subset \Gamma'$ and Γ' is consistent. □

Lemma 4.1 (Lindenbaümchen)
Let Π be the closure of a consistent β. If $\Gamma \subseteq \Pi$ is consistent, then there is a set $\Gamma' \subseteq \Pi$ such that $\Gamma \subseteq \Gamma'$ and Γ' is maximal consistent. □

Proof As Π is finite, the members of Π can be enumerated. Let us suppose $\#(\Pi) = k$ and that α_i $(1 \leq i \leq k)$is the i-th action sentence of this enumeration. Now define Γ_i $(0 \leq i \leq k)$ as follows:

$$\Gamma_0 = \Gamma$$
$$\Gamma_{i+1} = \begin{cases} \Gamma_i & \text{if } \Gamma_i \cup \{\alpha_{i+1}\} \text{ is inconsistent} \\ \Gamma_i \cup \{\alpha_{i+1}\} & \text{otherwise} \end{cases}$$

It is easily seen that Γ_k is maximal consistent in α. □

Definition 4.29 (Countermodel)
Let an action sentence α be given the countermodel for this sentence $\mathsf{M}_\alpha = (\mathsf{W}_\alpha, \mathsf{R}_\alpha, \mathsf{pre}_\alpha)$ is given by

- $\mathsf{W}_\alpha = \{\Gamma \subseteq \Pi \mid \Gamma \text{ is maximal consistent}\}$
- $\Gamma \mathsf{R}_\alpha(a)\Delta$ iff $\beta \in \Delta$ for all β such that $\Box_a\beta \in \Gamma$.
- $\mathsf{pre}_\alpha(\Gamma) = \bigwedge_{?\varphi\in\Gamma} \varphi$ □

Lemma 4.2 (Finite Valuation)
Let Γ be a maximal consistent set in Π.

1. If $\neg\beta \in \Pi$, then $\neg\beta \in \Gamma$ iff $\beta \notin \Gamma$.
2. If $\beta \wedge \gamma \in \Pi$, then $\beta \wedge \gamma \in \Gamma$ iff $\beta \in \Gamma$ and $\gamma \in \Gamma$.
3. If $\Box_a\beta \in \Pi$, then $\Box_a\beta \in \Gamma$ iff $\beta \in \Delta$ for all Δ with $\Gamma\mathsf{R}(a)\Delta$
4. If $C_{\mathcal{B}}\beta \in \Pi$, then $C_{\mathcal{B}}\beta \in \Gamma$ iff $\beta \in \Delta$ for all Δ such that $\Gamma\mathsf{R}(\mathcal{B})^+\Delta$ □

Proof see Dunin-Kęplicz and Verbrugge (2002). □

Lemma 4.3 (Finite Truth)
If $\Gamma \in W_\alpha$, then for all $\beta \in \Pi$ it holds that $(\mathsf{M}_\alpha, \Gamma) \models_{act} \beta$ iff $\beta \in \Gamma$. □

Proof By induction on β. Suppose β is of the form $?\varphi$ and $?\varphi \in \Pi$. If $(\mathsf{M}_\alpha, \Gamma) \models_{act} ?\varphi$, then $\mathsf{pre}_\alpha(\Gamma) \models_{stat} \varphi$. Both $\mathsf{pre}_\alpha(\Gamma)$, and φ are in $\mathscr{L}^{EC}_{\mathcal{PA}}$. Therefore by completeness of $\mathsf{K}^{EC}_{\mathcal{PA}}$ and $\mathsf{pre}_\alpha(\Gamma) \vdash_{stat} \varphi$. Since $\Gamma \vdash_{stat} ?\mathsf{pre}_\alpha(\Gamma)$, by repeated **Distr**$_?$ and **MP** we get $?\varphi \in \Gamma$. It also holds the other way around. For the other cases see Dunin-Kęplicz and Verbrugge (2002) □

Theorem 4.3 (Completeness)
If $\models_{act} \alpha$, then $\vdash_{act} \alpha$. □

Proof Suppose $\nvdash_{act} \alpha$, therefore i.e. $\neg\alpha$ is consistent, therefore there is a maximal consistent set Γ in the action closure Π of $\neg\alpha$ such that $\neg\alpha \in \Gamma$. Because of the finite truth lemma we may conclude that $(\mathsf{M}_{\neg\alpha}, \Gamma) \models_{act} \neg\alpha$, and therefore $(\mathsf{M}_{\neg\alpha}, \Gamma) \not\models_{act} \alpha$ □

Lecture or Amsterdam

We can already deal with some of the settings of the lecture or Amsterdam problem (see page 38). Suppose Anne opens the letter and reads the contents. Then afterwards Anne will know the contents. This can be formalized with ALL. Let a be Anne. The action of reading the letter will lead to either Anne learning that p or learning that $\neg p$. If Anne learns that p, this is an action that must satisfy $\Box_a ?p$. That is, she believes that a precondition of the action is that p holds. Therefore we would like to prove:

$$[\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$$

We can prove this in the following way

1. $[?p]p$ — see deduced theorems
2. $\Box_a [?p]p$ — $\Box_a$-necessitation 1
3. $\Box_a [?p]p \rightarrow [\Box_a ?p]\Box_a p$ — Knowledge-Action
4. $[\Box_a ?p]\Box_a p$ — Modus Ponens 2,3
5. $[\Box_a ?p](\Box_a p \vee \Box_a \neg p)$ — Modal Logic 4
6. $[\Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$ — Analogous to 1 – 5
7. $([\Box_a ?p](\Box_a p \vee \Box_a \neg p) \wedge [\Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)) \rightarrow [\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$ Choice
8. $[\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$ — Propositional Logic 5,6,7

Suppose now that Bill is also present at the table, and sees Anne opening the letter and reading it. Afterwards he knows that Anne knows what the letter said. Let b be Bill. Now we want to prove the following:

$$[\Box_b(\Box_a ?p \vee \Box_a ?\neg p)]\Box_b(\Box_a p \vee \Box_a \neg p)$$

This is quite simple given the previous result:

1. $[\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$ — Previous result
2. $\Box_b[\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p)$ — $\Box_b$-necessitation 1
3. $\Box_b[\Box_a ?p \vee \Box_a ?\neg p](\Box_a p \vee \Box_a \neg p) \leftrightarrow [\Box_b(\Box_a ?p \vee \Box_a ?\neg p)]\Box_b(\Box_a p \vee \Box_a \neg p)$ Knowledge-Action
4. $[\Box_b(\Box_a ?p \vee \Box_a ?\neg p)]\Box_b(\Box_a p \vee \Box_a \neg p)$ — Modus Ponens 2,3

Knowledge actions and public announcements

There are various subclasses of epistemic actions that are worth studying in their own right. For example, public announcements and knowledge actions, i.e. actions that preserve the $\mathsf{S5}$ axioms. One of the nice features of ALL is that we can try to characterize these simply by adding axioms, which would constrain the class of action frames at our disposal. Since completeness was proved we can do this by well-known modal techniques. The $\mathsf{S5}$ axioms can be added to get the class of knowledge actions for example. These axioms ensure that the accessibility relation in the action models are equivalence relations.

Public announcements can be seen as *one-world* models, where that world is accessible to all agents. The axiom **Triv** added to K is sound and complete with respect to the class frames where each world is accessible to itself and tot itself alone, i.e. $\forall w, v(wRv \leftrightarrow w = v)$. See Hughes and Cresswell (1996, pp.121,122).

Triv $\Box_a\alpha \leftrightarrow \alpha$

Each generated subframe of a frame that validates this axiom is a one-world frame. In that case for example, the following axiom is valid by the knowledge action axiom and the action static rule.

Perfect recall $\Box_a[\alpha]\varphi \rightarrow [\alpha]\Box_a\varphi$

The idea is that agents do not forget any information they had. See van Benthem (2001).

Contrary to Baltag, Moss, and Solecki (1999) the language does not need to be changed if we are interested in limited classes of epistemic actions. There is a big problem however. Although we have completeness for these logics, it seems very hard to provide a complete proof system for the whole system.

Incompleteness

There are serious problems with constructing a complete proof system for ALL. The question is whether it can be done at all. Let us look at the notion of completeness in general. Suppose we have a class of models $KS \subseteq K_{\mathcal{PA}}$. Let Λ be the set of all sentences in a certain language $\mathscr{L}$ that are true in every pointed model in KS, i.e. $\Lambda = \{\varphi \mid (M, w) \models \varphi$ for every $M \in KS\}$. A proof system S is said to be complete with respect to KS iff for every $\varphi \in \mathscr{L}$ it holds that $\varphi \in \Lambda$ implies that φ is provable in S. Note that we have the following in general: if $KS' \subseteq KS$, then $\Lambda \subseteq \Lambda'$. So the less models are allowed by the semantics, the more validities there are. But not only that: if a sentence was valid in all the models in the larger set, then certainly it is valid in all the models in the smaller set. If one wants to construct a complete proof system for an intended class of models that is a subset of a class of models for which there is a complete proof system, one adds axioms and rules to this system.

It seems this general picture does not apply to ALL. In this case there are two classes of models, static models $K_{\mathcal{PA}}$ and action models $\mathsf{K}_{\mathcal{PA}}$. Let us look at the following sentence:

$$\varphi = \langle \top \rangle \Box_a \bot$$

We can establish that this sentence is true in all static models in $K_{\mathcal{PA}}$ given the class of action models $\mathsf{K}_{\mathcal{PA}}$. Think of the one world action model where all accessibility relations are empty and the precondition of the only world is $\top$. This action model can be executed in any static model. The result will be a static model with empty accessibility relations for all agents. Therefore the sentence above is true in all models. Consequently this sentence should be deducible in a proof system for ALL.

Now suppose we want to limit actions to the class $\mathsf{S5}_{\mathcal{PA}}$ and static models to the class $S5_{\mathcal{PA}}$. Now the action described above is no longer available. So the sentence φ is not true in any models in $S5_{\mathcal{PA}}$ given the class $\mathsf{S5}_{\mathcal{PA}}$. This makes it very difficult to axiomatize this logic, if not impossible.

A similar problem arises when we limit actions to public announcements. In this case, the knowledge action axiom is no longer valid from right to left. The soundness proof depended essentially on adding a world to the action model. This is no longer allowed with public announcements.

The main problem seems to be that the validity of sentences of the form $[\alpha]\varphi$ depends on whether there exist models that satisfy α. It may well be that ALL is undecidable. One may try to prove this by trying to reduce a tiling problem to the question of satisfiability, or find an appropriate undecidable fragment of second order logic which can be faithfully translated to the language of ALL. See Blackburn, de Rijke, and Venema (2001, chapter 6).

4.4.2 Dyadic hybrid epistemic logic

The idea of treating states and actions in the same way is taken even further in ten Cate (2002). In his paper ten Cate puts static models and action models together in bigger models. This means that it is no longer the case that all action models are available, only the ones that occur in the bigger models. In most of the previous dynamic epistemic logics, executing an action meant that a new model was constructed which resulted from executing the action. In this case however we have both the static and the action model in the same model, and executing an action can be seen as a ternary relation between static worlds, action worlds, and static worlds. In that sense we have left dynamic logic and entered a branch of logics known as multi-dimensional logics, in this case a two dimensional logic, which has dyadic modal operators. The idea is to generalize usual modal operators. The modal operator $\Box_a$ for example gets its interpretation due to a binary relation defined on the set of worlds. By generalizing this one gets an n-ary modality for every $(n+1)$-ary relation on the set of possible worlds. Arrow logic, see van Benthem (1994), is an example of such a logic. See Blackburn, de Rijke, and Venema (2001) for an introduction to these kinds of modalities.

As we saw in the previous section much depends on the level at which one wants to describe actions. On the one hand one wants to be able to make assertions about actions of certain types, on the other hand one sometimes wants to make assertions about very concrete actions. The other feature of ten Cate's approach is to add the technical machinery of hybrid logic to the system. Hybrid logic has advanced considerably over the last years. Hybrid logics are extensions of modal logic, where a few key items are added to the language. The first thing that is added is a set of *nominals*, I. These are propositional variables $\mathsf{i}, \mathsf{j}, \mathsf{k}, \ldots$ whose extensions are singleton sets. So one could say that they name worlds. A world can have many names, or no name at all. Those names can be used just as individual variables can be used in first order logic. So a sentence $\varphi(\mathsf{i})$ could be interpreted as saying that i is a φ-world. But then we would have to interpret sentences on the level of models rather than on the level of pointed models. To keep things modal, and 'local', @-modalities are introduced. A sentence of the form $@_\mathsf{i}\varphi$ means that φ is true in the world called i, or 'at i φ holds'. More operators can be added such as $\downarrow$, where a sentence of the form $\downarrow \mathsf{i}\ \varphi$, means that φ is true if you call the actual world i. This operator will not be used here. The nice thing about hybrid logic is that one gains much of the expressivity of first-order logic, while preserving many of the nice computational properties of modal logic. Irreflexivity can for example be expressed as the axiom scheme $@_\mathsf{i}\Box_a\neg\mathsf{i}$. This cannot be expressed by any scheme in epistemic logic (see van Benthem (1984)). Moreover it has a nice axiomatization that can easily be extended for restricted classes of frames. But let us start with the language of Ten Cate's system.

Definition 4.30 (Language of DHEL)
Let a countable set of propositional variables $\mathcal{P}$ with a subset I of nominals, and a finite set of agents $\mathcal{A}$ be given. The language of DHEL $\mathscr{L}^{\rhd @_\mathsf{i} \mathsf{A}}_{\mathcal{P}\mathsf{I}\mathcal{A}}$ is given by the following rule in BNF:

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a\varphi \mid \varphi \rhd \varphi \mid @_\mathsf{i}\varphi \mid \mathsf{A}$$

where $p \in \mathcal{P}$, and $a \in \mathcal{A}$. Besides the usual abbreviations, $\langle\varphi\rangle\psi$ is an abbreviation for $\varphi\rhd\psi$, which has as a dual $\neg\langle\varphi\rangle\neg\psi$ (also written as $[\varphi]\psi$, but also $\varphi\lhd\psi$, which is another dual, namely $\neg(\neg\varphi\rhd\neg\psi)$. Note that these dualities differ in one negation. The sentence $\neg\mathsf{A}$ is abbreviated as S. $\Box$

Besides the usual connectives of epistemic logic, there are nominals i ($p \in \mathcal{P}$ also ranges over nominals), the @ operator, and the $\rhd$ connective, which is the operator that takes care of action execution. Besides these, there is also a special nullary modality A. A nullary modality is like a propositional variable, except that it is not variable. If one considers models based on the frame associated with the model, the interpretation of A is still the same. It indicates that a world is an action world and not a state.

Definition 4.31 (Models for dyadic hybrid epistemic logic)
A model for $\mathscr{L}_{\mathcal{PA}}^{\triangleright @_{\mathsf{i}} \mathsf{A}}$ is a quintuple $M = (W, \mathsf{W}, R, \mathsf{result}, V)$ such that:

- $W \neq \emptyset$; a set of states or possible worlds;
- $\mathsf{W} \neq \emptyset$; a set of action worlds;
- $R : \mathcal{A} \to 2^{W \times W} \cup 2^{\mathsf{W} \times \mathsf{W}}$; assigns an accessibility relation to each agent (in both states and actions);
- $\mathsf{result} : W \to (\mathsf{W} \rightharpoonup W)$; assigns a (partial) result function to each world;
- $V : \mathcal{P} \to 2^{W \cup \mathsf{W}}$; assigns a set of states and/or actions to each propositional variable and a single world to each nominal.
- $w \in W \cup \mathsf{W}$ □

Definition 4.32 (Semantics for $\mathscr{L}_{\mathcal{PA}}^{\triangleright @_{\mathsf{i}} \mathsf{A}}$)
Let an action model (M, w) be given.

$(M,w) \not\models \bot$		
$(M,w) \models p$	iff	$w \in V(p)$
$(M,w) \models \neg\varphi$	iff	$(M,w) \not\models \varphi$
$(M,w) \models (\varphi \wedge \psi)$	iff	$(M,w) \models \varphi$ and $(M,w) \models \psi$
$(M,w) \models \Box_a \varphi$	iff	$(M,v) \models \varphi$ for all v such that $wR(a)v$
$(M,w) \models \varphi \triangleright \psi$	iff	there is a $\mathsf{w} \in \mathsf{W}$ such that $(M,\mathsf{w}) \models \varphi$ and $\mathsf{result}(w)(\mathsf{w}) = v$ and $(M,v) \models \psi$
$(M,w) \models \mathsf{A}$	iff	$w \in \mathsf{W}$
$(M,w) \models @_{\mathsf{i}}\varphi$	iff	$(M,v) \models \varphi$ for $V(\mathsf{i}) = \{v\}$

Let me give an example of how these semantics work. We take the model from Figure 4.7, and convert it for this system. This is shown in figure 4.9. It is one of the intended models of the logic. But in the semantics there is no restriction on the result relation that ensures that the result relation is anything like the multiplication of the static part of the model with the action part of the model as defined in section 4.3.3, defintion 4.12. This matter will be dealt with below. For the semantics where the restrictions on result are relaxed so that it can be any ternary relation on $W \cup \mathsf{W}$, Ten Cate provides a Hilbert style axiomatization of hybrid logic with multidimensional modalities and a tableau calculus. Here I present the axiom system for the general logic given above.

Definition 4.33 (Proof system DHEL)
The proof system DHEL consists of all axioms and rules of $\mathsf{K}_{\mathcal{PA}}$ and the following axioms and rules

Distr$_{\cdot\triangleleft}$	$((p_1 \to p_2) \triangleleft p_3) \to ((p_1 \triangleleft p_3) \to (p_2 \triangleleft p_3))$	($\cdot\triangleleft$-distribution)
Distr$_{\triangleleft\cdot}$	$(p_1 \triangleleft (p_2 \to p_3)) \to ((p_1 \triangleleft p_2) \to (p_1 \triangleleft p_3))$	($\triangleleft\cdot$-distribution)
Distr$_{@_{\mathsf{i}}}$	$@_{\mathsf{i}}(p_1 \to p_2) \to (@_{\mathsf{i}}p_1 \to @_{\mathsf{i}}p_2)$	($@_{\mathsf{i}}$-distribution)

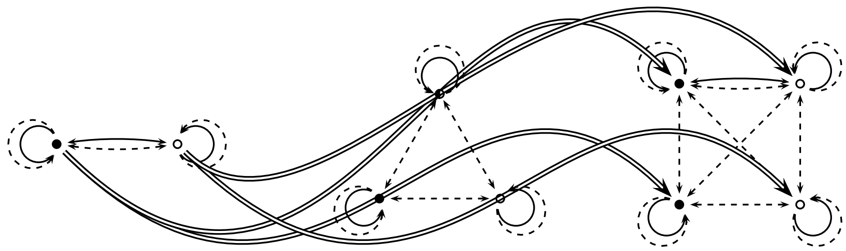


Figure 4.9: A model for dyadic hybrid epistemic logic. The cluster of worlds on the left and the cluster on the right are static possible worlds. The cluster of worlds in the middle are action worlds. The double-lined arrows indicate the ternary result relation. The solid nodes indicate that p is true and the open nodes indicate that p is false. The solid lines represent the accessibility relation for agent a, the dashed lines represent the accessibility relation for agent b. This picture models the same situation as figure 4.7

SD	$@_i p \leftrightarrow \neg @_i \neg p$	($@_i$-selfdual)
$@_i$**T**	$@_i i$	
AG	$@_i @_j p \leftrightarrow @_j p$	(agree)
IN	$i \rightarrow (p \leftrightarrow @_i p)$	(introduction)
SUB	$\dfrac{\varphi}{\varphi[\psi/p]}$	(uniform substitution)
Nec$_{\cdot\triangleleft}$	$\dfrac{\varphi}{\varphi \triangleleft \bot}$	($\cdot\triangleleft$-necessitation)
Nec$_{\triangleleft\cdot}$	$\dfrac{\varphi}{\bot \triangleleft \varphi}$	($\triangleleft\cdot$-necessitation)
Nec$_{@_i}$	$\dfrac{\varphi}{@_i \varphi}$	($@_i$-necessitation)
N	$\dfrac{@_i \varphi}{\varphi}$ where i does not occur in φ	(name)
P	$\dfrac{(@_i(j_1 \triangleright j_2) \wedge @_{j_1}\varphi_1 \wedge @_{j_2}\varphi_2) \rightarrow \psi}{@_i(\varphi_1 \triangleright \varphi_2) \rightarrow \psi}$ where $i \neq j_1$ and $i \neq j_2$ and j_1, j_2 do not occur in $\varphi_1, \varphi_2, \psi$	(paste)

This is only a general proof system for dyadic hybrid epistemic logic. It is sound and complete with respect to the class of models where the result relation is an arbitrary ternary relation.

There is a nice theorem in hybrid logic which states that any extension of the proof system presented above, with pure axioms or rules, i.e. rules in which no propositional variables occur in the sentences, is complete for the class of frames in which those axioms and rules are valid. So Ten Cate continues by capturing the additional restrictions on the models that make the result relation an action execution relation in pure axioms and rules. The logic for epistemic actions has some added features such as the nullary modality A. In the models introduced by Ten Cate an accessibility relation never crosses the bound between actions and states. Therefore the following axioms hold:

$$\begin{aligned} \mathsf{A} &\rightarrow \Box_a \mathsf{A} \\ \neg\mathsf{A} &\rightarrow \Box_a \neg\mathsf{A} \end{aligned}$$

Note that these are all pure sentences, because A is a fixed nullary modality. So the substitution rule is still valid.

The result relation is defined as a ternary relation where the first and last argument must be states and the second argument is an action. So the result relation always crosses the bounds between actions and states. This is captured by the following axioms.

$$\mathsf{A} \rightarrow [\top]\bot$$
$$[\neg\mathsf{A}]\bot$$
$$[\top]\neg\mathsf{A}$$

The result relation is a partial function. This is captured by the following pure axiom.

$$\langle \mathsf{i} \rangle \mathsf{j} \rightarrow [\mathsf{i}]\mathsf{j}$$

The result relation furthermore has what Ten Cate calls the *product property*. Basically it says that a world v' is accessible after the execution of an action w, iff there is a world v accessible now and there is an action v accessible from w such that v' is the result of executing v in v, and vice versa. This is captured by the following axiom and rule.

$$\Diamond_a \langle \mathsf{i} \rangle \mathsf{j} \rightarrow \langle \Diamond_a \mathsf{i} \rangle \Diamond_a \mathsf{j}$$

$$\frac{@_{\mathsf{i}}\langle \mathsf{j} \rangle \Diamond_a \mathsf{k} \wedge @_{\mathsf{i}} \Diamond_a \mathsf{i}' \wedge @_{\mathsf{j}} \Diamond_a \mathsf{j}' \wedge @_{\mathsf{i}'} \langle \mathsf{j}' \rangle \mathsf{k} \rightarrow \varphi}{@_{\mathsf{i}} \langle \mathsf{j} \rangle \Diamond_a \mathsf{k} \rightarrow \varphi}$$

Adding these makes the system complete for the semantics of definition 4.32.

There are things that are not captured yet by this system. Most other dynamic epistemic logics have actions that only affect information. The execution of an action does not change the truth value of propositional variables. This is captured by the atomic permanence axioms. A similar axiom in this system would be something like $p \rightarrow [T]p$. These axioms are essentially impure, because they have to mention propositional variables. The impurity of these axioms does not prevent them from being added to the axiom system, but completeness is no longer guaranteed. Also this would render the substitution rule invalid, because sentences with epistemic operators can change due to action execution: this is the feature that makes it worthwhile to study these logics. So we would have to look for other completeness proof techniques.

Moreover there are no longer explicit preconditions in the action models. It seems difficult to regain these in the semantics. One would also like the property that bisimulation is preserved by action execution. In the system as it currently stands this is not the case. Moreover it is problematic to incorporate the notion of common knowledge into the system. This is of course one of the desired features of a logic that deals with higher-order information.

4.5 Where do we go?

In conclusion we can say that there is no definitive logic for information and information change. It seems that a lot has been accomplished in the various systems presented in this chapter. It is quite clear what the effect of actions with epistemic aspects is on higher-order information. Still a large number of problems remain. The main desideratum is common to all logics, to have the following three things at the same time: a language that is expressive enough, clear semantics, and an elegant proof system.

As far as the language is concerned one would like to have two things. In the first place one would like the usual language of epistemic logic. This is not a major problem. All systems presented in this chapter have the usual epistemic modalities. Common knowledge is only present in some of these systems, although it is a highly desirable feature in any system that aims to describe higher-order information. In the second place one would like to be able to describe the actions that cause information change within the language. We have seen many examples of ways to describe actions. On the one hand the language should be able to describe the actions in such detail that it is clear what is going on. On the other hand it must also be abstract enough so that useful statements can be made about actions in general. There is a tension between these two desiderata. The one extreme is LEA, where action models were inserted into the language. On the other hand we saw ALL, where only very general classes of actions could be described. In the middle we have the systems where actions are described by means of PDL-like actions or by means of nominals that allow the description of action models in more detail. What the PDL-type action algebras are concerned, the one thing that has remained absent is the iteration operator or Kleene star (see chapter 3), which allows the description of while-loops and such. Recent results by Larry Moss and Joseph Miller (which have not been published) say that the logic of public announcements with Kleene star is undecidable. So it remains difficult to choose a good language.

As to the semantics: we would like it to be clear. The main problem for epistemic logics where things change is that sometimes you would like to add or to remove worlds and arrows, as was said earlier. In order to obtain extra worlds

Kripke models can be rolled out, or multiplied, or one can construct Kripke models in such a way that the resulting model is already present in it. An elegant way to solve this remains to be found. It is worth noting that adding the Kleene star to PDL-like actions is not a problem of semantics. One can easily define the reflexive transitive closure of a relation. The main problem of the Kleene star lies with the proof system.

As to proof systems, they remain to be the bottlenecks of most of the logics presented above. It turns out that it is very difficult to add common knowledge to a system and give a proof system where the interaction between common knowledge and actions is captured. A lot of work can be done in this area.

What of the future? Although these problems remain, it seems that due to all these systems we do have quite a clear picture of what effects the execution of actions can have on higher-order information. There is an issue that is not touched upon by any of these systems. Although they give an accurate account of the effects of actions, they do not provide an answer to the question how these actions come about in the first place. The motivation of many of these systems is the analysis of games involving higher-order information and semantics of natural language. In both cases the actions come about because of the intentions of the agents performing the actions. Players of a game often play because they want to win. Natural language is used because those who utter natural language wish to communicate. The actions in the systems presented in this chapter are completely detached from the agents performing them, although in the examples it is clear that they are performed by agents. One cannot even say whose turn it is.

If intentions were added we could view actions as strategies executed by agents. We would then also be able to compare the quality of two actions, given the intentions of the agents. It would allow us to study agent interactions at another level. In the multi-agent community there is much interest in intentions and goals of agents, including higher-order intentions, see Rao and Georgeff (1991, Dunin-Kęplicz and Verbrugge (2002). It would be very interesting to see what insights from that area could add to the study of information and information change.

Chapter 5

Intensional and statistical probability

5.1 Introduction

In the literature about the philosophy of probability theory it is stated that there is widespread consensus about the mathematics of probability theory, contrary to the philosophy of probability theory, where there is almost no consensus but fierce debate Gillies (2000, p.1), Weatherford (1982, p.5), Reichenbach (1949, pp.67–68).

In philosophy of probability many viewpoints are distinguished. One of the main classifications of these viewpoints is the distinctions between *objective* theories of probability – probability is taken to be a part of physical reality – and *epistemic* theories of probability – they see probability as degree of belief. (See Gillies (2000) or Weatherford (1982) for an overview.) This distinction seems related to a distinction made by Bacchus (1990) between *statistical* and *propositional* probabilities. A statistical probability statement is about the proportion of individuals that have a certain property, or in other words the probability that a *randomly chosen* individual has a certain property. A propositional probability statement expresses the probability that a particular individual has a certain property, or in more general terms the probability that a certain proposition is true. In Bacchus (1990) these two viewpoints are both developed into logics. These two logics are different, and the question is whether one can still maintain that there is consensus about the mathematics of probability, since logic is quite mathematical.

In general it is very enlightening when philosophically different notions are investigated formally such as in a logic, so that one can see the differences between related notions at a level that is difficult to attain to informally. Before I turn to these logics and the relationship between these I want to make some comments on the terminology I will use. There are a lot of terms for different notions of probability available in the literature: objective probabilities, subjective probabilities, epistemic probabilities, logical probabilities, statistical probabilities, propositional probabilities, personal probabilities, direct probabilities, indirect probabilities, probability_1, probability_2, etcetera. All are distinct. I wish to add one to this list. Bacchus uses the term propositional probability, I think, mainly for the reason that a propositional probability is a propositional attitude. The term however can be quite confusing in the context of logic. Bacchus develops a first-order logic for propositional probabilities. It could be quite confusing to call this "first-order propositional probability logic". Especially if one is also interested in "propositional propositional probability logic". Therefore I use the term *intensional probability* rather than propositional probability. On the one hand it does not interfere with any other terminology. On the other hand it is in line with modal approaches to other intensional concepts.

Intensional-with-an-s should not be confused with intentional-with-a-t. Intentionality is a term that is often used in the philosophy of mind where it is interpreted as "aboutness". Many mental attitudes such as fear, hope, doubt, belief, anger, intention, and others have this property. If someone hopes that the weather will be nice, then this hope is about the weather. Intensionality is an aspect of language, which often occurs in – what Quine calls – referentially opaque contexts Quine (1980, chapter 8). It is opposite to extensionality. In first-order logic the meaning of a predicate is defined as the set of objects to which the predicate applies: its extension. Consequently, if two predicates such as "has a heart" and "has kidneys" happen to have the same extension, then their meaning is the same and a sentence where one such predicate occurs must be equivalent to the sentence where that predicate has been replaced by the other. The *intension* or "conceptual content" however is quite different. There are contexts where one cannot interchange two terms with the same extension, but with different intensions. Such a context is said to be intensional and sometimes the terms that are used are also said to be intensional. (More on intensionality and extensionality can for instance be found in Gamut (1991, volume 2, page 14, 15).) For terms referring to objects that means that Leibniz's law (the principle of *indiscernability of identicals*), which states that if two terms refer to the same object one can substitute the one term for the other salva veritate, fails: it is necessary that nine is greater than seven. It is also the case that the number of planets is nine. However it is not necessary that the number of planets is greater than seven. Therefore necessity is considered to be an intensional concept. Other concepts such as belief, obligation, all have examples where extensionality fails. They have all been analyzed within modal logic. When probability is viewed as degree of belief extensionality fails for it just as it does for belief. Therefore it seems appropriate to call it intensional probability.

Let me give an example to make these notions more clear. Suppose that there is a lottery, where a winner is drawn from 200 tickets. Let us suppose that there are one hundred people who have bought tickets: 50 people bought one ticket and the remaining 50 people bought three tickets each. Now let us focus on the following sentences.

1. "The probability that someone wins equals $\frac{1}{100}$."
2. "The probability that someone wins equals $\frac{1}{200}$."
3. "The probability that someone wins equals $\frac{3}{200}$."
4. "The probability that someone wins equals 1."

Sentence number 1 is true if we consider it to express a *statistical* probability, because there is exactly one winner the proportion of winners is 1%.

We can interpret sentence 2, 3, and 4 to be true when we consider them to express *intensional* probabilities. As an intensional probability expresses a degree of belief, we must decide *whose* degree of belief we are concerned with. In most contexts it seems fine to take oneself or an ideal *rational agent* to be that person (where it is unclear whether these coincide). Usually an external perspective is chosen. So it is not the perspective of some participant of the lottery or the person who draws the tickets, but the perspective of someone who is completely ignorant about the exact situation, and does know the rules of the game (including the probability distributions).

In first-order modal logic there is a distinction between *de re* and *de dicto* readings of sentences that involve a modality and a quantifier. As an example Quine (1956) gives the sentence "Ralph believes that someone is a spy", which can be read as "Ralph believes that there are spies" and "There is someone whom Ralph believes to be a spy". These are the de dicto and the de re reading respectively. A similar ambiguity is found in the sentences about the lottery when these are taken to express intensional probabilities. We can give a de re and a de dicto reading of these sentence. The de re reading is "There is someone such that the probability that he or she wins equals q". In that case the sentence is true for both $q = \frac{1}{200}$ and $q = \frac{3}{200}$ (i.e. sentence 2 and 3), because there is someone such that the probability that he or she wins equals $\frac{1}{200}$ and there is someone such that the probability that he or she wins equals $\frac{3}{200}$, and there is nobody for whom the probability is different from $\frac{1}{200}$ or $\frac{3}{200}$. Now if we take the de dicto reading we find that sentence 4 is true, because it is certain that there is someone who wins.

These different notions of probability are often confused. This is not surprising since they are related, and sometimes equal. If in the lottery example every person would have had two tickets, then the statistical probability that someone wins would have been the same as the de re intensional probability for someone to win (i.e. $\frac{1}{100}$). Under certain conditions it is possible to derive an intensional probability from a statistical probability. This is called direct inference and is studied by Bacchus (1990). It seems rather natural that randomly selecting from a group of individuals yields possible worlds where in each world a different individual was selected. But these probabilities also interact in other ways. In mathematical statistics statements such as "the probability is 0.95 that between

55% and 65% of all birds fly" can be viewed as intensional probabilities about statistical probabilities.

In this chapter I want to make clear how intensional probability logic and statistical probability logic are related. In intensional probability logic probabilities are assigned to possible worlds. The probability operators are therefore very much like modal operators. In statistical probability logic probabilities are assigned to objects. The probability operators are very much like quantifiers. The relation between these logics therefore is very much like the relation between modal logic and first-order logic. Modal logic and predicate logic are related by what is called the *standard translation*. It was systematically studied by van Benthem (1984). See Blackburn, de Rijke, and Venema (2001) for a modern textbook introduction. One can view the set of possible worlds of a Kripke model for modal logic as the domain of a first-order model. The accessibility relation between possible worlds can be viewed as a relation between elements of the domain, and propositional variables that are true or false in a possible world can be viewed as unary predicates. The language of modal logic can be translated into the language of first-order logic in such a way that a sentence of modal logic is true in a certain model iff the translation of that sentence is true in the same model viewed as a first-order model. So the models are the same mathematically, but the languages that are interpreted in these models are not the same. The same is the case for intensional probability logic and statistical probability logic.

Abadi and Halpern (1994) translate between first-order intensional probability logic and statistical probability logic. In this chapter the focus is on translating *propositional* intensional probability logics to statistical probability logics. These translations make it possible to define fragments of statistical probability logics that have favourable computational properties.

5.2 Languages and Models

Before any kind of translation is presented it must first be clear which logics we are talking about. To get started I first present one of the simplest intensional probability logics: IPL. Then I present the statistical probability logic SPL into which the former logic can be translated.

5.2.1 The intensional probability logic IPL

The logic presented in this section is one of the simplest intensional probability logics. It can be found in Halpern (1991), which is based on Nilsson (1986). I call it IPL. The idea is that there is a set of possible worlds and every possible world has a certain probability. The probability of a sentence is equal to the sum of the probabilities of the worlds in which that sentence holds. The language with which we can express intensional probabilities is simply the language of propositional logic with special sentences to express what the probability of sentences is.

Definition 5.1 (Language of IPL)
Let a countable set of propositional variables $\mathcal{P}$ be given. The language of IPL $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ is given by the following rule in extended Backus-Naur form :

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid q_1\mathbf{P}(\varphi_1) + \cdots + q_n\mathbf{P}(\varphi_n) \geq q$$

where $p \in \mathcal{P}$ and $q_1, \ldots, q_n, q \in \mathbb{Q}$ (no distinction is made between rationals and the names of rationals). Besides the usual abbreviations, we have the following:

$$\textstyle\sum_{i=1}^{n} q_i\mathbf{P}(\varphi_i) \geq q \quad : \quad q_1\mathbf{P}(\varphi_1) + \cdots + q_n\mathbf{P}(\varphi_n) \geq q$$

We also use $>$, $<$, and $\leq$ in the usual way. □

A sentence of the form $\mathbf{P}(\varphi) \geq q$ can be read as "the probability that φ holds is greater than or equal to q". This language can be interpreted in models that are very much like Kripke models for modal logic. The main difference is that there is no accessibility relation, but a probability function that assigns probabilities to all possible worlds.

Definition 5.2 (Intensional probability models)
An intensional probability model M for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ is a triple (W, P, V) such that:

1. $W \neq \emptyset$; a set of possible worlds;
2. $P : W \to [0, 1]$; assigns a probability to each world such that
$$\sum_{w \in W} P(w) = 1$$
3. $V : \mathcal{P} \to 2^W$; assigns a set of worlds to each propositional variable;

Below we often use the notion of a pointed model (M, w). This is a model with a designated world, called its point, which is taken to be the actual world. We also want to generalize the probability function to sets of worlds. If E is a subset of W, then $P(E) = \sum_{v \in E} P(v)$. □

The semantics for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ is standard except for the last clause, which deals with probability sentences. Note that the truth value of these sentences does not depend on the world in which the sentence is evaluated. This is because the probability function is defined for the whole model. Note also that there is no restriction on the cardinality of the set of worlds, but the restriction on the probability function is such that the set of worlds for which the probability is non-zero is at most countable.

Figure 5.1: An intensional probability model for the tossing of a fair coin. The solid world indicates the outcome is heads. The open node indicates the outcome is tails.

Definition 5.3 (Semantics for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$)
Let a propositional probability model $M = (W, P, V)$ for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ and a world $w \in W$ be given.

$$\begin{array}{lll}
(M, w) \not\models \bot & & \\
(M, w) \models p & \text{iff} & w \in V(p) \\
(M, w) \models \neg\varphi & \text{iff} & (M, w) \not\models \varphi \\
(M, w) \models (\varphi \wedge \psi) & \text{iff} & (M, w) \models \varphi \text{ and } (M, w) \models \psi \\
(M, w) \models \sum_{i=1}^{n} q_i \mathbf{P}(\varphi_i) \geq q & \text{iff} & \sum_{i=1}^{n} q_i P(\varphi_i) \geq q
\end{array}$$

where $P(\varphi_i)$ is an abbreviation of $P\{v \in W \mid (M, v) \models \varphi_i\}$. □

As an example consider the situation where a fair coin has been tossed but the outcome has not been observed yet. Thus there are two possibilities: the outcome is heads or the outcome is tails. Because the coin is fair each of these outcomes is equally likely. A picture of the intensional probability model for this situation is given in Figure 5.1. Let p be the propositional variable that expresses that the outcome is heads. In this model the sentence $\mathbf{P}(p) = \frac{1}{2}$ is true, whether the actual world is the world where p holds or whether it is not.

5.2.2 The statistical probability logic SPL

The statistical probability logic SPL is a first-order logic, which views probabilities as terms. However, these terms do not denote elements of the domain of discourse, but they denote elements of the set of rational numbers or the reals. To take this into account the language of SPL has two kinds of terms: object terms and field terms. Object terms denote elements of the domain of discourse, field terms denote numbers (in $\mathbb{R}$). By distinguishing these, it is also possible to distinguish predicates and functions that apply to either object terms or field terms. The logic SPL can be found in Halpern (1990). It is different from the statistical probability logic presented in Bacchus (1990) in that predicates and functions are strictly typed and the set of predicates and functions for field terms is fixed and has the same interpretation in every model.

Definition 5.4 (Language of SPL)
Let a countable set X^{o} of object-variables $x^{\mathrm{o}}_1, x^{\mathrm{o}}_2, \ldots$, a countable set X^{f} of field-variables $x^{\mathrm{f}}_1, x^{\mathrm{f}}_2, \ldots$, a countable set F^n of n-ary object-function symbols $f^n_1, f^n_2, \ldots$

for each $n \in \mathbb{N}$, and a countable set R^n of n-ary object-predicate symbols $R_1^n, R_2^n, \ldots$ for each $n \in \mathbb{N}$ be given, where all these sets are disjoint. Let $\mathsf{X} = \mathsf{X}^{\mathrm{o}} \cup \mathsf{X}^{\mathrm{f}}$, $\mathsf{F} = \cup_{n\in\mathbb{N}}\mathsf{F}^n$, and $\mathsf{R} = \cup_{n\in\mathbb{N}}\mathsf{R}^n$. The language of SPL $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$ consisting of object terms τ^{o}, field terms τ^{f} and formulas φ is given by the following rules in extended Backus-Naur form:

$$\begin{array}{lll}
\tau^{\mathrm{o}} & ::= & x^{\mathrm{o}} \mid f^n(\tau_1^{\mathrm{o}}, \ldots, \tau_n^{\mathrm{o}}) \\
\tau^{\mathrm{f}} & ::= & x^{\mathrm{f}} \mid \mathbf{0} \mid \mathbf{1} \mid \tau_1^{\mathrm{f}} + \tau_2^{\mathrm{f}} \mid \tau_1^{\mathrm{f}} \times \tau_2^{\mathrm{f}} \mid \mathbf{P}x_1^{\mathrm{o}} \ldots x_n^{\mathrm{o}}(\varphi) \\
\varphi & ::= & \bot \mid R^n(\tau_1^{\mathrm{o}}, \ldots, \tau_n^{\mathrm{o}}) \mid \tau_1^{\mathrm{f}} = \tau_2^{\mathrm{f}} \mid \tau_1^{\mathrm{f}} > \tau_2^{\mathrm{f}} \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid \forall x^{\mathrm{o}}\varphi \mid \forall x^{\mathrm{f}}\varphi
\end{array}$$

where x^{o} ranges over X^{o}, x^{f} ranges over X^{f}, f ranges over F^n and R ranges over R^n. The 0-ary function symbols are called individual constants and the 0-ary predicate symbols are called propositional variables. Below I omit the various superscripts and subscripts where it is clear from context what is meant. Besides the usual abbreviations, natural numbers such as 3 are used as abbreviations for terms such as $(\mathbf{1}+\mathbf{1}+\mathbf{1})$. Formulas such as $\tau > q$, where q is a rational number (i.e. there are $k, n \in \mathbb{Z}$ such that $q = \frac{k}{n}$) are abbreviations for $n \times \tau > k$. □

A term of the form $\mathbf{P}x(\varphi)$ can be read "the probability of picking an individual x such that $\varphi(x)$ is true". The operator $\mathbf{P}x$ binds the variable x in $\varphi(x)$. Note that not all numbers in $\mathbb{R}$ are expressible in $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$, for there are only countably many expressions and uncountably many reals.

Definition 5.5 (Statistical probability models)
A statistical probability model for $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$ is a triple $M = (D, I, \mu)$ such that:

1. $D \neq \emptyset$ a set of individuals;

2. I: an interpretation function which assigns an n-ary function over D to every n-ary object-function symbol. It assigns an n-ary relation over D to every n-ary predicate symbol.

3. $\mu : D \to [0, 1]$; assigns a probability to each individual such that

$$\sum_{d\in D} \mu(d) = 1$$

Below we often use models with assignments (M, g), where g assigns an element of the domain to every object variable and an element of $\mathbb{R}$ to every field variable. We also generalize the probability measure to sets of tuples of individuals. If E is a subset of D^n, then $\mu(E) = \sum_{(d_1,\ldots,d_n)\in E} \mu(d_1) \times \cdots \times \mu(d_n)$. □

The idea behind the probability measure μ is that it gives the probability that an individual is chosen. In that sense μ is a selection function. This justifies taking the product measure for the probability of selecting a tuple. The assumption is made that selection is done with replacement. Therefore we want $\mathbf{P}xy(\varphi)$ to be equivalent to $\mathbf{P}yx(\varphi)$.

The language of SPL was defined by simultaneous induction over terms and formulas. Therefore the semantics are also defined for terms and formulas simultaneously.

Definition 5.6 (Semantics for $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$)
Let a statistical probability model $M = (D, I, \mu)$ for $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$ and an assignment g, which assigns an element of the domain to every object variable and an element of $\mathbb{R}$ to every field variable, be given.

$$
\begin{array}{lcl}
[\![x]\!]_{(M,g)} & = & g(x) \\
[\![f(\tau_1,\ldots,\tau_n)]\!]_{(M,g)} & = & I(f)([\![\tau_1]\!]_{(M,g)},\ldots,[\![\tau_n]\!]_{(M,g)}) \\
[\![\mathbf{0}]\!]_{(M,g)} & = & 0 \\
[\![\mathbf{1}]\!]_{(M,g)} & = & 1 \\
[\![\tau_1+\tau_2]\!]_{(M,g)} & = & [\![\tau_1]\!]_{(M,g)} + [\![\tau_2]\!]_{(M,g)} \\
[\![\tau_1\times\tau_2]\!]_{(M,g)} & = & [\![\tau_1]\!]_{(M,g)} \times [\![\tau_2]\!]_{(M,g)} \\
[\![\mathbf{P}x_1^{\mathrm{o}}\ldots x_n^{\mathrm{o}}(\varphi)]\!]_{(M,g)} & = & \mu\{(d_1,\ldots,d_n) \mid (M, g[x_1\mapsto d_1,\ldots,x_n\mapsto d_n]) \models \varphi\}
\end{array}
$$

$$
\begin{array}{lll}
(M,g) \not\models \bot & & \\
(M,g) \models R(\tau_1,\ldots,\tau_n) & \text{iff} & ([\![\tau_1]\!],\ldots,[\![\tau_n]\!]) \in I(R) \\
(M,g) \models \tau_1 = \tau_2 & \text{iff} & [\![\tau_1]\!] = [\![\tau_2]\!] \\
(M,g) \models \tau_1 > \tau_2 & \text{iff} & [\![\tau_1]\!] > [\![\tau_2]\!] \\
(M,g) \models \neg\varphi & \text{iff} & (M,g) \not\models \varphi \\
(M,g) \models (\varphi\wedge\psi) & \text{iff} & (M,g) \models \varphi \text{ and } (M,g) \models \psi \\
(M,g) \models \forall x\varphi & \text{iff} & (M, g[x\mapsto d]) \models \varphi \text{ for all } d \in D
\end{array}
$$

where $g[x \mapsto d]$ is the function that assigns d to x and differs from g at most with regard to the value of x. □

As an example consider a vase with nine marbles: five are black and four are white. Let us assume that one is equally likely to pick any marble. A picture of a statistical probability model for this situation is given in Figure 5.2. Suppose that the interpretation of the unary predicate R is the set of black marbles. The sentence $\mathbf{P}x(R(x)) = \frac{5}{9}$ is true in this model regardless the assignment. As another example the sentence $\mathbf{P}xy(R(x) \wedge \neg R(y)) = \frac{20}{81}$ is also true in this model.

5.3 The relation between IPL and SPL

The relation between IPL and SPL is that the language of IPL can be faithfully translated into the language of SPL. First observe that a model for intensional probability logic can be seen as a model for statistical probability logic if we take the set of worlds to be a domain of individuals, $D = W$, and if we view the propositional variables of intensional probability logic as unary predicates. For

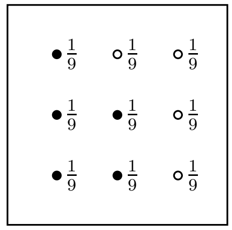


Figure 5.2: A statistical probability model for picking a marble from a vase. The solid nodes represent the black marbles, the open nodes indicate the white marbles.

every propositional variable $p \in \mathcal{P}$ there is a unary predicate $R_p \in \mathsf{R}^1$. The probability function over sets of worlds is then taken as a probability measure over the domain: $\mu = P$. The translation tells us that if the original sentence of intensional probability logic is true in the model at a world, then the translation of that sentence is true in the statistical probability model with an assignment that assigns the actual world to the variable with respect to which the sentence was translated. Thus, the variable with respect to which the sentence is translated has the role of the actual world.

Definition 5.7 (Translation from $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ to $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$)
The translation $t : (\mathsf{X}^{\mathrm{o}} \times \mathscr{L}^{\mathbf{P}}_{\mathcal{P}}) \to \mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$ is defined inductively as follows:

1. $t_x(\bot) = \bot$
2. $t_x(p) = R_p(x)$
3. $t_x(\neg\varphi) = \neg t_x(\varphi)$
4. $t_x(\varphi \wedge \psi) = t_x(\varphi) \wedge t_x(\psi)$
5. $t_x(q_1\mathbf{P}(\varphi_1) + \cdots + q_n\mathbf{P}(\varphi_n) \geq q) =$
$(q_1 \times \mathbf{P}x(t_x(\varphi_1))) + \cdots + (q_n \times \mathbf{P}x(t_x(\varphi_n))) \geq q$

where the first argument of t is written as a subscript. □

Theorem 5.1
Let $M = (W, P, V)$ be a intensional probability model for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$ and let $M' = (D, I, \mu)$ be a statistical probability model for $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R})$. Suppose $D = W$, $\mu = P$ and for every $p \in \mathcal{P}$ there is an $R_p \in \mathsf{R}^1$ such that $I(R_p) = V(p)$. Then:

$$(M, w) \models \varphi \text{ iff } (M', g[x \mapsto w]) \models t_x(\varphi)$$

for all $\varphi \in \mathscr{L}^{\mathbf{P}}_{\mathcal{P}}$. □

Proof By induction on φ. The case for absurdity, propositional variables, negation and conjunction is trivial. Suppose φ is of the form $\sum_{i=1}^{n} q_i \mathbf{P}(\varphi_i) \geq q$.

$(M, w) \models \sum_{i=1}^{n} q_i \mathbf{P}(\varphi_i) \geq q$
$\equiv$ {semantics of intensional probability logic}
$\sum_{i=1}^{n} q_i P(\{v \in W \mid (M, v) \models \varphi_i\}) \geq q$
$\equiv$ {induction hypothesis}
$\sum_{i=1}^{n} q_i \mu(\{v \in W \mid (M', g[x \mapsto v]) \models t_x(\varphi_i)\}) \geq q$
$\equiv$ {semantics of intensional probability logic}
$(M', g) \models \sum_{i=1}^{n} q_i \mathbf{P}x(t_x(\varphi_i)) \geq q$
$\equiv$ {by definition of t_x and x does not occur free in $\mathbf{P}x(t_x(\varphi_i))$}
$(M', g[x \mapsto w]) \models t_x(\sum_{i=1}^{n} q_i \mathbf{P}(\varphi_i) \geq q)$

□

So intensional probability models can be seen as a subclass of statistical probability models, and the language of intensional probability logic can be seen as a sublanguage of the language of SPL.

5.4 The probabilistic epistemic logic PEL

5.4.1 Language and semantics

There is a natural way to generalize intensional probability logic. When probability is viewed as degree of belief, then the agent who assigns these degrees of belief can be made explicit. Thus we enter the realm of multi-agent probability logics. In Fagin and Halpern (1994) and Halpern and Tuttle (1993) a very nice general approach to this is given. It also includes epistemic logic, which can be motivated in two ways. The first is that a distinction must be made between knowledge and certainty. By certainty I mean that if some agent a is certain of something, then the probability a assigns to it equals one (notation $cert_a(\varphi)$). Now when a coin is tossed repeatedly, it can be shown that the probability that a certain infinite sequence of coin tosses occurs is smaller than any positive rational number. Therefore every infinite sequence of coin tosses has probability zero. Hence it is certain for any infinite sequence that it will not occur. For example it is certain that an infinite sequence of heads will not occur. On the other hand one does not know that an infinite sequence of heads will not occur, because that it would occur happens to be consistent with all the information given. Therefore it is appropriate to distinguish certainty and knowledge at a formal level, even though this only matters within infinite contexts.

Secondly one might want to model ignorance about probabilities. This typically occurs when dealing with de re probabilities. Suppose for example that there are two coins. One is fair, and one lands heads one third of the times. Now suppose that an agent cannot distinguish the two coins and one of the coins is tossed, and she knows that it is one of those coins, but not which one. What is the probability according to her that the coin lands heads? If one reads this as a de re probability, then it seems natural to say that she does not know whether the probability of heads is one half or one third. Note that these two de re probabilities yield a unique statistical probability. Picking one of the coins at random and tossing it yields the statistical probability of the coin landing heads, which, if the coins are selected according to a uniform probability distribution, equals $\frac{5}{12}$. In order to capture the de re probabilities one needs to be able to express that an agent is ignorant about probabilities. Therefore one needs to add epistemic operators. The probabilistic epistemic logic presented in this section is called PEL. It will also play an important role in chapter 6.

Definition 5.8 (Language of PEL)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of PEL $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ is given by the following rule in extended Backus-Naur form :

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a\varphi \mid q_1\mathbf{P}_a(\varphi_1) + \cdots + q_n\mathbf{P}_a(\varphi_n) \geq q$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$ and $q_1, \ldots, q_k$ and q are rationals. Again, the usual abbreviations are used. □

A sentence of the form $\mathbf{P}_a(\varphi) \geq q$ should be read as "the probability agent a assigns to φ is greater than or equal to q". Note that in this language higher-order probability statements can be expressed, such as $\mathbf{P}_a(\mathbf{P}_b(\varphi) \geq q_1) \geq q_2$. This expresses that the probability a assigns to the sentence that the probability b assigns to φ is greater or equal to q_1, is greater or equal to q_2. This is higher-order in the sense that it expresses what information an agent has about the information of an(other) agent, completely analogous to the case in epistemic logic where sentences such as $\Box_a\Box_b p$ express that a has information about b's information.

This language is interpreted in probabilistic epistemic models. These are epistemic models together with a function P that assigns a probability function to each agent in each world. Fagin and Halpern define probabilistic epistemic models. For those readers familiar with probability theory, in their probabilistic epistemic models a probability space is assigned to each agent in each world. In this chapter I limit this to models where the σ-algebra of measurable sets is always the powerset of the sample space. Therefore the definition is a bit simpler. Most of the results in this chapter and the next however equally apply to the more general notion of probabilistic epistemic models.

Definition 5.9 (Probabilistic epistemic models)
A probabilistic epistemic model M for $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ is a quadruple (W, R, V, P) such that:

1. $W \neq \emptyset$; a set of possible worlds;

2. $R : \mathcal{A} \to 2^{W \times W}$; assigns an accessibility relation to each agent;
3. $V : \mathcal{P} \to 2^W$; assigns a set of worlds to each propositional variable;
4. $P : (\mathcal{A} \times W) \to (W \rightharpoonup [0, 1])$; such that

$$\forall a \in \mathcal{A} \forall w \in W \sum_{v \in dom(P(a,w))} P(a, w)(v) = 1$$

   assigns a probability function to each agent at each world such that its domain is a non-empty subset of the set of possible worlds. ($\rightharpoonup$ means that it is a partial function; some worlds may not be in the domain of the function.) □

We saw that in Fagin and Halpern (1994) more general models are presented where a probability space is assigned to each agent and each world. The probability functions of the definition above are a special case of these. They are a lot simpler, and in this chapter there is no need for the full generality of probability spaces. The main difference with intensional probability models is that the probability functions assigned to the agents need not be the same in the whole model, they can differ from world to world. Consequently, the question whether a probability sentence holds, does depend on the world in which the sentence is evaluated. Note that definition 5.9 leaves the connection between the epistemic accessibility relation and the probability function completely open.

Definition 5.10 (Semantics for $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$)
Let a probabilistic epistemic model $M = (W, R, V, P)$ for $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ and a world $w \in W$ be given.

$(M, w) \not\models \bot$		
$(M, w) \models p$	iff	$w \in V(p)$
$(M, w) \models \neg\varphi$	iff	$(M, w) \not\models \varphi$
$(M, w) \models (\varphi \wedge \psi)$	iff	$(M, w) \models \varphi$ and $(M, w) \models \psi$
$(M, w) \models \Box_a\varphi$	iff	$(M, v) \models \varphi$ for all v such that $wR(a)v$
$(M, w) \models \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q$	iff	$\sum_{i=1}^{n} q_i P(a, w)(\varphi_i) \geq q$

where $P(a, w)(\varphi_i) = P(a, w)(\{v \in dom(P(a, w)) \mid (M, v) \models \varphi_i\})$. □

5.4.2 Sample space assignments

Before we continue with the general case, let us first focus on an interesting subclass of probabilistic epistemic models which I will explain now: those generated by a sample space assignment and a prior probability distribution. The idea for models such as these is introduced in Halpern and Tuttle (1993) in the context of multi-agent systems (see section 4.2). A multi-agent system is a non-probabilistic system. A probability distribution can be defined on the runs of the system. This together with a sample space assignment determines the probabilities the agents assign to points in the system (a run at a certain time). In the context of probabilistic epistemic models a sample space assignment is a function $S : \mathcal{A} \times W \to 2^W$. Given a *prior* probability distribution P^{prior} on the set of worlds, the domain of the probability function P for each agent in each world is $S(a, w)$. The probability function for elements of the domain is then defined as:

$$P(a, w)(v) = \frac{P^{\text{prior}}(v)}{P^{\text{prior}}(S(a, w))}$$

In this way a probabilistic model is acquired iff $P^{\text{prior}}(S(a, w)) > 0$ for every agent a and every world w. We can view such a model as a statistical probability model in the following way. The domain D is the set of worlds W. For every propositional variable p there is a unary predicate R_p such that $I(R_p) = V(p)$. For every agent a there is a binary relation symbol R_a for the accessibility relation such that $I(R_a) = R(a)$. For every agent a there is also a binary relation symbol S_a for the sample space such that $I(S_a) = \{(w, v) \mid v \in S(a, w)\}$. And we take $\mu = P^{\text{prior}}$.

Now we can give the following translation for the language of PEL. In this definition conditional probabilities are used. Conditional probabilities can be calculated using Kolmogorov's definition Kolmogorov (1956):

$$\mathbf{P}(X|Y) = \frac{\mathbf{P}(X \cap Y)}{\mathbf{P}(Y)} \qquad \text{if } \mathbf{P}(Y) > 0$$

In the language of PEL sentences with conditional probabilities can be seen as abbreviations as will be shown below.

Definition 5.11 (Translation from $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ to $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R})$)
The translation $t : (\mathsf{X}^{\text{o}} \times \mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}) \to \mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R})$ is defined as follows:

1. $t_x(\bot) = \bot$
2. $t_x(p) = R_p(x)$
3. $t_x(\neg\varphi) = \neg t_x(\varphi)$
4. $t_x(\varphi \wedge \psi) = t_x(\varphi) \wedge t_x(\psi)$
5. $t_x(\Box_a\varphi) = \forall y(R_a(x, y) \to t_y(\varphi))$
6. $t_x(q_1\mathbf{P}_a(\varphi_1) + \cdots + q_n\mathbf{P}_a(\varphi_n) \geq q) =$
$q_1\mathbf{P}y(t_y(\varphi_1)|S_a(x, y)) + \cdots + q_n\mathbf{P}y(t_y(\varphi_n)|S_a(x, y)) \geq q$

where the last formula is an abbreviation for a much longer formula without conditional probabilities. For example $q_1\mathbf{P}y(t_y(\varphi_1)|S_a(x, y)) + q_2\mathbf{P}y(t_y(\varphi_2)|S_a(x, y)) \geq q$ is an abbreviation of

$$(q_1 \times \mathbf{P}y(t_y(\varphi_1) \wedge S_a(x, y))) + (q_2 \times \mathbf{P}y(t_y(\varphi_2) \wedge S_a(x, y))) \geq q \times \mathbf{P}y(S_a(x, y))$$

where there are no conditional probabilities. □

Note the resemblance between the clauses for the individual epistemic operators $t_x(\Box_a \varphi)$:

$$\forall y(R_a(x,y) \rightarrow t_y(\varphi))$$

versus the probability operators $t_x(\mathbf{P}_a(\varphi_1) \geq q)$:

$$\mathbf{P}y(t_y(\varphi)|S_a(x,y)) \geq q$$

It is quite tempting to think of the probability operator as a universal quantifier and conditional probability as an implication.When we look at the dual cases $t_x(\Diamond_a \varphi)$:

$$\exists y(R_a(x,y) \wedge t_y(\varphi))$$

versus $t_x(\neg(\mathbf{P}_a(\varphi_1) \geq q))$:

$$\mathbf{P}y(t_y(\varphi)|S_a(x,y)) < q$$

we see that the probability operator is also like an existential quantifier and conditional probability is like a conjunction.

Note also that now the variable with respect to which a sentence is translated is always free in the resulting sentence, contrary to definition 5.7.

Theorem 5.2
Let $M = (W, R, V, P)$ be a probabilistic epistemic model for $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ based on a sample space assignment S, and a prior probability distribution P^{prior}. Let $M = (D, I, \mu)$ be a statistical probability model for $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R})$. Suppose $D = W$; for every propositional variable p there is a unary predicate R_p such that $I(R_p) = v(p)$; for every agent a there is a binary relation symbol R_a such that $I(R_a) = R(a)$; for every agent a there is a binary relation symbol S_a such that $I(S_a) = \{(w,v) \mid v \in S(a,w)\}$. And suppose $\mu = P^{\text{prior}}$. Then

$$(M,w) \models \varphi \text{ iff } (M, g[x \mapsto w]) \models t_x(\varphi)$$

for all $\varphi \in \mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$. □

The proof is very similar to the proof of Theorem 5.1.

This result however holds only for the class of models where there is a common prior probability distribution.

5.4.3 The generalized statistical probability logic GSPL

The question is whether we can view every probabilistic epistemic models as a statistical probability model. The answer is no. In order to translate the language of PEL into the language of a statistical probability logic we need a more general statistical probability logic than SPL. In this section a new generalized statistical probability logic GSPL is presented. In probabilistic epistemic models a probability function is assigned to each agent and world. Essentially this probability function says what probability the agent assigns to some world that she is in a particular world. If we want to have a similar approach to this as we gave for SPL, we want to view the set of worlds of the probabilistic epistemic model as the domain of a statistical probability model. But now the probability of selecting an element d_2 after first selecting an element d_1 need not be the same as the probability of selecting d_2 first. Thus we need a more general approach to selection functions. A selection function determines the probability that an element is selected from the domain. There are a lot of choices to be made. For example is the selection made with replacement or without? If it is without replacement then the selection is not independent of any previous selection. In general we have a protocol that determines the probabilities of selecting elements from the domain, possibly dependent upon previous selection. Let us associate an agent with every such protocol.

Definition 5.12 (Language of GSPL)
Let a countable set X^{o} of object-variables $x^{\mathrm{o}}_1, x^{\mathrm{o}}_2, \ldots$, a countable set X^{f} of field-variables $x^{\mathrm{f}}_1, x^{\mathrm{f}}_2, \ldots$, a countable set F^n of n-ary object-function symbols $f^n_1, f^n_2, \ldots$ for each $n \in \mathbb{N}$, a countable set R^n of n-ary object-predicate symbols $R^n_1, R^n_2, \ldots$ for each $n \in \mathbb{N}$, and a finite set of agents $\mathcal{A}$ be given, where all these sets are disjoint. Let $\mathsf{X} = \mathsf{X}^{\mathrm{o}} \cup \mathsf{X}^{\mathrm{f}}$, $\mathsf{F} = \cup_{n\in\mathbb{N}} \mathsf{F}^n$, and $\mathsf{R} = \cup_{n\in\mathbb{N}} \mathsf{R}^n$. The language of GSPL $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R}, \mathcal{A})$ consisting of object terms τ^{o}, field terms τ^{f} and formulas φ is given by the following rules in extended Backus-Naur form:

$$\begin{array}{lll}
\tau^{\mathrm{o}} & ::= & x^{\mathrm{o}} \mid f(\tau^{\mathrm{o}}_1, \ldots, \tau^{\mathrm{o}}_n) \\
\tau^{\mathrm{f}} & ::= & x^{\mathrm{f}} \mid \mathbf{0} \mid \mathbf{1} \mid \tau^{\mathrm{f}}_1 + \tau^{\mathrm{f}}_2 \mid \tau^{\mathrm{f}}_1 \times \tau^{\mathrm{f}}_2 \mid \mathbf{P}_a x^{\mathrm{o}}_1 \ldots x^{\mathrm{o}}_n(\varphi) \\
\varphi & ::= & \bot \mid R(\tau^{\mathrm{o}}_1, \ldots, \tau^{\mathrm{o}}_n) \mid \tau^{\mathrm{f}}_1 = \tau^{\mathrm{f}}_2 \mid \tau^{\mathrm{f}}_1 > \tau^{\mathrm{f}}_2 \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid \forall x^{\mathrm{o}}\varphi \mid \forall x^{\mathrm{f}}\varphi
\end{array}$$

where x^{o} ranges over X^{o}, x^{f} ranges over X^{f}, f ranges over F^n, R ranges over R^n, and a ranges over $\mathcal{A}$. □

The only difference with the language of SPL $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R})$ provided in definition 5.4 is that now the probability operators have a subscript, which indicates which agent is performing the selection. So now a term of the form $\mathbf{P}_a x_1 \ldots x_n(\varphi)$ can be read as the probability that the sequence $x_1 \ldots x_n$ randomly selected by agent a satisfies φ. In this case we do not want $\mathbf{P}_a xy(\varphi)$ to be equivalent to $\mathbf{P}_a yx(\varphi)$, because earlier choices affect future choices.

Definition 5.13 (Generalized statistical probability models)
A statistical probability model for $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R}, \mathcal{A})$ is a triple $M = (D, I, \mu)$ such that:

1. $D \neq \emptyset$ a set of individuals;

2. I: an interpretation function which assigns an n-ary function over D to every n-ary object-function symbol. It assigns an n-ary relation over D to every n-ary predicate symbol.

3. $\mu : (\mathcal{A} \times \mathbb{N}) \to (D^n \to [0,1])$; assigns a probability to each element of D^n for each agent a and each $n \in \mathbb{N}$, such that for each $n \in \mathbb{N}$ and each $(d_1, \ldots d_{n-1})$
$$\sum_{d \in D} \mu(a,n)(d_1, \ldots d_{n-1}, d) = 1$$

Below we often use models with assignments (M, g), where g assigns an element of the domain to every object variable and an element of $\mathbb{R}$ to every field variable. We also want to generalize the measure μ in this case. If E is a subset of D^n, then $\mu(a,n)(E) = \sum_{(d_1,\ldots,d_n) \in E} \mu(a,1)(d_1) \times \cdots \times \mu(a,n)(d_1, \ldots, d_n)$. This yields a probability function over D^n. □

We can see that this is a generalization of the models given in Definition 5.5. In that case there was just one function assigning probabilities to the domain that was generalized to tuples by taking the product measure. Here we can do this by taking one agent where all the probability measures are independent of previous choices. The definition of the semantics is equal to Definition 5.6, but now in the clause

$$[\![\mathbf{P}_a x_1^{\circ} \ldots x_n^{\circ}(\varphi)]\!]_{(M,g)} = \mu(a)\{(d_1, \ldots, d_n) \mid (M, g[x_1 \mapsto d_1, \ldots, x_n \mapsto d_n]) \models \varphi\}$$

the function μ is interpreted differently, and an agent must be specified.

Let me give an example of how these semantics can be used to analyze simple problems of probability. This is an example adapted from one of the puzzle books by Smullyan (1997). There is a cabinet with three drawers. Each contains two jewels. One drawer contains two rubies. One drawer contains two emeralds. One drawer contains one emerald and one ruby. Now an agent a picks a drawer at random and picks a jewel from that drawer at random. Then he picks the other jewel from the same drawer. We can make a generalized statistical probability model for this. As the domain D we take the six jewels $\{j_1, \ldots, j_6\}$. We have two predicates R_1 and R_2 such that $I(R_1) = \{j_1, j_2, j_3\}$ (the rubies) and $I(R_2) = \{j_4, j_5, j_6\}$ (the emeralds). The probability of selecting any one of these initially is the same: thus $\mu(a,1)(j_i) = \frac{1}{6}$. However the probabilities of selecting the next jewel are not all the same. Let us say that the drawers contain $\{j_1, j_2\}$, $\{j_3, j_4\}$, and $\{j_5, j_6\}$. Then $\mu(a,2)$ should be defined as follows $\mu(a,2)(j_1, j_2) = 1$ and $\mu(a,2)(j_1, j_k) = 0$ for $j_k \neq j_2$. Similarly $\mu(a,2)(j_2, j_1) = 1$ and $\mu(a,2)(j_2, j_k) = 0$ for $j_k \neq j_1$. And we can define $\mu(a,2)$ analogously for the other drawers. We leave $\mu(a,3)$ and higher-probability measure functions undefined. Now we can answer questions such as: what is the probability of picking two rubies?

$$\mathbf{P}_a xy(R_1(x) \wedge R_1(y))$$

According to the semantics, the interpretation of this term is equal to

$$\mu(a,2)\{(d_1,d_2) \mid d_1 \in I(R_1) \text{ and } d_2 \in I(R_1)\}$$

This is equal to $\mu(a,2)\{j_1,j_2,j_3\}\times\{j_1,j_2,j_3\}$. However $\mu(a,2)$ equals 1 for (j_1,j_2) and (j_2,j_1) and is zero for the other pairs. Therefore these are the only factors that count. And so we multiply $\mu(a,1)(j_1) = \frac{1}{6}$ with 1 and add $\mu(a,1)(j_2) = \frac{1}{6}$ multiplied with 1. This yields $\frac{1}{3}$.

The question posed in the puzzle is what the probability is of picking another ruby given that you initially picked a ruby, i.e.

$$\mathbf{P}_a xy(R_1(y)|R_1(x))$$

By defining conditional probability in the standard way we get that the interpretation of this term is

$$\frac{\mu(a,2)\{(d_1,d_2) \mid d_1 \in I(R_1), d_2 \in I(R_1)\}}{\mu(a,2)\{(d_1,d_2) \mid d_1 \in I(R_1)\}} = \frac{2}{3}$$

Note that in the denominator d_2 is vacuous. This corresponds to the validity

$$\mathbf{P}_a x_1 \ldots x_n y(\varphi) = \mathbf{P}_a x_1 \ldots x_n(\varphi)$$

if y does not occur free in φ.

Now a probabilistic epistemic model can simply be seen as agents selecting worlds.

Definition 5.14 (Translation from $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ to $\mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R},\mathcal{A})$)
The translation $t : \mathscr{L}^{\mathbf{P}}_{\mathcal{PA}} \rightarrow \mathscr{L}^{\mathbf{P}}(\mathsf{X},\mathsf{F},\mathsf{R},\mathcal{A})$ is defined as follows:

1. $t_x(\bot) = \bot$
2. $t_x(p) = R_p(x)$
3. $t_x(\neg\varphi) = \neg t_x(\varphi)$
4. $t_x(\varphi \wedge \psi) = t_x(\varphi) \wedge t_x(\psi)$
5. $t_x(\Box_a\varphi) = \forall y(R_a(x,y) \rightarrow t_y(\varphi))$
6. $t_x(q_1\mathbf{P}_a(\varphi_1) + \cdots + q_n\mathbf{P}_a(\varphi_n) \geq q) =$
$$q_1\mathbf{P}_a yz(t_z(\varphi_1)|x=y) + \cdots + q_n\mathbf{P}_a yz(t_z(\varphi_n)|x=y) \geq q$$

where again the first argument is written as a subscript. □

The last clause again is an abbreviation of a much longer formula without conditional probability. The basic idea is best explained in a simple case where $n = 1$

$$t_x(\mathbf{P}_a(\varphi) \geq q) = \mathbf{P}_a yz(t_z(\varphi)|x = y) \geq q = \mathbf{P}_a yz(t_z(\varphi) \wedge x = y) \geq q \times \mathbf{P}_a y(x = y)$$

The identity $x = y$ is used such that x is still free in the translation. The analogy between the similarity of the translation of $\Box_a\varphi$ and $\mathbf{P}_n(\varphi) \geq q$ is a little more difficult to see than in the case where models were generated by a sample space assignment and a prior probability distribution presented in section 5.4.2, because the sample space assignment is no longer actually present, but hidden in the order of the variables. The idea of the translation is that we need the probability that the agent will select a φ world given that the actual world was first selected. To view a probabilistic epistemic model as a generalized statistical probability model we need the following:

1. $D = W$
2. for every $p \in \mathcal{P}$ there is a $R_p \in \mathsf{R}^1$ such that $I(R_p) = V(p)$
3. for every agent $a \in \mathcal{A}$ there is a $R_a \in \mathsf{R}^2$ such that $I(R_a) = R(a)$
4. $\mu(a, 1)(w) \neq 0$ for all $a \in \mathcal{A}$ and all $w \in W$
5. $\mu(a, 2)(w, v) = P(a, w, v)$ for all $a \in \mathcal{A}$ and all $w, v \in W$

The requirement on $\mu(a, 1)$ ensures that there are no divisions by zero. This requirement does affect the generality of the theorem stated below, because to be able to satisfy the requirement the set of worlds W must be countable. Now we can state the familiar theorem.

Theorem 5.3
Let $M = (W, R, V, P)$ be a probabilistic epistemic model for $\mathscr{L}^{\mathbf{P}}_{\mathcal{P}\mathcal{A}}$ and let $M = (D, I, \mu)$ be a generalized statistical probability model for $\mathscr{L}^{\mathbf{P}}(\mathsf{X}, \mathsf{F}, \mathsf{R}, \mathcal{A})$. Suppose the requirements stated above are satisfied. Then

$$(M, w) \models \varphi \text{ iff } (M, g[x \mapsto w]) \models t_x(\varphi)$$

for all $\varphi \in \mathscr{L}^{\mathbf{P}}_{\mathcal{P},\mathcal{A}}$. □

Again the proof is very similar to the proof of theorem 5.1.

5.5 Toward correspondence theory for probability logics

In the case of modal logic and predicate logic the standard translation has given rise to an area of research known as correspondence theory: a systematic study of the expressiveness of modal logic and first-order logic. In this section I sketch how these theorems can be used to do a similar investigation for probability logics. In Abadi and Halpern (1994) the relation between first-order intensional probability logic and statistical probability logic is investigated. The language of first-order intensional probability logic can also be translated into the language of statistical probability logic. Abadi and Halpern show that the reverse is also possible, which is quite interesting. These translations are also used to investigate the expressive power of these logics.

5.5.1 Correspondences

In probabilistic epistemic logics there are a number of axioms that are intuitively appealing. In Fagin and Halpern (1994) a number of these axioms are discussed. If one takes probability to represent the degree of belief it seems natural to introduce axioms for probability that also hold for belief. Let us for simplicity consider the class of models that are generated by a prior probability distribution and a sample space assignment. What is the probabilistic version of axiom **D** ($\Box_a\varphi \rightarrow \Diamond_a\varphi$)? One might expect it is $\mathbf{P}_a(\varphi) \geq q \rightarrow \neg\mathbf{P}_a(\neg\varphi) \geq q$. This axiom however does not hold in any probabilistic epistemic model. Take $\varphi = \top$ and $q = 0$. Now the probability of $\top$ is definitely greater that 0, however the probability of $\bot$ is not less than 0. So we have to consider another version of **D**. In the case of epistemic logic the axiom says that the accessibility relation is serial. In the context of probabilistic epistemic logic seriality means that the domain of the probability function is non-empty, but this is automatically the case. In fact we have to consider the following formulation of the axiom in terms of certainty

(PD) $cert_a\varphi \rightarrow \neg cert_a\neg\varphi$

Let us see what the translation of this axiom is. We want it to hold in all the worlds so we add a universal quantifier:

$$\forall x(\mathbf{P}y(t_y(\varphi) \wedge S_a(x,y))) = \mathbf{P}y(S_a(x,y)) \rightarrow \mathbf{P}y(\neg t_y(\varphi) \wedge S_a(x,y)) \neq \mathbf{P}y(S_a(x,y))$$

This should really be read as a formula where there is an implicit second order quantification over φ. In that case the formula holds in a statistical probability model iff $\forall x\mathbf{P}y(S_a(x,y)) > 0$, which means that the sample space assigned to a has a probability greater than zero, and therefore is non-empty. The lesson is that we have to be careful how we render the axioms for belief in a probabilistic context. The epistemic operators can not simply be replaced by probabilistic operators.

We can also see this with the axioms corresponding to positive and negative introspection. In epistemic logic these correspond respectively to transitivity and euclidicity of the accessibility relation. In the light of the analogue of axiom **D** one might want to formulate these as $cert_a\varphi \rightarrow cert_a cert_a\varphi$ and $\neg cert_a\varphi \rightarrow$

$cert_a \neg cert_a \varphi$. These axioms do indeed correspond to the property that the sample space assignment is transitive and euclidean, which can be seen by translating these axioms. However in probabilistic epistemic logic in general these axioms do not represent positive and negative introspection. Although it does ensure that the sample space assigned to a world is the same for all the worlds in its sample space, the probabilities assigned to these worlds may vary wildly. The point about introspection is that an agent is fully aware of his own beliefs. This is better expressed by the following two axioms, because these express that the agents know their degree of their own beliefs.

(P4) $\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q \rightarrow \mathbf{P}_a(\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q) = 1$

(P5) $\neg \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q \rightarrow \mathbf{P}_a(\neg \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q) = 1$

These together are called UNIF in Fagin and Halpern (1994). UNIF implies that the probabilities assigned to all the worlds in the sample space are the same for all the worlds in the sample space.

As was said earlier in the general case there is no connection between the epistemic accessibility relation of an agent and the probability function assigned to an agent. If one wants to impose a connection, one of the natural requirements is expressed by the following axiom:

CONS $\Box_a \varphi \rightarrow \mathbf{P}_a(\varphi) = 1$

The name for this axiom is also introduced in Fagin and Halpern (1994). It says that knowledge implies certainty. This axiom is valid on a frame iff the domain of the probability function assigned to a is a subset of the set of worlds accessible through the epistemic accessibility relation of the agent: $domP(a, w) \subseteq \{v \mid wR(a)v\}$ for all $w \in W$. This can easily be seen by the translation of the axiom into statistical probability logic.

$$\forall x(y(R_a(x, y) \rightarrow t_y(\varphi)) \rightarrow \mathbf{P}y(t_y(\varphi)|S_a(x, y)) = 1)$$

This must be seen as a universal second order formula, where $t_x(\varphi)$ is a free variable over relations. We can choose $R_a(x, y)$. Then the formula is equivalent to:

$$\forall x(\mathbf{P}y(R_a(x, y)|S_a(x, y)) = 1)$$

which is equivalent to $\forall xy(S_a(x, y) \rightarrow R_a(x, y))$.

It is nice to see these correspondences as they give a better picture of the notions of probability involved and the relationship between intensional and statistical probability. One of the interesting open questions is whether it is possible to express that the agents have a common prior, which was assumed for an interesting subclass of probabilistic epistemic models. A common prior is also often assumed in game theory Aumann (1976).

5.5.2 Complexity

One of the main differences between intensional probability logics and statistical probability logics is that the former are usually decidable whereas the latter are not. In Fagin, Halpern, and Megiddo (1990) and Fagin and Halpern (1994) a lot of interesting results are presented about the complexity of the decidability problems for intensional probability logics. It can be seen that the complexity of propositional intensional probability logic is very much like propositional logic, and the complexity of probabilistic epistemic logic is very much like modal logic. Both are decidable, but for most decision problems it is not computationally feasible to decide the problem.

The complexity of statistical probability logic is very much like first-order logic. In Abadi and Halpern (1994) the complexity issues regarding statistical probability logic and first-order intensional probability logic are investigated. One can translate first-order intensional probability logic into statistical probability logic, but it is also possible to do so the other way around. The complexity of these logics is also investigated in Abadi and Halpern (1994). In short, statistical probability logic is in general undecidable. It is an even worse case than first-order logic, where the set of validities is recursively enumerable, whereas this is not the case here. I suspect that the situation might again be worse for GSPL. When the class of models is restricted to rational numbers, then it gets a little better. And as with monadic first-order logic, monadic statistical probability logic is decidable. One of the interesting questions that arises when viewing these results is which fragment of statistical probability logic still has nice computational properties. It seems that the translations presented in this chapter could yield something interesting since the translation of probabilistic epistemic logic does contain binary predicate symbols and equality, although the models are restricted to rational numbers. The notion of bisimulation for probabilistic epistemic logic presented in section 6.5 might also shed some light on these issues.

5.6 Conclusion

In this chapter the relationship between intensional probability and statistical probability was investigated. In intensional probability logic probabilities are assigned to possible worlds, in statistical probability logic probabilities are assigned to individuals. When worlds are taken to be individuals one can translate the language of intensional probability logic into the language of statistical probability logic in such a way that truth is preserved. This makes it possible to investigate the expressive power and complexity of these logics.

Chapter 6

Probabilistic dynamic epistemic logic

6.1 Introduction

Epistemic logic is a modal logic used to reason about information, including higher-order information (see chapter 2). Dynamic epistemic logics are extensions of epistemic logic which can be used to reason about information and information change (see chapter 4). In probability theory Bayesian updating can be seen as a model for information change, but higher-order information is overlooked. This is a problem when one wants to formalize inferences about changing probabilistic higher-order information.

In this chapter I combine the probabilistic logic PEL (see section 5.4) with the dynamic epistemic logic DEL (see section 4.3.1) yielding a new logic, PDEL, that deals with changing probabilities and takes higher-order information into account. The semantics of PDEL are introduced in section 6.3. In section 6.4 I give a method for making models for specific situations and I provide a sound and complete proof system for PDEL. Bisimulation for probabilistic epistemic models is introduced in section 6.5. In section 6.6 some examples of application are discussed. Finally, in section 6.7 some conclusions are drawn and some directions for further research are indicated. But before that I want to make clear why I develop PDEL in the first place.

6.2 Motivation

In this section I will make clear why a combination of dynamic epistemic logic and probability theory is worthwhile and what the scope of this chapter is.

Just as in the case of epistemic logic, in probability theory the matter of incorporating newly acquired information has been investigated. In probability theory this is done by taking *posterior* probabilities instead of *prior* probabilities, i.e. the conditional probabilities given the new information, which is also called

Bayesian updating. Posterior probabilities can be calculated using Kolmogorov's definition (see page 89). The idea is that $\mathbf{P}(X|Y)$ gives one the probability of X after one gets the information that Y is the case. So posterior probability can be used to model information change.

Although the distinction between improbable and impossible events and ignorance about probabilities make PEL an appealing system for reasoning about probability, the main motivation for using PEL as the basis for PDEL is that probabilistic higher-order information can be studied in PEL, and that by making a dynamic version of PEL we can study probabilistic higher-order information change. It is interesting to note that both in dynamic epistemic logic and in probability theory, the incorporation of new information is studied. But they seem to come up with different answers to how this is to be done properly. The difference is that in dynamic epistemic logic more kinds of information change are distinguished that explicitly take higher-order information into account. The intriguing question that pops up is what these two fields could learn from each other with respect to information change.

Fortunately a formal connection between the two areas has been established (see Bacchus (1990) and Halpern (1991) for details), showing that probabilistic logic can be seen as an extension of (non-dynamic) epistemic logic. The language of epistemic logic can be seen as a fragment of the language of probabilistic logic. This is done by relating belief and certainty. Standard probability theory can be seen as an extension of KD45. Now let us focus on certainty and conditional certainty: $\mathbf{P}(\varphi) = 1$, which is abbreviated by $cert(\varphi)$ and $\mathbf{P}(\varphi \mid \psi) = 1$, which is abbreviated by $cert(\varphi \mid \psi)$. We get the following sentence for conditional certainty:

$$\mathbf{P}(\psi) > 0 \rightarrow (cert(\varphi \mid \psi) \leftrightarrow cert(\psi \rightarrow \varphi))$$

The consequent of this implication is very much like the Knowledge-Update axiom (also called the generalized Ramsey axiom) of dynamic epistemic logic (see figure 6.2 on page 109).

$$[\psi]\Box\varphi \leftrightarrow \Box(\psi \rightarrow [\psi]\varphi)$$

The only difference, apart from notation (see section 6.3 for the definitions), is that instead of $\psi \rightarrow \varphi$ in the case of probabilistic logic, we have $\psi \rightarrow [\psi]\varphi$ in dynamic epistemic logic. This crucial difference is due to a difference in perspective on information change: in dynamic epistemic logic learning that ψ can change the truth value of φ. In probabilistic logic this is assumed not to be the case. This difference in perspective only becomes apparent when one is interested in higher-order information. Assuming that learning something does not change facts (i.e. truth values of propositional variables), the truth value of φ can only change if an agent learns that ψ, if φ somehow involves a statement about the information the agent has. In this chapter I develop a probabilistic dynamic epistemic logic that does take into account that the truth value of sentences can change due to information change. To keep this chapter simple I limit updates to *public announcements*, i.e. all agents simultaneously get the same information it being common knowledge that they receive it. This simple dynamic epistemic logic is introduced in Gerbrandy and Groeneveld (1997) and I will call it DEL in this chapter. I think that probability theory could greatly benefit from the theory of information change provided by dynamic epistemic logic.

6.3 Language and semantics

We extend the language of PEL with update operators from DEL, thus obtaining a new language for reasoning about probability and information change.

Definition 6.1 (Language of PDEL)
Let a countable set of propositional variables $\mathcal{P}$ and a finite set of agents $\mathcal{A}$ be given. The language of PDEL $\mathscr{L}^{\mathbf{P}[\cdot]}_{\mathcal{PA}}$ is given by the following rule in extended Backus-Naur form :

$$\varphi ::= p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_a\varphi \mid [\varphi_1]\varphi_2 \mid q_1\mathbf{P}_a(\varphi_1) + \cdots + q_n\mathbf{P}_a(\varphi_n) \geq q$$

where $p \in \mathcal{P}$, $a \in \mathcal{A}$ and $q_1, \ldots, q_k$ and q are rationals. Besides the usual abbreviations, we have the following.

$$\begin{array}{lcl}
\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \geq q & : & q_1\mathbf{P}_a(\varphi_1) + \cdots + q_n\mathbf{P}_a(\varphi_n) \geq q \\
q_1\mathbf{P}_a(\varphi) \geq q_2\mathbf{P}_a(\psi) & : & q_1\mathbf{P}_a(\varphi) - q_2\mathbf{P}_a(\psi)) \geq 0 \\
\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \leq q & : & \sum_{i=1}^{n} - q_i\mathbf{P}_a(\varphi_i) \geq -q \\
\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) < q & : & \neg(\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \geq q) \\
\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) > q & : & \neg(\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \leq q) \\
\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) = q & : & (\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \leq q) \wedge (\sum_{i=1}^{n} q_i\mathbf{P}_a(\varphi_i) \geq q)
\end{array}$$

The language of PEL $\mathscr{L}^{\mathbf{P}}_{\mathcal{PA}}$ consists of those sentences of $\mathscr{L}^{\mathbf{P}[\cdot]}_{\mathcal{PA}}$ in which no update operators occur. □

A sentence of the form $[\varphi]\psi$ can be read as "ψ is the case, after everyone simultaneously and commonly learns that φ is the case". In order to interpret this language we have to give two definitions simultaneously, i.e. a truth-definition and a definition of updated models. These definitions are interdependent, but not circular.

Definition 6.2 (Semantics for $\mathscr{L}^{\mathbf{P}[\cdot]}_{\mathcal{PA}}$)
Let a probabilistic epistemic model $M = (W, R, V, P)$ and a world $w \in W$ be given (see section 5.4).

$$\begin{array}{lll}
(M, w) \models p & \text{iff} & w \in V(p) \\
(M, w) \models \neg\varphi & \text{iff} & (M, w) \not\models \varphi \\
(M, w) \models (\varphi \wedge \psi) & \text{iff} & (M, w) \models \varphi \text{ and } (M, w) \models \psi \\
(M, w) \models \Box_a\varphi & \text{iff} & (M, v) \models \varphi \text{ for all } v \text{ such that } wR(a)v
\end{array}$$

$$\begin{array}{lll} (M,w) \models [\varphi]\psi & \text{iff} & (M_\varphi, w_\varphi) \models \psi \text{ (see definition 6.3)} \\ (M,w) \models \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q & \text{iff} & \sum_{i=1}^{n} q_i P(a,w)(\varphi_i) \geq q \end{array}$$

where $P(a,w)(\varphi_i) = P(a,w)(\{v \in dom(P(a,w)) \mid (M,v) \models \varphi_i\})$. □

Definition 6.3 (Semantics for updates)
Let a probabilistic epistemic model $M = (W, R, V, P)$ and a world $w \in W$ be given. The *updated* model $M_\varphi = (W_\varphi, R_\varphi, V_\varphi, P_\varphi)$ is defined as follows.

$$\begin{array}{lll} W_\varphi & = & W \\ R_\varphi(a) & = & \{(u,v) \mid (u,v) \in R(a) \text{ and } (M,v) \models \varphi\} \\ V_\varphi & = & V \\ dom(P_\varphi(a,u)) & = & \begin{cases} dom(P(a,u)) & \text{if } P(a,u)(\varphi) = 0 \\ \{v \in dom(P(a,u)) \mid (M,v) \models \varphi\} & \text{otherwise} \end{cases} \\ P_\varphi(a,u)(v) & = & \begin{cases} P(a,u)(v) & \text{if } P(a,u)(\varphi) = 0 \\ \dfrac{P(a,u)(v)}{P(a,u)(\varphi)} & \text{otherwise given that } v \in dom(P_\varphi(a,u)) \end{cases} \end{array}$$

For a pointed model (M, w) the updated model is (M_φ, w) (i.e. $w_\varphi = w$). □

Announcing φ yields an updated model which is a copy of the original model. It is not an identical copy of the original model, for the accessibility relations and the probability functions differ. Worlds where φ does not hold are no longer accessible to any of the agents. The probability functions are treated similarly to accessibility relations. Worlds where φ does not hold are no longer in the domain of the function. Note that the announcement of φ does not presume that φ is actually true. Consequently an update can always be executed[1], i.e. it holds in general that $\langle\varphi\rangle\top$.

However an update only changes the probability functions of those agents who assign non-zero probability to φ. There are some approaches in probability theory for updating with sentences that have probability zero. The most common one is to leave it undefined. If we would take this approach in case of probabilistic logic, there would be truth value gaps, which would make it very difficult to give a complete proof system. Another approach is to assign probability zero to everything after an update with a sentence that has probability zero. This approach is found in Bacchus (1990), but also in probability theory, for example in Prohorov and Rozanov (1969). This would seem to go against the laws of modal logic; after learning a sentence with probability zero, even the truth would be assigned probability zero. By analogy to *ex falso sequitur quodlibet* it would be more appropriate to assign probability one to everything in that case. This on the other hand would go against the laws of probability theory. So both choices would

[1]There are other dynamic epistemic logics which limit public announcements to *truthful* public announcements where only true announcements can be made.

make it difficult to provide a complete proof system. There are more advanced approaches to updating with sentences with probability zero (see Halpern (2001) for an overview of the different approaches and references). All these approaches handle updating with sentences that have a non-empty set of worlds where that sentence holds. However updating with a sentence that does not hold in any world such as the absurdity remains a problem, and would still result in truth value gaps.

Dynamic epistemic logic cannot deal well with updates with inconsistent information as well. Typically, the accessibility relations become empty after an inconsistent update. A method of *revision* such as it is studied in belief revision is not available here. In PDEL too, we must also deal with updates with information that has probability zero in a way that is not intuitively appealing. The approach given in definition 6.3 is simply to ignore the information. This is to ensure that one does not divide by zero. There is no compelling philosophical reason for this choice, except maybe that the agent would just not believe the information received, and would therefore leave things as they were. This makes the proof system relatively simple.

Lemma 6.1
If (M, w) is a probabilistic epistemic model, then (M_φ, w_φ) is a probabilistic epistemic model too. □

Proof The only difficulty lies in whether P_φ assigns a probability function to each agent in each world. Take a world u and an agent a. If $P(a,u)(\varphi) = 0$, then $P_\varphi(a,u) = P(a,u)$ and therefore it is a probability function. If $P(a,u)(\varphi) \neq 0$, then the domain of $P_\varphi(a,u)$ is exactly the set of worlds in the original domain where φ holds, therefore:

$\sum_{v \in dom(P_\varphi(a,u))} P_\varphi(a,u)(v)$

$=$ {definition of the domain}

$\sum_{v \in dom(P(a,u)) \text{ and } (M,v) \models \varphi} P_\varphi(a,u)(v)$

$=$ {definition of the probability}

$\sum_{v \in dom(P(a,u)) \text{ and } (M,v) \models \varphi} \frac{P(a,u)(v)}{P(a,u)(\varphi)}$

$=$ {algebra}

$$\frac{\sum_{v \in dom(P(a,u)) \text{ and } (M,v) \models \varphi} P(a,u)(v)}{P(a,u)(\varphi)}$$

$=$ {definition $P(a,u)(\varphi)$}

$$\frac{P(a,u)(\varphi)}{P(a,u)(\varphi)}$$

$=\{P(a,u)(\varphi) \neq 0\}$

1

Moreover $P(a,u)(v) \in [0,1]$ for all $v \in dom(P_\varphi(a,u))$ and $P(a,u)(v) \leq P(a,u)(\varphi)$. Therefore $P_\varphi(a,u)(v) \in [0,1]$. □

As one can see this notion of updating is quite similar to Bayesian updating. In fact for many sentences it holds that

$$\mathbf{P}_a(\varphi|\psi) = q \text{ iff } [\psi]\mathbf{P}_a(\varphi) = q$$

Here notation is abused by adding conditional probabilities to the language, where conditional probability is defined as $\mathbf{P}_a(\varphi|\psi) = \frac{\mathbf{P}_a(\varphi\wedge\psi)}{\mathbf{P}_a(\psi)}$. The equivalence holds if the truth value of φ is not changed by learning that ψ. However the equivalence above does not hold in general. An example of this failure is in the case of an *unsuccessful update*. A successful update with φ will result in a state where the agents believe that φ. But for example when you get the information that 'you do not know that it is raining and it is raining', afterwards you will not believe that you do not know that it is raining[2]. In Gerbrandy (1999) this topic is discussed more extensively, including the *muddy children puzzle*, where an interesting example of unsuccessful updating occurs. The probabilistic version of this is quite similar. Suppose I flip a fair coin, such that you cannot see the outcome, but I can. Then I tell you that the probability you assign to heads is not zero and that the outcome is tails. After that update you do assign probability zero to the outcome being heads.

$$[\mathbf{P}_a(\text{heads}) > 0 \wedge \text{tails}]\mathbf{P}_a(\mathbf{P}_a(\text{heads}) > 0 \wedge \text{tails}) = 0$$

However

$$\mathbf{P}_a(\mathbf{P}_a(\text{heads}) > 0 \wedge \text{tails}|\mathbf{P}_a(\text{heads}) > 0 \wedge \text{tails}) = 1$$

This is due to the difference of perspective on information change in probability theory and dynamic epistemic logic as was explained in the introduction. Although after a public announcement it is common knowledge that φ is true at the time of the announcement, it need not be common knowledge that φ after the announcement, because φ may involve statements about information.

One remaining open question about DEL is to give a syntactic characterization of those sentences which may lead to an unsuccessful update. This is also unsolved for PDEL.

Although acquiring new information can now be modeled in a way that takes higher-order information into account, the language is not very sophisticated yet

[2]Some people argue that the occurrence of unsuccessful updates is due to the fact that these sentences are not properly labeled with time indices. In that case propositions can never change in truth value. In the context of dynamic logic however it seems more useful to have a notion of propositions that can change truth values.

with regard to how an update came about. When one models games, public announcements are made by the players using some sort of strategy. Perhaps they do not reveal all they know, or perhaps their actions depend on very complex protocols. For the kind of updates we are considering we have to assume that "the announcement that φ" came about by some process of which the result was either an announcement of φ or an announcement of $\neg\varphi$. For example when it is an answer to a question regarding φ. Consider the following example by Albers (2003, chapter 1). A fair die is thrown and one agent b can see the outcome, whereas a cannot. Now b can inform a about the outcome by saying either that the outcome is odd, or that the outcome is even, or that it is a multiple of three. Now if b truthfully states that the outcome is even, what is the probability that the outcome is 6? The answer depends on b's strategy or protocol. This kind of update cannot be dealt with in PDEL. For more sophisticated extensions of PDEL see section 6.7.

6.4 Reasoning about probability

In probability theory, inferences are often justified by making a model of the situation that is being investigated. Then the relevant propositions are analyzed in that model. In logic, inferences are usually shown to be valid by translating them into a formal language and showing that the conclusion can be deduced from the premises in a formal proof system. In this section I provide a way to make models of particular situations and a formal proof system for probabilistic dynamic epistemic logic. In section 6.6 an example of an application of each of these approaches is given.

6.4.1 Building a Model

Although I introduced probabilistic epistemic models in section 6.3, it is still not immediately clear how one can model a specific situation. In Halpern and Tuttle (1993) an approach for this is given. In this section I give a similar approach, which differs from the approach in Halpern and Tuttle (1993) in the sense that I introduce a new notion, namely purely probabilistic models. From that perspective one could say that in Halpern and Tuttle (1993) a special case of purely probabilistic models is considered, where only purely probabilistic models are considered that are S5 and connected. The interesting feature that these models have is that the agents have a *common prior*, which means that if they were to forget everything they have learned, then they would agree on all the probabilities. It is still an open question whether this class of models can be characterized by a sentence in the language of PDEL. The importance of having a common prior is that it is often assumed in game theory, see Aumann (1976).

As before, suppose an agent knows that a coin lands heads in one third of the cases or is fair, but she does not know which of these is the case. It is not easy to make a model of this at once. In this section I show how to construct a probabilistic epistemic model from two models: one for the *non-probabilistic* information (i.e. propositional and epistemic information) and another for the *probabilistic* information. It is often easier to think about these domains of information separately. The idea is to multiply an epistemic model with what I call a purely probabilistic model.

Definition 6.4 (Purely probabilistic models)
Let a nonempty set E and a finite set of agents $\mathcal{A}$ be given. A purely probabilistic epistemic model M is a triple $(\mathsf{W}, \mathsf{R}, \mathsf{P})$ such that:

- $\mathsf{W} \neq \emptyset$
- $\mathsf{R} : \mathcal{A} \rightarrow 2^{\mathsf{W} \times \mathsf{W}}$
- $\mathsf{P} : \mathsf{W} \rightarrow \{P \mid P \text{ is a probability functions with domain } E\}$ □

Thus a probability function is assigned to each world and the domain of all of these is E. I call these models *purely* probabilistic, because there are no propositional variables in them, but probability functions have a similar role. Nevertheless the accessibility relations will be interpreted epistemically.

Given an *epistemic* model M and a *purely probabilistic* model M we can make a *probabilistic epistemic* model $\mathfrak{M}$. Both models must be defined with respect to the same agents, and the set of possible worlds W of the epistemic model must be the domain of all probability spaces in the range of P (i.e. $E = W$). Worlds in the purely probabilistic model provide prior probability distributions over the set of worlds of the epistemic model. The probability an agent assigns to a set of worlds is its prior probability conditionalized on the agent's knowledge.

Definition 6.5 (Multiplication)
Given an epistemic model M and a purely probabilistic model M, such that $M = (W, R, V)$ and $\mathsf{M} = (\mathsf{W}, \mathsf{R}, \mathsf{P})$

$$M \otimes \mathsf{M} = \mathfrak{M} = (\mathfrak{W}, \mathfrak{R}, \mathfrak{V}, \mathfrak{P})$$

iff

$$\begin{array}{lcl}
\mathfrak{W} & = & W \times \mathsf{W} \\
\mathfrak{R}(a) & = & \{((w, \mathsf{w}), (v, \mathsf{v})) \mid wR(a)v \wedge \mathsf{w}\mathsf{R}(a)\mathsf{v}\} \\
\mathfrak{V}(p) & = & V(p) \times \mathsf{W} \\
dom(\mathfrak{P}(a, (w, \mathsf{w}))) & = & \{v \mid wR(a)v\} \times \{\mathsf{w}\} \\
\mathfrak{P}(a, (w, \mathsf{w}))(v, \mathsf{w}) & = & \dfrac{\mathsf{P}(\mathsf{w})(v)}{\sum_{(u,\mathsf{w}) \in dom(\mathfrak{P}(a,(w,\mathsf{w})))} \mathsf{P}(\mathsf{w})(u)}
\end{array}$$

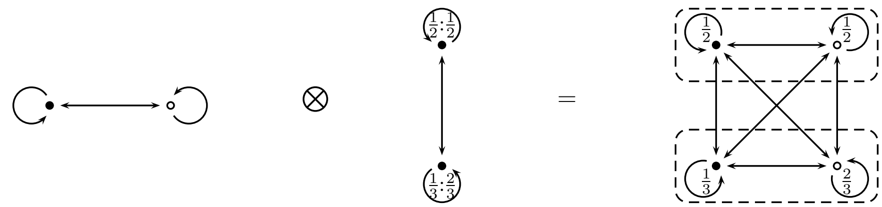


Figure 6.1: An example of multiplication. The epistemic model is on the left, the purely probabilistic model is in the middle, and the probabilistic epistemic model is on the right. In the probabilistic epistemic models the solid nodes indicate that the outcome is heads, and an open node indicates the outcome is tails. The dashed boxes indicate the domains of the probability functions.

The domain of the probability function that is assigned to an agent at a pair (w, w) contains those pairs (v, w) such that v is accessible to the agent in the epistemic model from w. So the domain is a probabilistic copy of the set of worlds accessible to the agent in the epistemic model. The probability assigned to a world by an agent is its conditional probability given that it is in the domain of the agent's probability function (disregarding the second element of the pair).

Now we can deal with the initial example: a coin is tossed and an agent a does not know the outcome. So she cannot distinguish worlds where the outcome is heads from worlds where it is tails. She knows the coin is fair or that it lands heads one third of the times, but she does not know which is the case. Hence she cannot distinguish worlds where the probability of heads is $\frac{1}{2}$ from worlds where it is $\frac{1}{3}$. I can make an *epistemic* model for a's information about the outcome and a *purely probabilistic* model for a's information about the coin. These two models and the result of multiplying these models are shown in figure 6.1. Now we can see that a sentence $\mathbf{P}_a(\varphi) \geq q$ should *not* be read as 'the probability a assigns to φ is greater than or equal to q,' because there need not be a *unique* probability q assigns to φ. In the example a cannot distinguish two probability distributions. $\mathbf{P}_a(\varphi) \geq q$ should be read as 'the probability a *should* assign to φ is greater than or equal to q, given the "actual" probability distribution over the worlds and given a's other information.' Hence we should be interested in sentences of the form $\Box_a(\mathbf{P}_a(\varphi) \geq q)$. Such a sentence holds iff a knows the probability she should assign to φ is greater than or equal to q.

There is one requirement the underlying models should meet for multiplication to work: the sets of worlds accessible to the agents should have non-zero probabilities. This ensures that the probability functions are well-defined, because it ensures that the set of worlds accessible to an agent is not empty and that no division by zero occurs.

6.4.2 Proof system, soundness and completeness

The proof system PDEL provided in this section is based on the proof system DEL in Gerbrandy (1999) for dynamic epistemic logic, and the proof system AX_{MEAS} in Fagin and Halpern (1994) for probabilistic epistemic logic. These two systems joined with the axioms Probability-Update 1 and Probability-Update 2 constitute the proof system for probabilistic dynamic epistemic logic.

Definition 6.6 (Proof System)
The proof system of probabilistic dynamic epistemic logic, PDEL, is provided in Figure 6.2.

Let us call the axioms and rules for propositional logic, epistemic logic and update logic without the Probability-Update axioms DEL, and the axioms and rules for propositional logic, epistemic logic, linear inequalities and probability logic PEL. In Gerbrandy and Groeneveld (1997) the soundness and completeness of DEL is proved, although it is proved with respect to non-well-founded objects, the correspondence between these and epistemic models implies it is also sound and complete for epistemic models. In Fagin and Halpern (1994) the soundness and completeness of PEL is proved. Although it is proved for a more general class of models than the models of definition 5.9, the models of definition 5.9 form a subclass, therefore PEL is still sound. Furthermore, in the completeness proof, the countermodels are probabilistic epistemic models in the sense of definition 5.9 (their system has the finite model property). Therefore this logic is also complete for these models.

The axiom Probability-Update 1 clarifies the relationship between conditional probability and the notion of updating probabilities in this chapter. The relationship is best captured by the following equivalence, which was pointed out to me by Johan van Benthem:

$$[\psi](\mathbf{P}_a(\varphi) = q) \qquad \text{iff} \qquad \mathbf{P}_a([\psi]\varphi \mid \psi) = q$$

Note that it is not a normal modal logic, because we do not have universal substitution. This is due to the existence of unsuccessful updates. For example, $\vdash [p]\Box_a p$ is a theorem, but $\vdash [\neg\Box_a p \wedge p]\Box_a(\neg\Box_a p \wedge p)$ is not, although it is a substitution instance. There are more principles in dynamic epistemic logics which are valid but not derivable schematically (see van Benthem (2002b)).

Theorem 6.1 (Soundness)
If $\vdash \varphi$ then $\models \varphi$. □

Proof The soundness of the axioms of DEL and PEL I do not prove. The proof of their soundness can be found in Gerbrandy (1999) or Gerbrandy and Groeneveld (1997), and Fagin and Halpern (1994).

Propositional Logic	
PC	$\vdash \varphi$ where φ is an instance of a propositional tautology
Epistemic Logic	
$\Box_a$-distribution	$\vdash \Box_a(\varphi \to \psi) \to (\Box_a\varphi \to \Box_a\psi)$
$\Box_a$-necessitation	From $\vdash \varphi$, infer $\vdash \Box_a\varphi$
Update Logic	
$[\varphi]$-distribution	$\vdash [\varphi](\psi \to \chi) \to ([\varphi]\psi \to [\varphi]\chi)$
Functionality	$\vdash \neg[\varphi]\psi \leftrightarrow [\varphi]\neg\psi$
Atomic Permanence	$\vdash p \leftrightarrow [\varphi]p$
Knowledge-Update	$\vdash [\varphi]\Box_a\psi \leftrightarrow \Box_a(\varphi \to [\varphi]\psi)$
Probability-Update 1	$\vdash \mathbf{P}_a(\varphi) > 0 \to (([\varphi]\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q) \leftrightarrow (\sum_{i=1}^n q_i\mathbf{P}_a(\varphi \wedge [\varphi]\varphi_i) \geq q\mathbf{P}_a(\varphi)))$
Probability-Update 2	$\vdash \mathbf{P}_a(\varphi) = 0 \to (([\varphi]\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q) \leftrightarrow (\sum_{i=1}^n q_i\mathbf{P}_a([\varphi]\varphi_i) \geq q))$
$[\varphi]$-necessitation	From $\vdash \psi$, infer $\vdash [\varphi]\psi$
Linear Inequalities	
0 terms	$\vdash \sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q \leftrightarrow (\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i)) + 0\mathbf{P}_a(\varphi_{k+1}) \geq q$
Permutation	$\vdash \sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q \to \sum_{i=1}^n q_{j_i}\mathbf{P}_a(\varphi_{j_i}) \geq q$ where $j_1, \ldots, j_k$ is a permutation of $1 \ldots k$
Addition	$\vdash \sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q \wedge \sum_{i=1}^n q'_i\mathbf{P}_a(\varphi_i) \geq q' \to \sum_{i=1}^n (q_i + q'_i)\mathbf{P}_a(\varphi_i) \geq (q + q')$
Multiplication	$\vdash (\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q) \leftrightarrow (\sum_{i=1}^n dq_i\mathbf{P}_a(\varphi_i) \geq db)$ where $d > 0$
Dichotomy	$\vdash (t \geq q) \vee (t \leq q)$
Monotonicity	$\vdash (t \geq q) \to (t > q')$ where $q > q'$
Probability Logic	
Nonnegativity	$\vdash \mathbf{P}_a(\varphi) \geq 0$
Probability of truth	$\vdash \mathbf{P}_a(\top) = 1$
Additivity	$\vdash \mathbf{P}_a(\varphi \wedge \psi) + \mathbf{P}_a(\varphi \wedge \neg\psi) = \mathbf{P}_a(\varphi)$
Equivalence	From $\vdash \varphi \leftrightarrow \psi$, infer $\vdash \mathbf{P}_a(\varphi) = \mathbf{P}_a(\psi)$

Figure 6.2: The proof system PDEL for probabilistic dynamic epistemic logic

For Probability-Update 1, first of all note that, if $(M, w) \models \mathbf{P}_a(\varphi) > 0$:

$$\begin{aligned}&\{v \mid (M, v) \models [\varphi]\psi \wedge \varphi \text{ and } v \in dom(P(a, w))\}\\ &=\\ &\{v_\varphi \mid (M_\varphi, v_\varphi) \models \psi \text{ and } v_\varphi \in dom(P_\varphi(a, w_\varphi))\}\end{aligned} \tag{6.1}$$

Suppose $(M, w) \models \mathbf{P}_a(\varphi) > 0$. Now the following equivalences hold:
$(M, w) \models [\varphi]\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q)$
$\equiv$ {truth definition}
$(M_\varphi, w_\varphi) \models \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q)$
$\equiv$ {truth definition}
$\sum_{i=1}^{k} q_i P_\varphi(a, w_\varphi)(\varphi_i) \geq q$
$\equiv$ {truth definition}
$\sum_{i=1}^{k} q_i P_\varphi(a, w_\varphi)(\{v_\varphi \mid (M_\varphi, v_\varphi) \models \varphi_i \text{ and } v_\varphi \in dom(P_\varphi(w_\varphi))\}) \geq q$
$\equiv$ {By (6.1)}
$\sum_{i=1}^{k} q_i P_\varphi(a, w_\varphi)(\{v \mid (M, v) \models [\varphi]\psi \wedge \varphi \text{ and } v \in dom(P(a, w))\}) \geq q$
{the definition of P_φ, and $(M, w) \models \mathbf{P}_a(\varphi) > 0$}

$$\sum_{i=1}^{k} q_i \frac{P(a, w)([\varphi]\varphi_i \wedge \varphi)}{P(a, w)(\varphi)} \geq q$$

$\equiv$ {algebra}
$\sum_{i=1}^{k} q_i P(a, w)([\varphi]\varphi_i \wedge \varphi) \geq q P(a, w)(\varphi)$
$\equiv$ {truth definition}
$(M, w) \models \sum_{i=1}^{n} q_i \mathbf{P}_a([\varphi]\varphi_i \wedge \varphi) \geq q \mathbf{P}_a(\varphi))$

The soundness of probability-update 2 is immediate from the definition of update, because if φ has probability zero, nothing happens to the domain of the probability function after updating with φ.

□

To prove completeness I provide a translation of the sentences of probabilistic dynamic epistemic logic to the sentences of probabilistic epistemic logic. Given that PEL is complete for probabilistic epistemic logic, it then suffices to show that a sentence is provably equivalent in PDEL to its translation.

Definition 6.7 (Translation from $\mathscr{L}_{\mathcal{PA}}^{\mathbf{P}[\cdot]}$ to $\mathscr{L}_{\mathcal{PA}}^{\mathbf{P}}$)
The translation $t : \mathscr{L}_{\mathcal{PA}}^{\mathbf{P}[\cdot]} \rightarrow \mathscr{L}_{\mathcal{PA}}^{\mathbf{P}}$ is defined as follows:

1. $t(p) = p$
2. $t(\neg\varphi) = \neg t(\varphi)$
3. $t(\varphi \wedge \psi) = t(\varphi) \wedge t(\psi)$
4. $t(\Box_a \varphi) = \Box_a t(\varphi)$

5. $t(\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q) = (\sum_{i=1}^{n} q_i \mathbf{P}_a(t(\varphi_1)) \geq q)$
6. $t([\varphi]p) = p$
7. $t([\varphi]\neg\psi) = \neg t([\varphi]\psi)$
8. $t([\varphi](\psi \wedge \chi)) = t([\varphi]\psi) \wedge t([\varphi]\chi)$
9. $t([\varphi]\Box_a\psi) = \Box_a(t(\varphi) \rightarrow t([\varphi]\psi))$
10. $t([\varphi](\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q)) =$
11. $t([\varphi][\psi]\chi = t([\varphi]t([\psi]\chi))$

$$(\mathbf{P}_a(t(\varphi)) > 0 \wedge (\sum_{i=1}^{n} q_i \mathbf{P}_a(t(\varphi) \wedge t([\varphi]\varphi_i)) \geq q\mathbf{P}_a(t(\varphi))))$$
$$\vee$$
$$(\mathbf{P}_a(t(\varphi)) = 0 \wedge \sum_{i=1}^{n} q_i \mathbf{P}_a(t([\varphi]\varphi_i)) \geq q)$$

Note that although the update operator has an infinitary character (it has effects for the entire model), when evaluating a sentence the effect only needs to be given for the finite intention depth (the number of stacked modal operators). The proofs about this translation uses the following complexity measure.

Definition 6.8 (Complexity)
The complexity of sentences is defined as follows:

1. $c(p) = 1$
2. $c(\neg\varphi) = 1 + c(\varphi)$
3. $c(\varphi \wedge \psi) = 1 + \max(c(\varphi), c(\psi))$
4. $c(\Box_a\varphi) = 1 + c(\varphi)$
5. $c(\sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q) = 1 + \max_{1 \leq i \leq k} c(\varphi_i)$
6. $c([\varphi]\psi) = c(\varphi) + c(\psi)$ □

Lemma 6.2
For every sentence φ of PDEL, the translation of that sentence $t(\varphi)$ is a sentence in probabilistic epistemic logic to which it is provably equivalent in PDEL.

Proof By induction on the complexity of φ

base Trivial

induction hypothesis For every sentence φ of PDEL such that $c(\varphi) \leq n$, the translation of that sentence $t(\varphi)$ is a sentence in probabilistic epistemic logic and φ and $t(\varphi)$ are provably equivalent in PDEL.

induction step Suppose $c(\varphi) = n+1$. The cases where φ is a negation, conjunction, or an epistemic, or probabilistic sentence are fairly straightforward.

1. Suppose φ is a sentence of the form $[\psi]p$. Then $t([\psi]p) = p$. Therefore $t([\psi]p)$ is a sentence of probabilistic epistemic logic. It follows from the axiom of atomic permanence that these are equivalent.

2. Suppose φ is a sentence of the form $[\psi]\neg\chi$. Then $t([\psi]\neg\chi) = \neg t([\psi]\chi)$. By the induction hypothesis $t([\psi]\chi)$ is a sentence of probabilistic epistemic logic. Therefore $t([\psi]\neg\chi)$ is a sentence of probabilistic epistemic logic. By the induction hypothesis $t([\psi]\chi)$ is equivalent to $[\psi]\chi$. Therefore $\neg t([\psi]\chi)$ is equivalent to $\neg[\psi]\chi$, which by the functionality axiom is equivalent to $[\psi]\neg\chi$.

3. Suppose φ is a sentence of the form $[\psi](\chi \wedge \xi)$. Then $t([\psi](\chi \wedge \xi)) = t([\psi]\chi) \wedge t([\psi]\xi)$. By the induction hypothesis $t([\psi]\chi)$ and $t([\psi]\xi)$ are sentences of probabilistic epistemic logic. Therefore $t([\psi](\chi \wedge \xi))$ is a sentence of probabilistic epistemic logic. By the induction hypothesis $t([\psi]\chi)$ and $t([\psi]\xi)$ are equivalent to $[\psi]\chi$ and $[\psi]\xi$. Some modal reasoning suffices to see that $[\psi](\chi \wedge \xi)$ is equivalent to $[\psi]\chi \wedge [\psi]\xi$.

4. Suppose φ is a sentence of the form $[\psi]\Box_a\chi$. Then $t([\psi]\Box_a\chi) = \Box_a(t(\psi) \rightarrow t([\psi]\chi))$. By the induction hypothesis $t(\psi)$ and $t([\psi]\chi)$ are sentences of probabilistic epistemic logic. Therefore $t([\psi]\Box_a\chi)$ is a sentence of probabilistic epistemic logic. By the induction hypothesis $t(\psi)$ and $t([\psi]\chi)$ are equivalent to ψ and $[\psi]\chi$ respectively. By the knowledge update axiom we have that $[\psi]\Box_a\chi$ is equivalent to $\Box_a(\psi \rightarrow [\psi]\chi)$

5. Suppose φ is a sentence of the form $[\psi](\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q)$. Then $t([\psi](\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q)) =$

   $$\begin{array}{l}(\mathbf{P}_a(t(\psi)) > 0 \wedge (\sum_{i=1}^n q_i\mathbf{P}_a(t(\psi) \wedge t([\psi]\varphi_i)) \geq q\mathbf{P}_a(t(\psi)))) \\ \vee \\ (\mathbf{P}_a(t(\psi)) = 0 \wedge \sum_{i=1}^n q_i\mathbf{P}_a(t([\psi]\varphi_i)) \geq q))\end{array}$$

   By the induction hypothesis $t(\psi)$ and $t([\psi]\varphi_i)$ are sentences in probabilistic epistemic logic. Therefore $t([\psi]\sum_{i=1}^n q_i\mathbf{P}_a(\varphi_i) \geq q))$ is a sentence of probabilistic epistemic logic. By the induction hypothesis $t(\psi)$ and all $t([\psi]\varphi_i)$ are equivalent to ψ and $[\psi]\varphi_i$ respectively. It is deducible that $\mathbf{P}_a(\varphi) > 0 \vee \mathbf{P}_a(\varphi) = 0$. By the probability-update axioms we get the desired result.

6. Suppose φ is a sentence of the form $[\varphi][\psi]\chi$. Then $t([\varphi][\psi]\chi) = t([\varphi]t([\psi]\chi))$. By applying the induction hypothesis to $t([\psi]\chi)$ we get a formula with a lower complexity than $[\psi]\chi$, then we can apply the induction hypothesis to $t([\varphi]t([\psi]\chi))$ to get the desired result.

□

Theorem 6.2 (Completeness)
If $\models_{\mathsf{PDEL}} \varphi$, then $\vdash_{\mathsf{PDEL}} \varphi$. □

Proof Suppose that $\models_{\mathsf{PDEL}} \varphi$. From lemma 6.2 and soundness we get that $\models_{\mathsf{PDEL}} t(\varphi)$. Therefore $\models_{\mathsf{PEL}} t(\varphi)$, and by completeness of PEL we get $\vdash_{\mathsf{PEL}} t(\varphi)$ Now we can conclude from lemma 6.2 that $\vdash_{\mathsf{PDEL}} \varphi$. □

Corollary 6.1
The language of probabilistic dynamic epistemic logic is just as expressive as the language of probabilistic epistemic logic. □

Corollary 6.2
The validity problem for probabilistic dynamic epistemic logic is decidable. □

Proof This follows directly from the decidability result in Fagin and Halpern (1994). □

As to complexity, the validity problem for probabilistic epistemic logic is complete for polynomial space. However the translation from definition 6.7 is exponential in space in the depth of probabilistic operators after an update, i.e. sentences of the form $[\varphi](\mathbf{P}_a(\mathbf{P}_a(\ldots)))$. We can of course conclude that polynomial space is a lower bound on complexity and exponential space is an upper bound on complexity.

Strong completeness fails, because probabilistic epistemic logic is not compact.

Lemma 6.3 (Non-compactness)
Probabilistic epistemic logic is not compact.

Proof Consider the following set of sentences

$$\Gamma = \{\mathbf{P}_a(p) > 0\} \cup \{\mathbf{P}_a(p) \leq 2^{-i} \mid i \in \mathbb{N}\}$$

Every finite subset of Γ has a model, but the whole set Γ does not. □

This lemma is based on Keisler (1985), where a similar result is proved for a probability logic for statistical probabilities. This 'Zeno'-example applies equally well to other probability logics. In Keisler (1985) the main focus is on infinitary probability logics, where compactness can be regained in some cases.

6.5 Bisimulation for probabilistic dynamic epistemic logic

Bisimulation is a useful notion in modal logic. It generally holds that if two structures are bisimilar, then they are behaviorally indistinguishable. In the case of probabilistic epistemic models, behaviorally indistinguishable means satisfying the same sentences. A well-known result in modal logic is that if two pointed models are bisimilar, then they satisfy the same sentences (see for example Blackburn, de Rijke, and Venema (2001) for a textbook explanation of this notion.) In this section I show that such a result holds for probabilistic dynamic epistemic logic as well.

Definition 6.9 (Bisimulation)
I use the following abbreviations.

$$\begin{array}{lcl}\mathsf{forth}(E,E') & := & \forall x \in E \exists y \in E'(xBy) \\ \mathsf{back}(E,E') & := & \forall y \in E' \exists x \in E(xBy)\end{array}$$

Let two probabilistic epistemic models M and M' be given. A relation $B \subseteq W \times W'$ is a bisimulation iff for all $w \in W$ and $w' \in W'$, if wBw', then for all $n \in \mathcal{A}$ the following hold:

atoms $w \in V(p)$ iff $w' \in V'(p))$ for every $p \in \mathcal{P}$

forth $\mathsf{forth}(\{v \mid wR(a)v\}, \{v' \mid w'R'(a)v'\})$

back $\mathsf{back}(\{v \mid wR(a)v\}, \{v' \mid w'R'(a)v'\})$

pforth For every $E \subseteq dom(P(a,w))$ there is an $E' \subseteq dom(P'(a,w'))$ such that

$$P(a,w)(E) \leq P'(a,w')(E') \text{ and } \mathsf{back}(E,E')$$

pback For every $E' \subseteq dom(P'(a,w'))$ there is an $E \subseteq dom(P(a,w))$ such that

$$P'(a,w')(E') \leq P(a,w)(E) \text{ and } \mathsf{forth}(E,E')$$

I write $(M,w) \leftrightarrows (M',w')$, if there is a bisimulation between M and M' linking w and w'. □

Atoms, **forth** and **back** are the usual conditions for bisimulation. I added **pforth** and **pback** to accommodate probabilistic sentences. For those readers familiar with probability theory, this definition can easily be extended to the more general notion of probabilistic epistemic models given in Fagin and Halpern (1994) with probability spaces, where one takes the inner measure instead of the probability function in **pforth** and **pback**. The theorem below also holds for these models.

Theorem 6.3
For all models (M, w) and (M', w') and for all sentences φ, if $(M, w) \leftrightarroweq (M', w')$, then $(M, w) \models \varphi$ iff $(M', w') \models \varphi$ $\square$

Proof By induction on φ. Suppose $(M, w) \leftrightarroweq (M', w')$. The base case and cases for conjunction, negation and individual epistemic operators $\Box_a$ are straightforward. By lemma 6.2, we get the case for updates for free.

Suppose uBu' and $(M, u) \models \sum_{i=1}^{n} q_i \mathbf{P}_a(\varphi_i) \geq q$. Let

$$E_i = \{v \in dom(P(a, u)) \mid (M, v) \models \varphi_i\}$$

and

$$E'_i = \{v' \in dom(P'(a, u')) \mid (M', v') \models \varphi_i\}$$

If we show that $P(a, u)(E_i) \leq P'(a, u')(E'_i)$ we are done. From uBu' and **pforth** it follows that there is an $S' \subseteq dom(P'(a, w'))$ such that

$$P(a, w)(E_i) \leq P'(a, w')(S') \text{ and } \mathsf{back}(E_i, S')$$

The induction hypothesis together with $\mathsf{back}(E_i, S')$ imply that $(M', v') \models \varphi_i$ for every $v' \in S'$. Therefore $S' \subseteq E'_i$ and therefore $P'(a, u')(S') \leq P'(a, u')(E'_i)$. Now we conclude that

$$P(a, u)(E_i) \leq P'(a, w')(S') \leq P'(a, u')(E'_i)$$

The case for right to left is analogous. Which gives as an additional result that $P(a, u)(E_i) = P'(a, u')(E'_i)$.

Therefore for all models (M, w) and (M', w'), if $(M, w) \leftrightarroweq (M', w')$, then for all sentences φ: $(M, w) \models \varphi$ iff $(M', w') \models \varphi$ $\square$

The converse of theorem 6.3 also holds when the models are finite or when one uses an infinitary language which allows conjunctions over arbitrary sets of sentences.

The notion of bisimulation presented in this chapter can also be applied to probability spaces, which can be seen as special cases of probabilistic epistemic Kripke models, and therefore it is also interesting for probability theory, to see whether two models of the same experiment are equivalent. It would be worthwhile to investigate the mathematics of this further.

There are richer languages for reasoning about probability which are able to distinguish bisimilar models, such as the language of SPL (see chapter 5). Dependent upon the language which is used in reasoning about probability, one might wonder whether there is information being modeled which is not needed. This leads to the question whether one can define minimal models. In modal logic one can define a minimal model with respect to an arbitrary Kripke model by identifying all bisimilar worlds. The result for modal logic seems to be folklore. This can also be done for probabilistic epistemic models.

Definition 6.10 (Minimal models)
Let a probabilistic epistemic model $M = (W, R, V, P)$ be given. The minimal model associated with M is the model $M' = (W', R', V', P')$, where:

- $W' = \{E \subseteq W \mid \text{for all } w, v \in E : (M, w) \leftrightarrows (M, v)\}$
- $R'(a) = \{(E, E') \in (W' \times W') \mid \text{there is a } w \in E \text{ and a } v \in E' \text{ such that } wR(a)v\}$
- $V'(p) = \{E \mid \text{there is a } w \in E \text{ such that } w \in V(p)\}$
- $dom(P'(a, E)) = \{E' \in W' \mid \text{there is a } w \in E \text{ and a } v \in E' \text{ such that } v \in dom(P(a, w))\}$
- $P'(a, E)(E') = \sup\{q \in \mathbb{R} \mid \text{there is a } w \in E \text{ such that } q = P(a, w)(E' \cap dom(P(a, w)))\}$

where in the last clause $E' \in dom(P'(a, E))$. □

Lemma 6.4
A minimal model is an probabilistic epistemic model. □

Proof Let M' be a minimal model. The only difficulty lies in showing that P' is a probability function. We have to show that

$$\forall a \in \mathcal{A} \forall E \in W' \sum_{E' \in dom(P'(a,E))W'} P'(a, E)(E') = 1$$

Take an arbitrary agent $a \in \mathcal{A}$, an arbitrary world $E \in W'$, and an arbitrary $E' \in dom(P'(a, E))$. The probability $P'(a, E)(E')$ is defined as

$$\sup\{q \in \mathbb{R} \mid \text{there is a } w \in E \text{ such that } q = P(a, w)(E' \cap dom(P(a, w)))\}$$

Suppose there are two worlds w and v in E. Suppose moreover that $P(a, w)(E' \cap dom(P(a, w))) \neq P(a, v)(E' \cap dom(P(a, v)))\}$. Assume without loss of generality that the former is greater than the latter. Because w and v are both in E they are bisimilar. Let B be a bisimulation that establishes $(M, w) \leftrightarrows (M, v)$. It follows from **pforth** that there is a set $F' \subseteq dom(P(a, v)$ such that

$$P(a, w)(E' \cap dom(P(a, w)))(a, w) \leq P(a, v)(F')$$

and $\mathsf{back}(E' \cap dom(P(a, w)), F')$. From $\mathsf{back}(E' \cap dom(P(a, w)), F')$ it follows that $F' \subseteq E' \cap dom(P(a, v))$, since all worlds bisimilar to world in E' are also in E'. But then it cannot be the case that $P(a, w)(E' \cap dom(P(a, w))) > P(a, v)(E' \cap dom(P(a, v)))\}$. Therefore $\{q \in \mathbb{R} \mid \text{there is a } w \in E \text{ such that } q = P(a, w)(E' \cap dom(P(a, w)))\}$ is a singleton set. Therefore $\sum_{E' \in dom(P'(a,E))W'} P'(a, E)(E') = 1$.
□

Lemma 6.5
Every model M is bisimilar to the minimal model M' associated with it. $\Box$

Proof Let $M = (W, R, V, P)$ and $M' = (W', R', V', P')$ as in definition 6.10. Now we will show that the $\in$ relation on $W \times W'$ is a bisimulation. The case for **atoms**, **forth**, and **back** are straightforward.

For the case for **pforth** assume that $w \in E$, where $E \in W'$. Suppose $S \subseteq \mathit{dom}(P(a, w))$. Now we have to show that there is a subset $\mathbb{S}$ of the domain of $P'(a, E)$ such that the probability assigned to it is greater or equal to the probability assigned to S and $\mathsf{back}(S, \mathbb{S})$. Let $\mathbb{S} =$

$$\{E' \in W' \mid \text{there is a } v \in S \text{ such that } v \in E'\}$$

From the definition of $\mathit{dom}(P'(a, E))$ it follows that $\mathbb{S} \subseteq \mathit{dom}(P'(a, E))$. It is also easily seen that $\mathsf{back}(S, \mathbb{S})$. From the definition of $P'(a, E)$ it follows that $P(a, w)(E' \cap \mathit{dom}(P(a, w))) \leq P'(a, E)(E')$, for every $E' \in \mathbb{S}$. Therefore $P(a, w)(S) \leq P'(a, E)(\mathbb{S})$.

For the case of **pback** assume that $w \in E$, where $E \in W'$. Suppose $\mathbb{S} \subseteq \mathit{dom}(P'(a, E))$. Now we have to show that there is a subset S_w of the domain of $P(a, w)$ such that the probability assigned to it is greater or equal to the probability assigned to $\mathbb{S}$ and $\mathsf{forth}(S, \mathbb{S})$. Let $S_w =$

$$\{v \mid \text{there is an } E' \in \mathbb{S} \text{ such that } v \in E' \text{ and } v \in \mathit{dom}(P(a, w))\}$$

It is obviously the case that $S_w \subseteq \mathit{dom}(P(a, w))$ and $\mathsf{forth}(S_w, \mathbb{S})$. Suppose, towards a contradiction that $P'(a, E)(\mathbb{S}) > P(a, w)(S_w)$. Therefore there is a $u \in E$ such that for the set $S_u =$

$$\{v \mid \text{there is an } E' \in \mathbb{S} \text{ such that } v \in E' \text{ and } v \in \mathit{dom}(P(a, u))\}$$

it is the case that $P(a, u)(S_u) > P(a, w)(S_w)$. But because both w and u are in E they must be bisimilar, given that $E \in W'$. Let B be a bisimulation that establishes that $(M, w) \leftrightarrows (M, u)$. From **pback** it follows that for S_u there is a set $S \subseteq \mathit{dom}(P(a, w))$ such that $P(a, u)(S_u) \leq P(a, w)(S)$ and $\mathsf{forth}(S, S_u)$. Therefore for every world v in S there is a world v' in S_u bisimilar to it. Therefore v and v' end up in the same world in $E' \in \mathbb{S}$. Since $S \subset \mathit{dom}(P(a, w))$, it is the case that $S \subseteq S_w$. But this leads to a contradiction, because it now follows that $P(a, w)(S) \leq P(a, w)(S_w)$. But we had assumed that $P(a, w)(S_w) < P(a, u)(S_u)$, and had concluded that $P(a, u)(S_u) \leq P(a, w)(S)$.

Therefore $M \leftrightarrows M'$. $\Box$

In the literature on probabilistic transition systems, notions of probabilistic bisimulation have also been put forward: notably those by Larsen and Skou (1991), who introduce a notion of bisimulation for discrete systems, and de Vink and

Rutten (1999), which is a generalization of Larsen en Skou's approach to general probabilistic transition systems. There are some small differences between these notions of probabilistic bisimulation and the notion presented in this chapter, and the question whether the notions coincide, or that one is more general than the other requires further investigation. However as far as I know, the result that bisimilarity of two probabilistic epistemic models implies that they have the same probabilistic dynamic epistemic theory is new, as well as the result about minimal models.

6.6 Example

In this section an example will be given that applies the theory presented in this chapter, which links up with the approach of building models of section 6.4.1. It shows that when higher-order information is involved it can be useful to formalize the inferences that are involved in the example with the language introduced in this chapter. In chapter 7 more examples can be found.

Let us look at the following game, which is based on an example by Van Rooy (2003), which is in turn based on an example by Hirshleifer and Riley (1992, p. 220). Suppose there are two players, a and b. A coin is tossed and the players have to guess the outcome. Player b guesses first, and after hearing player b's guess player a guesses the outcome. If they guess the same outcome both players receive a payoff of 12 euros regardless of the outcome, otherwise, when the guesses differ, the player who guessed the outcome correctly receives 30 euros.

Suppose that it is not known to player a whether the coin is fair, or whether the coin lands heads with probability one third, but player b does know, and this is common knowledge. Consequently player a does not really know which game she is playing: it is a game of incomplete information (see Binmore (1992, chapter 11)). One can construct a probabilistic epistemic model for this situation in the way described in section 6.4.1 (see Figure 6.3 below) . A game theoretical analysis of this situation tells us that when the coin is fair and player a is risk neutral, the best strategy for a is to guess the opposite of player b's guess. The expected payoff is 15 euros. But when the coin is not fair and player a is risk neutral, the best strategy for player a is to guess the same outcome as player b. This strategy guarantees an outcome of 12 euros. If player a would guess differently, the expected outcome is only 10 euros, because player b's best strategy is to guess the outcome that is most likely (in this case tails). In that case the probability that player a wins 30 euros is $\frac{1}{3}$.

Now suppose a public announcement is made as to which probabilities player b assigns to the outcome tails. Afterwards player a will know what strategy to follow. Note that the announcement does not say anything about what the actual outcome is. Player a only learns about the probabilities player b assigns to the outcomes. This higher-order information determines what the best strategy is.

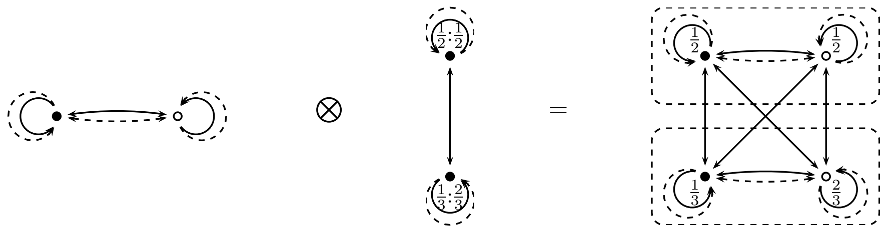


Figure 6.3: The construction of a probabilistic epistemic model for the situation where player a does not know whether the coin is fair or not, but player b does. In the (probabilistic) epistemic models the solid nodes indicate that the outcome is heads, and an open node indicates the outcome is tails. The solid lines represent the accessibility relation of player a, the dashed lines represent the accessibility relation of player b. The dashed boxes indicate the domains of the probability functions.

6.7 Conclusion and further research

In this chapter I presented a probabilistic dynamic epistemic logic, which can be used to reason about probability, information, and information change. The difference between information change as it is modeled in this logic and as it is modeled in probability theory is that higher-order information is taken into account. Besides semantics I have provided a method to build models and a sound and complete proof system. Moreover the notion of bisimulation has also been defined for this logic. It can be applied to game situations such as card games with public announcements, but also the Monty Hall Dilemma (see chapter 7).

The principal advantage PDEL has with respect to probability theory is that it can be used to formalize inferences into a formal language, such that standard logical tools can be used to see whether it is a good inference. Therefore it is very suitable to model reasoning. In probability theory probabilities are assigned to sets of worlds. These sets appear in the 'language' of probability theory, which means one is always working with a specific model. This makes it difficult to assess whether inferences hold in all models, which is exactly what logic provides. Moreover, by having a language such as PDEL one can explicitly deal with higher-order information.

As a contribution to the study of information change, PDEL provides a novel approach to probabilistic updating. Updating in probability theory and PDEL are very similar. In probability theory new information is usually represented as a set of possible worlds. One learns that the actual world is an element of that set. In PDEL new information is represented as a sentence. By definition 6.2 every sentence is associated with a set of possible worlds. One learns that the actual world is an element of that set, just as in probability theory. But by having this linguistic component that can express higher-order information, we can take into account that the truth value of sentences can change due to an update. Updating with the same sentence twice may yield different results than updating once. Updating with one sentence and then another may be different from updating the other way around. These phenomena are not taken into account in probability theory, where receiving the same information twice is always the same as receiving it once, and the order in which information is received does not matter.

Now let us turn our attention to further research. Publicly learning a sentence is not the only way one can acquire new information. There are changes in information that cannot be modeled with PDEL. In the future we want to model game actions such as: one player showing another player a card, while a third player can see this is going on, but cannot see which card is being shown. To be able to handle these kinds of actions, we need to bring more of dynamic epistemic logic into PDEL, by making the dynamic operators more program-like (in the style of PDL (see chapter 3)). An extension with test, non-deterministic choice, sequential composition, and subgroup updates does not involve many difficulties. Subgroup updates are updates where some agents get new information, whereas the other agents do not get that information. The same proof technique, i.e. by a translation, for completeness applies.

There are more phenomena we would like to capture such as common knowledge, because common knowledge plays an important role in many game situations. This poses some problems for the proof system. In Baltag, Moss, and Solecki (1999) a complete proof system for dynamic epistemic logic with common knowledge is provided, which gives good hopes that it can also be added to PDEL.

Another direction for further research is to develop a logic along the lines of Baltag (2002), where epistemic actions are viewed as epistemic action models that can be multiplied with epistemic models, yielding the result of executing the action. All these models could be made into probabilistic models. A step in this direction is made in van Benthem (2002a). However the problem mentioned at the end of section 6.3 cannot be solved yet. This problems arises when the preconditions of the actions do not form a partition of the set of possible worlds. In order to solve this problem one would have to be able to model strategies and protocols.

In probability theory there are other, more complex ways of incorporating new information, such as Jeffrey's rule of conditioning, see Jeffrey (1983), Dempster's rule of combination, see Dempster (1967), and cross entropy, see Kullback and Leibler (1951). All these ideas were born out of different kinds of dissatisfaction with conditional probability as a model for incorporating new information. It would also be interesting to investigate to what kinds of dynamic epistemic updates these kinds of information change correspond.

This chapter is a first step in combining epistemic logic with probability theory and there are many more steps to make.

Chapter 7

The Monty Hall dilemma

7.1 Introduction

The Monty Hall Dilemma is a puzzle that often leads to furious discussions. The puzzle received worldwide attention when it was discussed in Marilyn vos Savant's column in *Parade Magazine* (Vos Savant (1990)). In her column 'Ask Marilyn' she answers questions sent in by the readers.

> Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?
>
> Craig. F. Whitaker
> Columbia, MD

The Monty Hall Dilemma got its name from the American game show host Monty Hall, who figured in the problem in an article by Selvin (1975). Other versions of this puzzle circulated at least as early as the sixties, see Mosteller (1965, p.4, for example). Answering this question, Vos Savant argued that it is to your advantage to switch. If you switch you lose in one third of the cases and win in two thirds of the cases.

The claim made by Vos Savant, who is listed in the Guinness Book of World Records for the highest IQ, can be argued as follows. If you switch you get a goat in one third of the cases and win the car in two third of the cases. This could be argued as follows. Suppose you initially pick the door with the car, then you should not switch. This happens in one third of the cases. Suppose on the other hand you initially pick a door that contains a goat, which happens in two third of the cases. Monty Hall cannot open the door with the car and he cannot open the door you picked. He has to open the other door with a goat. So, if you pick a door with a goat, Monty Hall only has one option. After he opens that door, the remaining unopened door you did not pick must contain the car. Therefore, if you initially pick a door with a goat, switching will guarantee that you win the car. You pick such a door in two third of the cases. Hence by switching you lose in one third of the cases and you win in two third of the cases.

Many people did not agree with this solution. They argued the chances of winning do not increase when you switch. Among them are some considered to be experts in the field of probability. Three Ph.D.'s wrote letters explaining that Vos Savant was wrong. The mathematician Paul Erdös also did not want to believe switching is to your advantage. The discussion made its way to the Netherlands after Rob van den Berg reported the discussion in the newspaper NRC-Handelsblad (May 18th, 1995). The response to his article was overwhelming. People called to say they could not sleep because they were thinking about the puzzle and demanded an explanation, many e-mails were sent and the newspaper received over eighty letters. Some of these letters were published (the 1st, 8th, and 15th of June). I found the letter by H. von Saher to be the most surprising:

> I will show with an analogous example that her thesis is incorrect. Marilyn vos Savant wins a quiz (naturally), at the back of the stage a wall is placed with 100 doors. She takes her stand in front of door 1; in that case she has a 1% chance of standing in front of the right door and there is a 99% chance that the prize is behind one of all the other doors. Then the quiz master comes along; he opens all doors from 2 to 99, so 98 doors in total. Behind none of these doors is the prize. It must be behind door 1 or 100. According to Marilyn vos Savant the whole 99% chance passes to door 100. It is evident that this is absurd. Here too another situation has arisen with only two alternatives and therefore equal chances for each of the remaining doors. (NRC June 1st, 1995, my translation)

Although it seems that some of the crucial information eludes Von Saher, namely that the quiz master only opens a door if he knows it does not contain a prize, he apparently presents a good argument for Vos Savants thesis, instead of against it. Suppose you pick door number one and all the other doors except door number 53 are opened; to me it seems even more obvious that in that case you should switch than in the three door case. These arguments seem to be difficult to understand and can even make people quite angry. Marilyn vos Savant was ridiculed in many of the letters written to her.

There were also letters by people who reported that they had actually played the game or had made a computer simulation of the game confirming Vos Savant's claim. Several good simulations of the Monty Hall Dilemma can be found on the Internet by searching for the Monty Hall Dilemma. Theo Kuipers wrote that playing the game with his wife not only convinced his wife, but also made them find an elegant solution, see Wouters (1991). Yet some who reported that their findings concurred with Vos Savant were still not convinced. Some even offered their computer programs to be distributed hoping that someone could find a bug. Experimental data can give a correct answer to the question whether you should switch, but experimental data alone do not yield an intuitively appealing analysis of the problem.

The Monty Hall Dilemma is a puzzle for which intuitions fail many people. It is surprising that these wrong intuitions are very strong. But there are many puzzles and paradoxes where one can have strong intuitions that are wrong. One might think for example that there are more natural numbers than there are prime numbers, but this is not true. The best way to show that such counterintuitive results are in fact correct is to use some formal method such as logical analysis. In this chapter I use the logic that was introduced in chapter 6 to analyze this puzzle.

7.2 A semantical analysis

In section 6.4.1 a method was introduced to build a model for specific situations. In this section this method is applied to the Monty Hall Dilemma. The set of propositional variables $\mathcal{P}$ is the union of the three sets $A = \{A_1, A_2, A_3\}$ (where A_i means that the car is behind door number i), $C = \{C_1, C_2, C_3\}$ (where C_i means that the contestant initially chooses door number i), and $O = \{O_1, O_2, O_3\}$, (where O_i means that door number i is opened by Monty Hall).

There are twelve possible outcomes, which can be seen as follows. There are three possible locations for the car. For each of these the contestant can choose three doors. For each of these choices Monty Hall can either open one or two doors. If the same door is chosen as where the car is, both the other doors can be opened. If another door is chosen, the remaining door that does not contain the car must be opened. Initially the contestant cannot distinguish any of these possibilities, and Monty Hall just knows where the car is. An epistemic model for this situation is given in figure 7.1.

The question is what would be considered an appropriate prior probability distribution over these worlds. In the face of ignorance, the usual approach is to consider a uniform probability distribution. But that is inappropriate. There are twelve possible worlds and all would get a probability of $\frac{1}{12}$. If we were to calculate the probabilities assigned to worlds according to Monty Hall we get something odd. Suppose the car is behind door number one. In this case, the domain of the probability function assigned to Monty Hall consists of those worlds where the car is behind door number one only. These are four worlds and each would get probability $\frac{1}{4}$. Consequently Monty Hall would assign probability $\frac{1}{2}$ that the contestant initially chooses door number one. But it seems more appropriate that

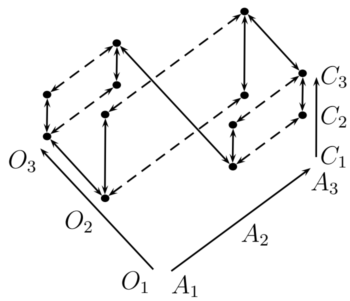


Figure 7.1: An epistemic model of the *initial* situation of the Monty Hall Dilemma. The solid arrows indicate indistinguishability for both the contestant and Monty Hall. The dashed arrows indicate indistinguishability for the contestant only. Transitive and reflexive arrows are omitted.

Monty Hall would assign probability $\frac{1}{3}$ to that. What the contestant is concerned a uniform distribution would also be inappropriate. Suppose that in that case the player chooses door number one. After updating with his own choice there are only four worlds remaining with probability $\frac{1}{4}$ each. But that would mean that although the contestant initially assigned probability $\frac{1}{3}$ to the car being behind door number one, after choosing door number one it would be $\frac{1}{2}$. It seems that afterwards it should still be $\frac{1}{3}$. So a uniform distribution is not appropriate.

So the requirements seem to be that Monty Hall should assign probability $\frac{1}{3}$ to all choices the contestant can make and the contestant should assign probability $\frac{1}{3}$ to all the possible locations of the car, before and after choosing a door. There are many ways to satisfy these constraints, but these requirements do fix the probability assigned to worlds where the door chosen by the contestant is not the door with the car: this probability must be $\frac{1}{9}$. The rest depends on what Monty Hall does if the contestant chooses the door with the car. In that case he can open either of the two remaining doors. Let us assume that in that case he chooses according to a uniform distribution over those two doors. This seems reasonable. The model is shown in figure 7.2.

Now we can use the semantics of probabilistic dynamic epistemic logic to see what happens to the information of both the contestant and Monty Hall during the quiz. The contestant chooses a door, let us assume it is door number one and Monty Hall opens a door containing a goat, let us assume it is door number three. Let us also assume that in fact the car is behind door number two. So now we want to know whether the sentence $[C_1][O_2](\mathbf{P}_c(A_1) = \frac{1}{3} \wedge \mathbf{P}_c(A_2) = \frac{2}{3})$ is true in the world where the car is behind door number two, door number one is chosen and door number three is opened. Therefore we have to calculate the results of the updates. This is shown in figure 7.3, and indeed in the resulting model the probability that the car is behind door number one is one third according to the

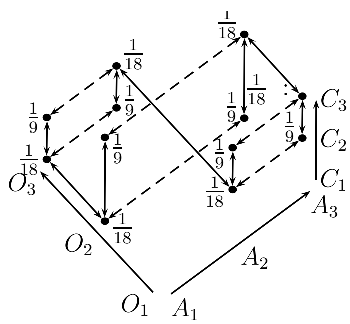
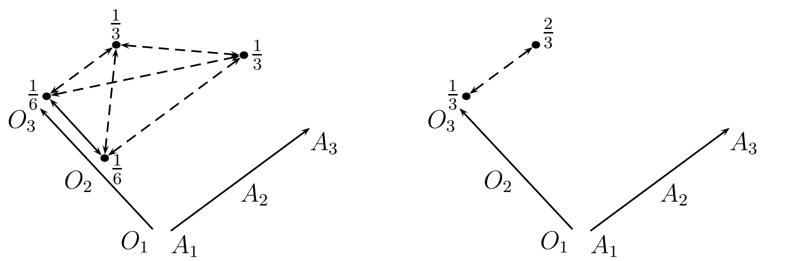


Figure 7.2: The probabilistic epistemic model of the initial situation of the Monty Hall Dilemma. The probabilities shown are the prior probabilities of the worlds.



a. The relevant part of the probabilistic epistemic model after the contestant chooses door number one.

b. The relevant part of the probabilistic epistemic model after Monty Hall opens door number three.

Figure 7.3:

contestant and the probability that the car is behind door number two is two thirds. Therefore he should switch.

7.3 A syntactic analysis

In a syntactic approach the idea is that we try to represent the inference in the language of probabilistic epistemic logic, and show that it is valid. So, what are the premises of the inference. One of the rules is that there is only one car behind the doors, the contestant may only choose one door, and Monty Hall may only open one door. We use the same sets of propositional variable as in the previous section.

$$\begin{aligned} \mathsf{onecar} &= \bigoplus A \\ \mathsf{onechoice} &= \bigoplus C \\ \mathsf{oneopen} &= \bigoplus O \end{aligned}$$

Where $\bigoplus$ means exclusive or. I assume that the contestant should assign a probability of $\frac{1}{3}$ to the car being behind a particular door. This is an assumption that has to be made to get vos Savant's answer. Moreover I assume the contestant does not learn anything about the location of the car by picking a door. Therefore the contestant should still assign a probability of $\frac{1}{3}$ after picking a door: the contestant's choice is independent of where the car is.

$$\begin{aligned} \mathsf{equal} &= \bigwedge_{i \in \{1,2,3\}} \mathbf{P}_c(A_i) = \tfrac{1}{3} \\ \mathsf{independentAC} &= \bigwedge_{j \in \{1,2,3\}} [C_j]\mathsf{equal} \end{aligned}$$

This is a nice way of expressing independence. This assumption remains implicit in most other analyses I found of the Monty Hall Dilemma. The crucial part of the analysis of the Monty Hall Dilemma is to see under what conditions Monty Hall opens a door. He opens exactly one door such that the contestant did not pick it and the car is not behind it.

$$\mathsf{conditions} = \bigwedge_{i,j=\{1,2,3\}} [C_i](O_j \leftrightarrow (\neg A_j \wedge \neg C_j \wedge \bigwedge_{\substack{k \in \{1,2,3\} \\ k \neq j}} \neg O_k))$$

Let us use $\mathsf{initial}$ as an abbreviation for the conjunction of onecar, $\mathsf{onechoice}$, $\mathsf{oneopen}$, equal, $\mathsf{independentAC}$, and $\mathsf{conditions}$.

The question is whether the contestant should switch or not:

$$\mathsf{switch} = [C_1][O_3]\mathbf{P}_c(A_1) \leq \mathbf{P}_c(A_2)$$

If this sentence is true, then the chances that the contestant wins the car do not decrease by switching. It turns out that $\mathsf{initial}$ is not enough to deduce this result. What is needed is that the contestant is informed about the game: $\mathbf{P}_c(\mathsf{initial}) = 1$. We also need two other very natural assumptions, namely that $\mathbf{P}_c(C_1) > 0$ and $[C_1]\mathbf{P}_c(O_3) > 0$. This suffices to deduce switch.

The $\mathsf{independentAC}$ assumption implies that $[C_1]\mathbf{P}_c(A_1) = \frac{1}{3}$, and therefore:

$$[C_1]\mathbf{P}_c(O_3 \wedge A_1) \leq \tfrac{1}{3}$$

By $\mathsf{conditions}$, $\mathsf{onechoice}$, and $\mathsf{oneopen}$ we get $[C_1]\mathbf{P}_c(A_2 \rightarrow O_3) = 1$. Some probabilistic reasoning gives us that $[C_1]\mathbf{P}_c(O_3 \wedge A_2) = \mathbf{P}_c(A_2)$. This, together with $[C_1]\mathbf{P}_c(A_2) = \frac{1}{3}$ (from $\mathsf{independentAC}$), allows us to infer that $[C_1]\mathbf{P}_c(O_3 \wedge A_2) = \frac{1}{3}$, which yields

$$[C_1]\mathbf{P}_c(O_3 \wedge A_1) \leq \mathbf{P}_c(O_3 \wedge A_2)$$

By atomic permanence we get

$$[C_1]\mathbf{P}_c(O_3 \wedge [O_3]A_1) \leq \mathbf{P}_c(O_3 \wedge [O_3]A_2)$$

From the Probability-Update 1 axiom it follows (using some rewriting and the 0-terms axiom) that

$$\mathbf{P}_c(O_3) > 0 \to (([O_3]\mathbf{P}_c(A_1) \leq \mathbf{P}_c(A_2)) \leftrightarrow (\mathbf{P}_c(O_3 \wedge [O_3]A_1) \leq \mathbf{P}_c(O_3 \wedge [O_3]A_2)))$$

By applying necessitation with $[C_1]$, distribution, the assumption that $[C_1]\mathbf{P}_c(O_3) > 0$, and propositional reasoning we get:

$$[C_1][O_3]\mathbf{P}_c(A_1) \leq \mathbf{P}_c(A_2)$$

Thus far we have made no assumptions about the strategy used by the contestant or Monty Hall. We do not need this to deduce switch, but we do need to assume something about the strategy of Monty Hall if we want to deduce that the probability that the contestant wins the car by switching equals two thirds. Then we need to assume that if Monty Hall can choose between opening two doors (if the door the contestant picked is the same as where the car is), then the probability he opens one door is the same as the probability he opens the other door. This boils down to:

$$\mathsf{equalopen} = \bigwedge_{\substack{\{i,j,k\} = \{1,2,3\} \\ i \neq j \\ i \neq k}} [C_i]\mathbf{P}_c(O_j) = \mathbf{P}_c(O_k)$$

With this we can deduce:

$$\mathsf{Savant} = [C_1][O_3]\mathbf{P}_c(A_1) = \tfrac{1}{3} \wedge \mathbf{P}_c(A_2) = \tfrac{2}{3}$$

In order to deduce this, first of all we have to see that using conditions and onecar we can get

$$[C_1]\mathbf{P}_c(O_3) = \mathbf{P}_c(O_3 \wedge A_1) + \mathbf{P}_c(O_3 \wedge A_2)$$

Moreover given similar reasoning as before we get $[C_1]\mathbf{P}_c(O_3 \wedge A_1) = \mathbf{P}_c(A_1)$. Therefore using independentAC we get

$$[C_1]\mathbf{P}_c(O_3) = \mathbf{P}_c(O_3 \wedge A_1) + \tfrac{1}{3}$$

Moreover from oneopen and equalopen and conditions we get that $\mathbf{P}_c(O_3) = \frac{1}{2}$. Therefore $\mathbf{P}_c(O_3 \wedge A_1) = \frac{1}{6}$. In order to apply the Probability-Update 1 axiom we need to get the right form. But given what we have we can deduce

$$[C_1]\mathbf{P}_c(O_3 \wedge A_1) = \tfrac{1}{3} \times \mathbf{P}_c(O_3)$$

with atomic permanence we get

$$[C_1]\mathbf{P}_c(O_3 \wedge [O_3]A_1) = \tfrac{1}{3} \times \mathbf{P}_c(O_3)$$

As before, using Probability-Update 1, we get

$$[C_1][O_3]\mathbf{P}_c(A_1) = \tfrac{1}{3}$$

Given conditions we have that

$$[C_1][O_3]\mathbf{P}_c(A_1) + \mathbf{P}_c(A_2) = 1$$

From this we can deduce Savant.

7.4 Conclusion

There are other formal methods one might use to analyze the Monty Hall Dilemma. One might use a Bayesian approach. One can also think of decision theory or game theory. I do not at all reject these methods as successful means of analyzing the Monty Hall Dilemma. There is a vast literature about the Monty Hall Dilemma. In my view a logical approach to the problem is the best.

Why does the solution of the Monty Hall Dilemma seem so counterintuitive? The question whether the contestant should switch door in order to increase his chance of winning the car is usually answered by stating that it does not matter whether you should switch or not. There is one very natural way by which to arrive at this answer. The information Monty Hall provides by opening door number three is simply seen as an update with the sentence $\neg A_3$. If one updates with this sentence, then the probability that the car is behind any of the remaining doors equals $\frac{1}{2}$. But this is not the only information Monty Hall provides. The conditions that must hold in order for him to be able to provide the information that the car is not behind door number three are also provided. The question is whether the formalism employed to analyze the dilemma is rich enough to be able to express these conditions. In other words, one should be able to write them down. As was seen in the previous section the logic presented in chapter 6 is such a system. Because logic deals with the question whether an inference is valid, I think logic is the best way to analyze problems such as the Monty Hall Dilemma.

Chapter 8

Mastermind

The subject of this chapter lies somewhat outside the scope of the central issues in this thesis. Logic and higher-order information do not really play a role. Information change is the only link to the rest of the thesis. However I find this chapter worthwhile on its own, and therefore it is included in this thesis.

8.1 Introduction

Mastermind is a two-player zero-sum game of imperfect information. player i must choose a combination of four pawns drawn from six colors (hence there are 6^4 possible combinations). Player II does not know the choice player I made. Then player II can ask eight questions in the form of a combination. If she asks the secret combination, she wins the game, otherwise player I wins the game. Each time player II asks a question, she gets an answer that expresses the accuracy of the question. She gets to know how many colors are in the right place and she gets to know how many colors are in the combination, but not in the right place. For example:

$AABB$	the secret combination
$BBAB$	the question

The answer given in this case is: 1 in the right place and 2 are the right color but are not in the right place (I will abbreviate this as (1,2)).

In this chapter some strategies that can be calculated quite easily are discussed including their quality. The game mastermind is so interesting because it provides examples in which one can play around with notions of quality of different questioning stategies. These strategies could be generalized to more general question-answer settings.

8.2 Game theoretic analysis

Although giving an answer seldom gives rise to any problems for people playing the game, the rules of the game formulate quite vaguely how an answer should be determined. My interpretation is the following. Take the set of all combinations that could be made from the pawns that the question was made of (the set of anagrams of the question). Take from this set the element that matches the secret combination best. This yields the maximum number of pawns one could put in the right place, given the pawns of one's question. This number minus the number of pawns in the right place is the second number. It can be calculated by counting for every color the number of pawns that occur in the question and the number of pawns that occur in the secret combination of that color. Take the minimum of these numbers. Add them all up, and subtract the number of pawns in the right place. Formally: let p pawns be drawn from a set C of c different colors (Let us also assume that C is (alphabetically) ordered and let c_i be the i-th element of that ordering). The answer function $a : (C^p \times C^p) \rightarrow (\mathbb{N} \times \mathbb{N})$ can be defined as follows. Let $s = (s_1, \ldots, s_p) \in C^p$ be a secret combination and $g = (g_1, \ldots, g_p) \in C^p$ be a question

$$\begin{aligned} a(s,g) &= (\#(\{j \mid s_j = g_j\}), \\ &\quad (\sum_{i=1}^{c} \min(\#(\{j \mid s_j = c_i\}), \#(\{j \mid g_j = c_i\}))) - \#(\{j \mid s_j = g_j\})) \end{aligned}$$

As I mentioned, people playing the game do not find the definition vague. So for the example where $s = AABB$ and $g = BBAB$, the first number of the answer is 1, because s and g are equal on the fourth pawn only. There are two A's in s and there is one A in g, the minimum of which is one. For B we get two. From the sum of these the first number of the answer is subtracted, which yields two.

In principle mastermind can be analyzed using standard game-theoretic techniques. Simply draw the game tree and apply Zermelo's algorithm to calculate the value of the game (see Binmore (1992)). For mastermind we have a finite tree. First player I chooses a combination and then player II can choose a combination 8 times. So the tree is finite. It contains at most $(6^4)^9 \approx 1.10^{28}$ terminal nodes. This is quite a large tree.

For mastermind, however, the value of the game is known, because strategies have been found that need eight or less questions. Because one wants to be able to say something about the quality of these strategies, other than being a strategy that will win the game, one has to change some of the rules. For example, one could imagine that the players split eight dollars by playing mastermind. Player I gets as many dollars as the number of questions it takes player II to win and player I gets the rest. Or one could limit the number of questions, for example to five questions. The first of these is more often taken as the measure of quality of a strategy. The optimal strategy for this measure of quality has been found by

Koyoma and Lai (1993) by a depth first computer search. Assuming a uniform distribution over the possible combinations, the value of the game for player II is 4.340. The strategy needs a total of 5625 questions. This strategy is difficult to generalize. Therefore I will focus on strategies that are more easy to calculate.

8.3 A simple strategy

The first strategy I want to discuss is by Shapiro (1983) (it is also published in Sterling and Shapiro (1994)). What his algorithm does is the following. The combinations are somehow ordered (usually alphabetically), then it asks the first. The answer is received. The next question is the first one in the ordering that is consistent with the answers given so far. And so on until the combination is cracked. A crucial drawback to this strategy, however, is that it looks at the informativity of questions very marginally. one can only be certain that one does not know the answer already, but that is all. We will see in section 8.6 that a question, which cannot be the secret combination can be very informative.

8.4 Looking one step ahead

In mastermind a question partitions the set of possible combinations. This can be seen in the following example. Consider a simplified mastermind game with two pawns and four colors. The set of possible combinations can be represented as follows:

$$\begin{array}{cccc} DA & DB & DC & DD \\ CA & CB & CC & CD \\ BA & BB & BC & BD \\ AA & AB & AC & AD \end{array}$$

The questions AA and DA can be represented by the corresponding answers as:

$$\begin{array}{cccc} 1,0 & 0,0 & 0,0 & 0,0 \\ 1,0 & 0,0 & 0,0 & 0,0 \\ 1,0 & 0,0 & 0,0 & 0,0 \\ 2,0 & 1,0 & 1,0 & 1,0 \end{array} \qquad \begin{array}{cccc} 2,0 & 1,0 & 1,0 & 1,0 \\ 1,0 & 0,0 & 0,0 & 0,1 \\ 1,0 & 0,0 & 0,0 & 0,1 \\ 1,0 & 0,1 & 0,1 & 0,2 \end{array}$$

It is obvious that the second question is more informative than the first, but how can we motivate this intuition?

Let us look at the partitions for the standard mastermind game; four pawns and six colors.

	AAAA	*AAAB*	*AABB*	*AABC*	*ABCD*
(0,0)	625	256	256	81	16
(0,1)	0	308	256	276	152
(0,2)	0	61	96	222	312
(0,3)	0	0	16	44	136
(0,4)	0	0	1	2	9
(1,0)	500	317	256	182	108
(1,1)	0	156	208	230	252
(1,2)	0	27	36	84	132
(1,3)	0	0	0	4	8
(2,0)	150	123	114	105	96
(2,1)	0	24	32	40	48
(2,2)	0	3	4	5	6
(3,0)	20	20	20	20	20
(4,0)	1	1	1	1	1

One can look at this table from the perspective of wanting to minimize the *maximum* number of questions required to win the game and one can look at it from the perspective of wanting to minimize the *average* number of questions required to win. From the first perspective, a worst case perspective, one wants to look at the maximum number of questions required for each of the resulting partition elements. As we are looking only one step ahead, the only estimate we can make of this must be based on the size of the partition element. One can assume that the larger the element is, the more questions are required.

	AAAA	*AAAB*	*AABB*	*AABC*	*ABCD*
size largest partition element	625	317	256	276	312

So the first question we should ask is a question like $AABB$. This idea is worked out in the strategy by Knuth (1977).

From the perspective of wanting to minimize the *average* number of questions, other considerations play a role. There are several approaches one can take here. First one could say that now the selection should not be based on the worst case, but on the 'average case'. In other words, one should look at the expected size of the partition element one ends up in. The expected size of a partition element is the probability of getting the answer corresponding to that partition element multiplied with the size of the partition element. This expectation is defined as follows for the first question. Let A be the set of possible answers to questions.

Let g be a question, then the expected size of the partition element one ends up in is:

$$\sum_{a_i \in A} \mathbf{P}_g(a_i).\#(\{x \mid x \in C^p \wedge a(x,g) = a_i\})$$

where $\mathbf{P}_g(a_i)$ is the probability that the answer to g is a_i. If one assumes the distribution over the possible combinations to be uniform, then:

$$\mathbf{P}_g(a_i) = \frac{\#(\{x \mid x \in C^p \wedge a(x,g) = a_i\})}{\#(C^p)}$$

For example $\mathbf{P}_{AAAA}(0,0) = \frac{625}{1296}$, because $6^4 = 1296$. So the expected size is

$$\sum_{a_i \in A} \frac{\#(\{x \mid x \in C^p \wedge a(x,g) = a_i\})^2}{\#(C^p)}$$

For the first question the expected sizes are shown in the table below

	AAAA	*AAAB*	*AABB*	*AABC*	*ABCD*
expected size	511.9	235.9	204.5	185.3	188.2

From this pont of view, one should select *AABC* as the first question. This approach is taken by Irving (1979).

One can also take another approach, which can be motivated as follows. Suppose that a player has to guess which card another player chose from an ordinary deck of cards. Let us assume that this is random in the sense that the probability distribution is uniform. The player wins 1$ if the guess is correct. There are 52 possibilities. Before the player guesses she can ask one question. She must divide the cards into two (nonempty) piles and ask to which of the piles the card belongs (for example red and black, or spades and non-spades, or even the queen of hearts and not the queen of hearts). Which question is best? They are all equally good. This can be seen as follows. Suppose the two piles have sizes x and y. The card is in group x with probability $\frac{x}{52}$. The probability of guessing the right card if it is in this group is $\frac{1}{x}$. The card is in group y with probability $\frac{y}{52}$. The probability of guessing the right card if it is in this group is $\frac{1}{y}$. Hence, the expected gain is:

$$\frac{x}{52} \cdot \frac{1}{x} \cdot 1\$ + \frac{y}{52} \cdot \frac{1}{y} \cdot 1\$ = \frac{2}{52}\$$$

So it does not matter what the sizes of x and y are.

This can be generalized. Suppose there is a set A and we have to guess what element of A we are dealing with. We also have to assume that the probability distribution on A is uniform. Before we guess we can ask a question that can be seen as a partition $V = \{V_i, \ldots, V_n\}$. The probability of guessing correctly, once we learn in which part of V the element is:

$$\sum_{i=1}^{n} \frac{\#(V_i)}{\#(A)} \cdot \frac{1}{\#(V_i)} = \frac{n}{\#(A)}$$

So in these cases the sizes of the elements of the partition do not matter. The only thing that matters is the size of the partition, i.e. the number of elements of the partition.

This can be generalized to games with more rounds. Assume that the probability of guessing the element of any set S correctly in a game with r rounds is the number of parts that one can partition S in with r questions, divided by the cardinality of S, where the player's question can depend on the answer to the previous questions. Let us look at a game with $r+1$ rounds. A player can ask $r+1$ questions, and then has to guess which element of A the other player chose. Let the first question be a partition $V = \{V_1 \ldots V_n\}$. Let n_i indicate the number of parts V_i can be partitioned in with the rest of the questions. Using the induction hypothesis we infer that when the game is played in r rounds for V_i the probability of guessing the element of V_i equals n_i divided by the cardinality of V_i. Then the probability of guessing correctly in $r+1$ rounds for the set A is:

$$\sum_{i=1}^{m} \frac{\#(V_i)}{\#(A)} \cdot \frac{n_i}{\#(V_i)} = \sum_{i=1}^{m} \frac{n_i}{\#(A)}$$

So the probability of guessing the element of set A correctly in a game with $r+1$ rounds equals the number of parts that one can partition A in with $r+1$ questions divided by the cardinality of A. By induction one can conclude that this holds for any r and any set S.

So in mastermind, if one wants to maximize the number of combinations for which one would win in a certain round, then one should maximize the number of parts the set of all combinations is partitioned in, in the previous round. It is still not feasible to calculate this for an interesting number of rounds, such as five, but it can be used as a motivation for a strategy. It is interesting both from the point of view of minimizing the average number of questions required as from the point of view of minimizing the maximum number of questions required. Let us look at the first question again, then we see:

	AAAA	*AABB*	*AABB*	*AABC*	*ABCD*
number of parts	5	11	13	14	14

So this strategy should start with either *AABC* or *ABCD*. When one writes a computer-program, however, one has to make a choice. In most of the literature an alphabetical ordering is used and those combinations are preferred that are still possible. This strategy, which maximizes the number of partition elements, is not discussed in the literature on mastermind that I found.

One might argue that the strategy where only the number of parts matter is not specific enough. Consider again the example of the cards given above. Surely, when considering a game with more than one round, a question splitting the cards in exactly two equally large piles is better than a question splitting it into one pile with one card and the other cards in the other pile, because if the answer is the smaller pile, on cannot ask any non-trivial question. Therefore the partition would not be maximal given a limited number of rounds.

There is a measure that gives an ordering on partitions which is called entropy (see Cover and Thomas (1991)), which in the case of the cards would select a question with two piles of 26 cards. This can be motivated by the following example. Suppose we have another guessing game. Player I picks a card randomly from a deck of cards. Player II has to determine which card Player I picked using a minimal number of yes/no questions. If there are eight cards for example, one needs three questions to determine which card it is. This is the $\log_2(8)$. The logarithm gives an approximation of the expected number of yes/no questions needed. (It is not exactly the expected number of questions, because one should look at the logarithm as the limit of the expected number of questions, if one can play a number of these games simultaneously.) Suppose we have a partition $V = \{V_1, \ldots, V_m\}$ of a set A. Let p_i be $\frac{\#(V_i)}{\#(A)}$. (This is the probability that an element of V_i is in A. If the probability distribution is not uniform another definition is needed.) Then the expected number of yes/no questions could be represented as

$$\sum_{i=1}^{n} p_i \log(\#(V_i))$$

Trying to minimize this measure is the same as trying to maximize the entropy, S, which is defined as

$$S(V) = -\sum_{i=1}^{n} p_i \log(p_i)$$

since $\log p_i = \log(\frac{\#(V_i)}{\#(A)}) = \log(\#(V_i)) - \#(A)$. In figure 8.1 a graph displaying the entropy for partitions with two elements is drawn. The variable for the x-axis is the probability p of one of the elements of the partition, the entropy is given on the y-axis. So the graph shows the function $-p \log p + -(1-p)\log(1-p)$.

Let us look at the entropies of the first questions.

	AAAA	*AAAB*	*AABB*	*AABC*	*ABCD*
entropy	1.498	2.693	2.885	3.044	3.057

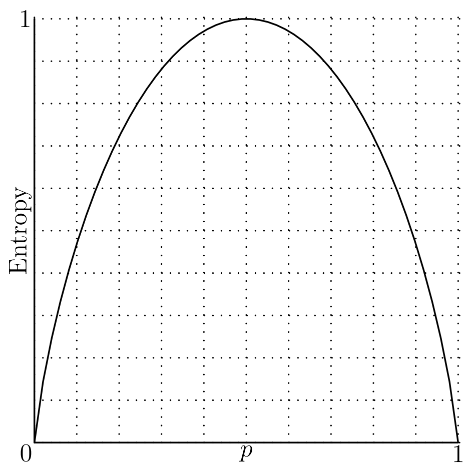


Figure 8.1: Entropy of a partition with two elements.

This strategy is one of the strategies studied by Neuwirth (1982). He has introduced another strategy that seems to perform quite well. I do not discuss it here.

8.5 Empirical results

The first table shows for each strategy for how many combinations the game is won in a particular round of the game. Or put in other words: each strategy produces a game tree, the table shows for each depth of the tree how many leafs (nodes without successors) there.

Round number	1	2	3	4	5	6	7	8	9
Shapiro	1	4	25	108	305	602	196	49	6
Maximum size	1	6	62	533	694	0	0	0	0
Expected size	1	10	54	645	583	3	0	0	0
Most parts	1	12	72	635	569	7	0	0	0
Entropy	1	4	71	612	596	12	0	0	0

The second table shows the same results, but shows for how many combinations the game has been won before or at the end of a particular round, i.e. the numbers in the table above are added.

Round number	1	2	3	4	5	6	7	8	9
Shapiro	1	5	30	138	443	1045	1241	1290	1296
Maximum size	1	7	69	602	1296	1296	1296	1296	1296
Expected size	1	11	65	710	1293	1296	1296	1296	1296
Most parts	1	13	85	720	1289	1296	1296	1296	1296
Entropy	1	5	76	688	1284	1296	1296	1296	1296

The third table shows how many questions are needed in total in the strategy and the average number of questions needed (the average length of a path to a leaf). The numbers in the second column are rounded.

	total number of questions	average number of questions
Shapiro	7471	5.765
Maximum size	5801	4.476
Expected size	5696	4.395
Most parts	5668	4.373
Entropy	5722	4.415

8.6 Evaluation

In this section I will try to say something more about the empirical results. It seems quite surprising that Shapiro's strategy performs so badly regarding the maximum number of rounds required and the average number of rounds required. It does not even guarantee that one wins in eight rounds. It seems that the first question that is asked is not a good choice. This can easily be improved by choosing another combination than $AAAA$ to be the first combination that is asked and let the rest be ordered alphabetically. $AABB$ for example gives the following results:

Round number	1	2	3	4	5	6	7	8	9
Shapiro starting with $AABB$	1	12	71	253	588	286	78	7	0

which is considerably better. But it still performs badly in comparison to the other strategies. Why the maximum is higher can be explained by the following example. If one uses the entropy strategy and the secret combination is $CCCC$, the following game will be played.

round	question	answer
1	$ABCD$	$(1, 0)$
2	$BEEF$	$(0, 0)$
3	$AAAC$	$(1, 0)$
4	$CCCC$	$(4, 0)$

After the second question, only three secret combinations are possible, $AAAA$, $CCCC$, and $DDDD$. Now look at the following table, where for each of these remaining possibilities the answers is shown for asking the question in the column.

	$AAAA$	$CCCC$	$DDDD$	$AAAC$
$AAAA$	$(4,0)$	$(0,0)$	$(0,0)$	$(3,0)$
$CCCC$	$(0,0)$	$(4,0)$	$(0,0)$	$(1,0)$
$DDDD$	$(0,0)$	$(0,0)$	$(4,0)$	$(0,0)$

A consistent question would not be able to distinguish all three combinations, but the question $AAAC$ can, as can be seen in the table. In this way the maximum number of questions required can be reduced. In all other strategies except Shapiro's inconsistent questions occur. An interesting question related to this example is whether any set of three combinations can be separated with one question. Maybe even every set of four or five.

One of the other interesting results is that, although strategies often have no theoretic way to distinguish two questions, but only alphabetic ways of distinguishing, the empirical results give a different answer. Due to a programming error the first tests that I ran had strategies that picked the alphabetically last optimal combination if a unique optimal combination was not in the set of remaining possibilities. These give a slightly different picture.

Round number	1	2	3	4	5	6	7	8	9
Maximum size	1	8	65	522	696	4	0	0	0
Expected size	1	10	54	646	582	3	0	0	0
Most parts	1	12	72	636	568	7	0	0	0
Entropy	1	4	70	613	596	12	0	0	0

Shapiro's strategy has been left out of this table, because these considerations do not affect his strategy. These differences are very small. They are greatest in case of Knuth's strategy of minimizing on the maximum size of the partition elements. I think this means none of the strategies proposed here can be defended theoretically in a satisfactory manner.

Why the results are so very different in the Knuth's case is because of the following. After the first question has been answered, the number of ways the set of remaining possibilities can be partitioned in is quite large. As we know there are only five types of question that can be asked in the initial state. But after the first question has been answered there are much more. The following table shows the number of questions that can be asked if the first answer is $(1,0)$.

question	answer	number of different questions
$AAAA$	$(1,0)$	12
$AAAB$	$(1,0)$	53
$AABB$	$(1,0)$	34
$AABC$	$(1,0)$	125
$ABCD$	$(1,0)$	52

So in Knuth's strategy, there are already 34 different kinds of partitions that can be made. His strategy only looks at one aspect of these partitions and apparently this is not fine-grained enough to result in a robust strategy. If there are already 34 questions that can be asked after the first question, this will be worse after more.

Irving's strategy of minimizing on the expected size of the set of remaining possibilities is straightforward, but his paper contains a number of strange (irreproducible) results. First of all he claims that a closer investigation of Knuth's strategy reveals that the total number of questions required for all 1296 combination is 5804, whereas it is 5801 according to my calculations. This can be explained by a minor programming error (the same that I made), but I cannot explain any of his other results. He says his strategy selects the first two questions on the basis of the expected number of remaining possibilities and the rest by exhaustive search. When I look at the second question he makes after the first reply I disagree with him on five questions. In four of those it is simply the case that he does not take the first one out of the list that is available to him. In one case it is simply wrong. His first question is $AABC$. If the reply to this question is (3,0), according to Irving the next question should be $FBAC$. (One immediately wonders why not $DBAC$.) According to my calculations, the expected size of the set of remaining possibilities after this question is 4.7. However, if one asks $ABCC$ the expected size is 3.6, which is quite different. One difference between these two questions is that Irving's question partitions the remaining possibilities in 8 parts, whereas $ABCC$ partitions the set of remaining possibilities in 7 parts. So it might be the case he took the average number of remaining possibilities, instead of the expected size, but I still cannot reproduce his results.

The "most parts" strategy results in the best strategy when one looks at the average number of questions, the only problem is that the theory behind it tells you that the number of rounds really matters, whereas this is ignored in selecting a question. Each time one only looks one step ahead. I found the following looking two steps ahead.

	AAAA	*AAAB*	*AABB*	*AABC*	*ABCD*
(0,0)	14	14	14	13	8
(0,1)	0	14	14	14	13
(0,2)	0	9	12	14	14
(0,3)	0	0	7	10	11
(0,4)	0	0	1	2	4
(1,0)	13	14	14	14	13
(1,1)	0	13	14	14	14
(1,2)	0	7	10	11	11
(1,3)	0	0	0	4	4
(2,0)	11	12	12	12	12
(2,1)	0	9	10	11	9
(2,2)	0	3	4	4	4
(3,0)	5	8	8	8	7
(4,0)	1	1	1	1	1
total	44	104	121	132	125

The numbers in the table is the number of different answers one could get by asking a question, after the initial question given and the initial answer. So the total number at the bottom is the total number of parts of the partition that results from asking two questions. So if the game consists of three rounds, it is best to start with *AABC*.

8.7 Conclusion

My conclusion is that Shapiro's strategy is not very good, but that I cannot order the other strategies in any way, when one wants to apply these strategies to other areas. To be able to say something more about this, empirical results about more difficult variants of mastermind (more pawns or more colors) are needed. I suspect that if one makes it easier (less pawns or less colors) the different strategies will again lead to approximately the same results.

Chapter 9

Trying to resolve the two-envelope problem

The subject of this chapter lies somewhat outside the scope of the central issues in this thesis. Higher-order information does not really play a role. Although logic and probability theory do play a role, and the focus is on other issues regarding logic and probability theory than in the rest of this thesis. However I find this chapter worthwhile on its own, and therefore it is included in this thesis.

9.1 Introduction

The two-envelope paradox is a problem that can baffle people. The earliest form of the paradox appeared in Kraitchik (1943) (see also Nalebuff (1989)). He discusses the paradox of the neckties.

> 'Each of two persons claims to have the finer necktie. They call in a third person who must make a decision. The winner must give his necktie to the loser as consolation. Each of the contestants reasons as follows: 'I know what my tie is worth. I may lose it, but I may also win a better one, so the game is to my advantage'. How can the game be to the advantage of both?'

Kraitchik also formulated a variant where two people compare the number of pennies in their purses. This form also appears in Gardner (1982), where it is called the wallet game. It is unclear who gave the problem its modern form. Zabell (1988a), (1988b) heard it from Steve Budrys (see Nalebuff (1989)). It goes along the following lines.

There are two indistinguishable envelopes. Both envelopes contain a check, upon which an amount of money is written. One of the checks is worth twice as much as the other check. One picks one of the envelopes, opens it and observes the amount on the check. Then one is allowed to decide between

1. keeping the envelope one has;

2. returning the envelope and take the other one.

Two variants are distinguished: the discrete case, where the amount is a natural number; the continuous case, where the amount is a positive real number.

In the *Encyclopedia of Philosophy*, Van Heijenoort defines a paradox, especially a logical paradox, as follows:

> A paradox, in the original sense of the word, is a statement that goes against generally accepted opinion. In logic the word has taken on a more precise meaning. A logical paradox consists of two contrary or even contradictory propositions to which we are led by apparently sound arguments. The arguments are considered sound because when used in other contexts they do not seem to create any difficulty. (Van Heijenoort, 1967)

Take the liar paradox for example:

> L: The sentence L is false.

There is a sound argument that the sentence is true: suppose L is false, therefore it is not the case that L is false. Therefore, by *tertium non datur* L is true. There is also a good argument that L is false: suppose that L is true, but L says that L is false. Therefore it is false. The two propositions "L is true" and "L is false" are contradictory. The line of argument that led to both conclusions seems unexceptionable.

In case of the two-envelope problem, there are also two contrary propositions to which we are led by apparently sound arguments. The first proposition is that one should switch. Let us first introduce some notation. Let y be the lesser amount on the checks. Let z be 1 if one picks the envelope with y, and 2 otherwise. The amount found in the envelope is $x(= zy)$. Now one can argue as follows: the probability of picking the envelope with the larger amount is $\frac{1}{2}$. One observes an amount of x. Therefore the probability that the other envelope will contain $\frac{1}{2}x$ equals $\frac{1}{2}$ and the probability that it will contain $2x$ is also $\frac{1}{2}$. The *expected value* of the amount in the other envelope therefore equals $\frac{5}{4}x(= \frac{1}{2} \times \frac{1}{2}x + \frac{1}{2} \times 2x)$. The expected value is greater than x, therefore one should switch.

The other proposition is that it does not matter whether one switches or not. The strange thing is that it follows from the argument above. It is clear that the argument above applies regardless of the observed amount x. Therefore one should switch anyway. In that case it seems that one could just as well choose the other envelope initially. As this is true for both envelopes it does not matter which envelope one chooses, and therefore it does not matter whether one switches or not.

This chapter is the result of joint work with Casper Albers and Willem Schaafsma (see Albers, Kooi, and Schaafsma (2002)). It contains three parts, in which three distinct problems are discussed. In section 9.2, the two-envelope paradox is regarded as a purely logical paradox. In section 9.3 the two-envelope paradox is regarded as a paradox in probability theory. Even when the logical paradox is solved, this paradox remains. Finally a few words are spent on the two-envelope *problem*. Even when the two-envelope *paradox* is solved, there still is the problem of choosing between keeping the envelope one has and returning the envelope and taking the other one.

9.2 A purely logical paradox

Smullyan (1997) maintained that 'probability is really quite inessential to the heart of this paradox'. He presents it as a purely logical paradox with no reference to probability or expected values.

Proposition 1: The amount you will gain by trading, if you do gain, is greater than the amount you will lose, if you do lose.

Proposition 2: The two amounts are really the same.

He proves both of them in the following way[1]: "To prove Proposition 1, let x be the amount you are now holding. Then the other envelope either contains $2x$ or $x/2$. If you gain by trading, you will gain x dollars (moving from x to $2x$), whereas if you lose by trading, you will lose only $x/2$. Since x is greater than $x/2$, then Proposition 1 is established.

To prove Proposition 2, let d be the difference between the two amounts in the envelopes (or what is the same thing, the lesser of the two amounts). Well, if you gain on the trade, you will gain d dollars. If you lose on the trade, you will lose d dollars. Since d is equal to d, then Proposition 2 is established." (Smullyan (1997, p.174))

Thus we are truly dealing with a logical paradox. The solution of a paradox must meet a number of conditions. Susan Haack indicates these in her book *Philosophy of Logics*.

> [...] This suggests two requirements on a solution; that it should give a consistent formal theory [...] in other words, indicate which apparently unexceptionable premises or principle of inference must be disallowed (the *formal* solution); and that it should, in addition, supply some explanation of *why* that premise or principle is, despite appearances, exceptionable (the *philosophical* solution). [...] Further

[1] In his book Smullyan uses the letter n in his proof of proposition 1. We have replaced it with x to keep a uniform notation.

> requirements concern the scope of a solution; it should not be so broad as to cripple reasoning we want to keep (the 'don't cut off your nose to spite your face' principle); but it should be broad enough to block all relevant paradoxical arguments (the 'don't jump out of the frying pan into the fire' principle); the 'relevant', of course, glosses over some problems. (Haack (1978, pp. 138 – 139))

This indicates several ways of tackling a paradox. One can develop a philosophically well motivated formal system in which at least one of the arguments is flawed. One can claim the propositions are not truly contrary, by for example claiming that they are not propositions at all. In the case of the logical two-envelope paradox I claim that the propositions are not contrary because they are ambiguous. This will be argued by using Lewis's theory of counterfactuals.

Lewis's theory of counterfactuals was developed by a certain dissatisfaction with *material implication* and *strict implication*. Material implication, which is denoted by $\varphi \rightarrow \psi$, is equivalent to $\neg(\varphi \wedge \neg\psi)$. Strict implication, which is denoted by $\varphi \prec \psi$, is equivalent to $\Box(\varphi \rightarrow \psi)$ in ordinary alethic modal logic. Yet neither can capture the conditional expressed in the following sentence.

> If kangaroos had no tails, they would topple over.

It is clear that these kinds of sentences are not to be interpreted as material implications. Lewis argues that they cannot be interpreted as strict implications either, because of their non-monotone behavior. (See Lewis (1973)).

In Lewis's system these kind of conditionals are denoted by $\varphi \mathrel{\Box\!\!\rightarrow} \psi$. They differ from both material implication and strict implication. If one would consider the conditional above to be a material implication it would be true merely because the antecedent is false. This is rather counterintuitive. On the other hand if it is read as a strict implication, then there are other counterintuitive results. This is best illustrated by the following sentences.

> If Oswald did not kill Kennedy, then someone else did.
>
> If Oswald had not killed Kennedy, someone else would have.

The first sentence is best read as a strict implication: in all possible worlds where Oswald did not kill Kennedy, someone else did kill him, because in fact he was killed. The other sentence is different. It cannot be read as a strict implication. The antecedent expresses a *counterfactual situation*, and assumes that in fact Oswald did kill Kennedy. Lewis reads this sentence as follows. Someone else killed Kennedy in all those possible worlds that are as similar as possible to the actual world, except that Oswald did not kill Kennedy. Assuming that there was no conspiracy, and that no one else tried to kill Kennedy, those possible worlds include some where no one killed Kennedy and he would still be alive, or died in some other way than by someone killing him. It is quite vague what it

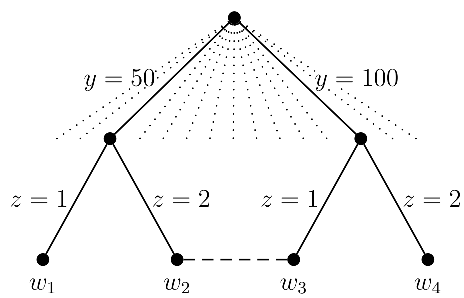


Figure 9.1: A picture of the game tree of the two-envelope problem. The dotted lines indicate other possibly choices for y.

means to be 'as similar as possible' to the actual world, but the important thing is that this notion of conditionals seems to capture a lot of our intuition about counterfactuals.

Now let us return to the logical two-envelope paradox. Consider the following term:

the amount you will gain by trading, if you do gain

Let us assume that the condition in this term expresses a counterfactual situation and that in fact you do not gain by trading. (Using first order logic and the iota-operator (see (Gamut 1991), volume 1, section 5.2) with counterfactuals this term could be represented as $\iota x(\exists y G(y) \mathrel{\Box\!\!\rightarrow} G(x))$, where $G(x)$ is to be read as "you gain x by trading".) What possible worlds are as similar as possible to the actual world, where you do gain by trading? To make clear what choices we can make, let us look at a picture of the game tree for this problem shown in Figure 9.1. First an amount y and an amount $2y$ are put into two envelopes. Then you pick one of these envelopes, containing the check with $2y$, then $z = 2$, or the check containing y, then $z = 1$. Let us assume that in fact $y = 50$ and that $z = 2$. In that case you lose by trading. The actual world is thus w_2. In both world w_1 and world w_3 you gain by trading. Which of these is more similar to w_2? There are three possible answers: w_1 is more similar, w_3 is more similar, and they are equally similar (assuming that similarity defines a total order). Any choice that is made resolves the paradox. In the case where another world is the actual world we can resolve the paradox analogously.

It seems a matter of perspective which of these worlds is regarded as most similar. To you world w_3 seems closer, because according to the information you have it is still possible. It is a world in which you have the same information, and therefore is similar to the actual world. In this case the amount you would gain would be 100 dollars. On the other hand from the perspective of the envelopes that were available to you, world w_1 seems closer. If you had chosen the other envelope, you would have ended up in w_1. It is a world which has most history in common with the actual world, and therefore it is similar to the actual world. In this case the amount you would gain would be 50 dollars. If you consider both worlds as similar as possible to the actual world there is no fixed amount you win by trading. It is either 50 or 100 dollars. So the term "the amount you will gain by trading, if you do gain" is ambiguous. This ambiguity is exploited in the two arguments for Proposition 1 and Proposition 2. In the proof of Proposition 1 the perspective is taken that w_3 is more similar to the actual world. In the proof of Proposition 2, w_1 is more similar to the actual world.

A similar analysis is given by Chase (2002), but he argues that one interpretation is better than the other. I do not think that one interpretation is better than the other. It is clear that it depends on which meaning one gives to the words, and one can choose as one likes. It is important to note that whichever interpretation is chosen, this does not resolve the other paradox, where we were led to the conclusion that it is better to switch and to the conclusion that it does not matter if one switches. We now turn to resolving this.

9.3 The two-envelope paradox

As many authors have noted the two-envelope paradox can be explained quite easily Zabell (1988a), Zabell (1988b), Nalebuff (1989), Nalebuff (1988), Broome (1995), Clark and Shackel (2000), Chalmers (1994), Jackson, Menzies, and Oppy (1994), Linzer (1994), McGrew, Shier, and Silverstein (1997), and many, many more. The mistake in the argument that the expected value of the amount on the cheque in the other envelope is $\frac{5}{4}x$, is that in the calculation the *prior* probability of picking the larger and the lesser amount is used, instead of the *posterior* probabilities of picking the larger or the lesser amount after observing x. To calculate the expected utility of switching as a general strategy one must know the prior probability that x and $2x$ are in the envelopes, and the prior probability that $\frac{1}{2}x$ and x are in the envelopes. For example if there is always 100 and 50 in the envelopes, the expected value of switching, after finding 100, equal -50 rather than 125. In short, one must know what the distribution over y is, before one can calculate the expected value of switching. And one does not know this distribution, so one cannot calculate the expected value.

9.4 The two-envelope problem

So the paradox is solved, but a problem remains. What should one do? Should one switch or not in a particular case? This is discussed extensively in Albers, Kooi, and Schaafsma (2002) and I will not go into details here. The problem is discussed from the perspective of economy, psychology, logic, probability, and mathematical statistics, respectively, as well as through an in-depth contribution from game theory. The conclusion is that the two-envelope problem does not allow a satisfactory solution.

Chapter 10

Conclusion

In this thesis I have investigated the notions of knowledge, chance, and change, separately and I have investigated how they interact. The logic of knowledge and information, epistemic logic, is a subject worth studying on its own. There are many approaches to representing information and reasoning about information. Epistemic logic deserves a special status among all these approaches because it deals with higher-order information in a very good way. The concept of common knowledge is a prime example of this.

The logic of change, dynamic logic, was studied in chapter 3. One of the best known dynamic logics PDL has a property called non-compactness. This makes it difficult to provide a strongly complete proof system for it. An infinitary system was presented and studied in chapter 3. And generalizations are considered for epistemic logic with common knowledge. Moreover it was shown that the canonical model for this proof system lacked program harmony. There are still interesting application areas remaining.

Epistemic logic can only provide a static picture of knowledge and information. If one is interested in information change there are also many approaches available. If one is especially interested in the change of higher-order information, one enters the realm of dynamic epistemic logics. As we saw in chapter 4, there are many approaches and there is not one approach that is clearly superior to the others. There are also many problems that have not been solved in this area. It remains worth studying.

Another way to generalize epistemic logic is to incorporate probabilistic information. Probability can be interpreted as a degree of belief. There are also other views of probability, but in this thesis the focus is on 'degree of belief' as a generalization of epistemic logic. The relation between this view of probability and the notion of statistical probability was investigated in chapter 5. Of course as a generalization of epistemic logic, multi-agent probability logic and higher-order probability are very interesting subjects. Especially when one is interested in the change of higher-order probabilities one notices that taking conditional probabilities is just one way of handling higher-order probability change. A novel approach to this was presented in chapter 6.

A few problems in the area where knowledge, chance, and change meet were studied in chapters 8, 7, and 9. When higher-order information does not play a role in a certain problem most approaches to representing information and information change seem to be equivalent. For example expansion in belief revision comes up with the same results as updates in dynamic epistemic logic, as far as information about the state of the world is concerned. There is a paper by Segerberg (1999) which links these two approaches. Also as we saw in chapter 6 on page 100, when higher-order information does not play a role, reasoning about conditional certainty is the same as reasoning about information change. Therefore one might just as well use an alternative to (dynamic) epistemic logic when one is not interested in higher-order information. In the analysis of mastermind for instance higher order information plays no role whatsoever. Sjoerd Druiven and Maarten Grachten have proposed a variant of the game where higher-order information does play a role: MastersMinds. In MastersMinds two players try to find out the other player's secret combination, but when they ask a question they also have to answer it themselves. In this way they have to reason about what the other player knows. But when it does play a role (dynamic) epistemic logic provides the best understanding of higher order information.

But what are the remaining challenges in these areas? There are two that I think are very much worthwhile.

A unified theory of information As was indicated in chapter 4 there are a lot of approaches to information. It would be very good to know what the relations between all these approaches are. At the moment it seems that many scientists working in these separate areas of research are very much focused on their own approach. For a large part I do not think this is problematic. This is the case in many sciences. A field can be so broad that someone specialized in one area does not even have to know of the existence of another very specialized area. Let me use an analogy. In a car factory there are several very specialized tasks. Someone who puts the windows into a car does not need to know how the engine works to do this efficiently. Someone who puts the engine into a car does not need to know how the windows work. However they *know* this: it is common knowledge. They do have a general picture of the car that tells them that the proper functioning of the engine has nothing to do with the windows and vice versa. In the study of information, however, the engineers do not seem to know how all the different theories go together: there seem to be interrelations, but it is not clear how one area is related to another. This needs to be found out!

Philosophy of probability for all In chapter 5 I sketched some of the positions in the philosophy of probability. It seems, however, that consensus about the subject is still difficult to find. In my view probability logic can be the bridge between the mathematics of probability and the philosophy of probability. Different philosophies can bring about different logics. But the interrelations between these logics can be studied. I think this can be so successful that probability theorists would sit up and take notice. For example in the two envelope problem as it was presented by Raymond Smullyan it seems that it is unclear which value one should take as fixed. Should one consider the amount in the envelope you chose to be fixed? Or should one consider the difference between the amounts in the envelopes to be fixed? In probability theory these amounts are taken to be random variables. From the point of view of first-order modal logic one would consider them to be nonrigid desigators (referring to different things in different worlds). The lesson from logic that things might go astray when one applies Leibniz' law to nonrigid designators, is quite useful in this context.

Samenvatting

Stel je bent doorgedrongen tot de laatste ronde van een televisiequiz. Je kunt een auto winnen, die achter een van drie deuren staat. De quizmaster vraagt je een deur te kiezen. Nadat je een deur gekozen hebt, moet de quizmaster je helpen. Hij weet waar de auto staat en moet een deur openmaken die je niet gekozen hebt en waar de auto niet achter staat. Nu mag je nog eens kiezen tussen de overgebleven deuren. Als de auto achter de deur staat die je uiteindelijk kiest win je de auto. De vraag is of het voordelig is om van je initiële keuze af te wijken en dus de andere deur te kiezen. Dit is het Monty Hall Dilemma, genoemd naar een quizmaster uit de Verenigde Staten. Het was een van de problemen die de aanleiding vormde voor mijn promotieonderzoek.

De voor velen tegenintuïtieve oplossing van het Monty Hall Dilemma is dat het voordelig is om van je initiële keuze af te wijken. Je wint dan de auto in twee derde van de gevallen. Om deze conclusie te trekken moet rekening gehouden worden met de kennis van de quizmaster, de waarschijnlijkheid dat de auto achter een deur staat en de informatieverandering die optreedt ten gevolge van het openen van een deur. Dit proefschrift gaat over de logica van kennis, waarschijnlijkheid, en verandering.

Logica wordt gebruikt bij onder andere informatica, wiskunde, wijsbegeerte, taalkunde, rechtsgeleerdheid en kunstmatige intelligentie. In de logica houdt men zich bezig met redeneringen en dan vooral met de vraag of een redenering klopt. Bij het beantwoorden van deze vraag wordt niet gekeken naar specifieke redeneringen, maar kijkt men naar de abstracte vorm van redeneringen. Dit gebeurt door redeneringen te vertalen naar een logische taal. Zo kunnen de redeneringen "Alle mensen zijn sterfelijk en Socrates is een mens. Dus Socrates is sterfelijk" en "Alle honden zijn vals en Lassie is een hond. Dus Lassie is vals" na vertaling dezelfde vorm hebben. Zo'n logische taal is toegesneden op redeneringen waarbij een bepaald aspect centraal staat. Zo ontstaan allerlei logica's, elk met een eigen toepassingsgebied. Bij kennislogica, probabilistische logica en dynamische logica staan respectievelijk kennis, waarschijnlijkheid en verandering centraal; de aspecten die bij het Monty Hall Dilemma een rol spelen. Deze logica's komen ook aan bod in dit proefschrift.

In *kennislogica* ligt de nadruk op redeneringen waarbij kennis centraal staat. Redeneren over kennis is vooral interessant in situaties waarbij meerdere personen betrokken zijn. Dan kan iemand iets weten over wat iemand anders weet of niet weet. Dergelijke *hogere-orde kennis* speelt in veel situaties een cruciale rol. Zo begon tijdens de eerste Golfoorlog de operatie “Desert Storm” met de vernietiging van een aantal radarsystemen van Irak. Op die manier hadden de Verenigde Staten een voordeel boven Irak, omdat ze gebruik konden maken van hun kennis dat Irak bepaalde informatie niet meer had. Hoofdstuk 2 is een inleiding op de kennislogica.

Hoofdstuk 5 gaat over *probabilistische logica*. Hierbij richt men zich op redeneringen over waarschijnlijkheden. In filosofische debatten over waarschijnlijkheid worden theorieën, over de vraag wat waarschijnlijkheid is, in twee categorieën verdeeld. In de ene categorie wordt waarschijnlijkheid als deel van de werkelijkheid gezien. Als bijvoorbeeld een knikker uit een vaas met knikkers wordt gepakt is de waarschijnlijkheid dat de knikker groen is, de relatieve verhouding van groene knikkers. Bij de andere wordt waarschijnlijkheid gezien als iets wat de graad van vertrouwen van iemand uitdrukt. Als bijvoorbeeld een dobbelsteen onder een beker ligt, is de waarschijnlijkheid dat de uitkomst zes is de relatieve verhouding van zes in alle uitkomsten die de waarnemer voor mogelijk houdt. In het eerste geval worden waarschijnlijkheden toegekend aan objecten (aan iedere knikker wordt een waarschijnlijkheid toegekend om gepakt te worden), in het tweede geval aan situaties (aan iedere mogelijke uitkomst wordt een waarschijnlijkheid toegekend door de waarnemer). Ik laat zien hoe de logica’s die bij deze twee perspectieven horen aan elkaar gerelateerd zijn.

Dynamische logica spitst zich toe op redeneringen over veranderingen, zoals die bijvoorbeeld optreden in een computer ten gevolge van het uitvoeren van een computerprogramma. Men kan met dynamische logica bewijzen dat een computerprogramma doet wat het behoort te doen. Dat is voor veel programma’s van essentieel belang. In hoofdstuk 3 worden enkele technische resultaten over propositionele dynamische logica gepresenteerd. Met propositionele dynamische logica worden redeneringen over eenvoudige computerprogramma’s bestudeerd. Stel bijvoorbeeld dat gegeven is dat p geldt nadat het programma a wordt uitgevoerd. Bovendien is gegeven dat p ook geldt nadat het programma a twee keer wordt uitgevoerd. Enzovoorts. Uit al deze gegevens – oneindig veel – kan geconcludeerd worden dat p geldt nadat programma a willekeurig vaak herhaald wordt. In de gebruikelijke bewijssystemen voor deze logica kan dit echter niet bewezen worden omdat daarin slechts eindige redeneringen bestudeerd worden. In het bewijssysteem dat in hoofdstuk 3 gepresenteerd wordt, kunnen oneindige redeneringen wel bestudeerd worden.

Elk van de bovengenoemde logica’s richt zich dus op een van de aspecten die bij het Monty Hall Dilemma van belang zijn. De laatste jaren is grote interesse uitgegaan naar logica’s voor twee van deze aspecten, namelijk combinaties van dynamische logica en kennislogica. Dynamische kennislogica, waarmee redeneringen over verandering van kennis bestudeerd kunnen worden, wordt in hoofdstuk 4 behandeld. Dynamische kennislogica onderscheidt zich van andere vakgebieden die zich bezighouden met informatieverandering doordat de verandering van hogere-orde kennis centraal staat, dat wil zeggen dat rekening wordt gehouden met de verandering van kennis over andermans kennis.

Voor het Monty Hall Dilemma is echter een combinatie van kennislogica, probabilistische logica en dynamische logica nodig. Deze logica wordt ontwikkeld in hoofdstuk 6. Ik definieer een logische taal waarin redeneringen over zowel kennis, waarschijnlijkheden als verandering uitgedrukt kunnen worden. Tevens geef ik aan wanneer een uitspraak in die logische taal waar is en wanneer onwaar. Bovendien wordt een bewijssysteem verschaft dat bestaat uit axioma's en afleidingsregels. Hiermee kan stap voor stap worden aangetoond dat een redenering klopt. Deze logica is uitermate geschikt voor de analyse van redeneringen over kennis, waarschijnlijkheid en verandering.

De analyse van het Monty Hall Dilemma met een probabilistische dynamische kennislogica staat in hoofdstuk 7. De redenering, die als conclusie heeft dat de auto met een kans van twee derde achter de overgebleven deur zit die in eerste instantie niet gekozen werd, wordt vertaald naar de logische taal van probabilistische dynamische kennislogica. Met het bewijssysteem kan aangetoond worden dat deze redenering klopt.

References

Abadi, M. and J. Y. Halpern (1994). Decidability and expressiveness for first-order logics of probability. *Information and Computation 112*(1), 1–36.

Aczel, P. (1988). *Non-Well-Founded Sets*. Stanford, California: CSLI Publications.

Albers, C. J. (2003). *Distributional Inference: The Limits of Reason*. Ph. D. thesis, Institute for Mathematics and Computing Science, Groningen.

Albers, C. J., B. P. Kooi, and W. Schaafsma (2002). Trying to resolve the two-envelope problem. Submitted tot *Statistical Science*.

Alchourrón, C. E., P. Gärdenfors, and D. Makinson (1985). On the logic of theory change: partial meet functions for contraction and revision. *Journal of Symbolic Logic 50*, 510–530.

Aumann, R. J. (1976). Agreeing to disagree. *Annals of Statistics 4*(6), 1236 – 1239.

Bacchus, F. (1990). *Representing and Reasoning with Probabilistic Knowledge, A Logical Approach to Probabilities*. Cambridge, Massachusetts: The MIT Press.

Baltag, A. (2002). A logic for suspicious players: epistemic action and belief-updates in games. *Bulletin of Economic Research 54*(1), 1–45.

Baltag, A., L. S. Moss, and S. Solecki (1999). The logic of public announcements, common knowledge, and private suspicions. Technical Report SEN-R9922, CWI, Amsterdam.

Barwise, J. and J. Perry (1983). *Situations and Attitudes*. Cambridge, Massachusetts: MIT Press.

Bernardo, J. M., M. H. DeGroot, D. V. Lindley, and A. F. M. Smith (Eds.) (1988). *Bayesian Statistics 3, Proceedings of the Third Valencia International Meeting*, Oxford. Clarendon Press. held June 1–5 1987.

Binmore, K. (1992). *Fun and Games*. D.C. Heath and Company.

Blackburn, P., M. de Rijke, and Y. Venema (2001). *Modal Logic*, Volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge: Cambridge University Press.

Broome, J. (1995). The two-envelope paradox. *Analysis 55*(1), 6–11.

Chalmers, D. J. (1994). The two-envelope paradox: a complete analysis? `http://www.u.arizona.edu/~chalmers/papers/envelope.html`.

Chase, J. (2002). The non-probabilistic two envelope paradox. *Analysis 62*(2), 157–160.

Chvátal, V. (1983). Mastermind. *Combinatorica 3*, 325 –329.

Clark, M. and N. Shackel (2000). The two-envelope paradox. *Mind 109*(435), 415–442.

Cover, T. and J. Thomas (1991). *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley & Sons Inc.

de Vink, E. and J. Rutten (1999). Bisimulation for probabilistic transition systems: a coalgebraic approach. *Theoretical Computer Science 221*, 271–293.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics 38*, 325–329.

Dunin-Kęplicz, B. and R. Verbrugge (2002). Collective intentions. *Fundamenta Informaticae 51*(3), 271–295.

Emerson, E. A. (1990). Temporal and modal logic. In J. van Leeuwen (Ed.), *Handbook of theoretical computer science*, Volume B, Chapter 16, pp. 995–1072. Amsterdam New York Oxford Tokyo Cambridge, Massachusetts: Elsevier Science Publishers B.V. and MIT Press.

Fagin, R. and J. Halpern (1994). Reasoning about knowledge and probability. *Journal of the Association for Computing Machinery 41*(2), 340–367.

Fagin, R., J. Halpern, Y. Moses, and M. Vardi (1995). *Reasoning About Knowledge*. Cambridge, Massachusetts: MIT Press.

Fagin, R., J. Y. Halpern, and N. Megiddo (1990). A logic for reasoning about probabilities. *Information and Computation 87*(1,2), 78–128.

Fagin, R. and M. Y. Vardi (1985). An internal semantics for modal logic: preliminary report. *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, 305 – 315.

Flood, M. (1985-86). Mastermind strategy. *Journal of Recreational Mathematics 18*(3), 194–202.

Flood, M. (1988a). Sequential search sequences with Mastermind variants – part 1. *Journal of Recreational Mathematics 20*(2), 105 – 126.

Flood, M. (1988b). Sequential search sequences with Mastermind variants – part 2. *Journal of Recreational Mathematics 20*(3), 168 – 181.

Freudenthal, H. (1969). Formulering van het 'som-en-product'-probleem. *Nieuw Archief voor Wiskunde 17*, 152.

Freudenthal, H. (1970). Oplossing van het 'som-en-product'-probleem. *Nieuw Archief voor Wiskunde 18*, 102–106.

Gamow, G. and M. Stern (1958). *Puzzle-Math*. London: Macmillan.

Gamut, L. T. F. (1991). *Logic, Language, and Meaning*. Chicago and London: The University of Chicago Press.

Gärdenfors, P. and H. Rott (1995). Belief revision. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson (Eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Volume 4, Epistemic and temporal reasoning, pp. 35–132. Oxford: Clarendon Press.

Gardner, M. (1982). *aha! Gotcha: paradoxes to puzzle and delight*. New York: W.H. Freeman and Company.

Gerbrandy, J. (1999). *Bisimulations on Planet Kripke*. Amsterdam: ILLC Dissertation Series.

Gerbrandy, J. and W. Groeneveld (1997). Reasoning about information change. *Journal of Logic, Language, and Information 6*, 147 –196.

Gillies, D. (2000). *Philosophical Theories of Probability*. Philosophical issues in science. London and New York: Routledge.

Goldblatt, R. (1982). *Axiomatising the Logic of Computer Programming*, Volume 130 of *Lecture Notes in Computer Science*. Berlin, Heidelberg, New York: Springer-Verlag.

Goldblatt, R. (1992). *Logics of Time and Computation*, Volume 7 of *CSLI Lecture Notes*. Stanford, California: CSLI. Second edition, revised and expanded.

Goldblatt, R. (1993). *Mathematics of Modality*, Volume 43 of *CSLI Lecture Notes*. Stanford, California: CSLI Publications.

Gray, J. N. (1978). Notes on data base operating systems. In R. Bayer, R. M. Graham, and G. Seegmüller (Eds.), *Operating Systems: an advanced course*, Volume 60 of *Lecture Notes in Computer Science*, pp. 393–481. Berlin, Heidelberg, New York: Springer-Verlag.

Haack, S. (1978). *Philosophy of Logics*. Cambridge: Cambridge University Press.

Halpern, J. and M. Tuttle (1993). Knowledge, probability, and adversaries. *Journal of the Association for Computing Machinery 40*(4), 917–962.

Halpern, J. Y. (1990). An analysis of first-order logics of probability. *Artificial Intelligence 46*, 311–350.

Halpern, J. Y. (1991). The relationship between knowledge, belief, and certainty. *Annals of Mathematics and Artificial Intelligence 4*, 301–322.

Halpern, J. Y. (2001). Lexicographic probability, conditional probability, and nonstandard probability. In *Proceedings of the Eighth Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 17–30.

Harel, D., D. Kozen, and J. Tiuryn (2000). *Dynamic Logic*. Foundations of Computing. Cambridge, Massachusetts: MIT Press.

Hintikka, J. (1962). *Knowledge and Belief, An Introduction to the Logic of the Two Notions*. Ithaca & London: Cornell University Press.

Hirshleifer, J. and J. G. Riley (1992). *The analytics of uncertainty and information*. Cambridge surveys of economic literature. Cambridge: Cambridge University Press.

Hughes, G. E. and M. J. Cresswell (1996). *A New Introduction to Modal Logic*. London and New York: Routledge.

Irving, R. (1978-1979). Towards an optimum Mastermind strategy. *Journal of Recreational Mathematics 11*(2), 81 − 87.

Jackson, F., P. Menzies, and G. Oppy (1994). The two envelope 'paradox'. *Analysis 54*(1), 43–45.

Jeffrey, R. C. (1983). *The logic of decision*. Chicago London: University of Chicago Press.

Keisler, H. J. (1985). Probability quantifiers. In J. Barwise and S. Feferman (Eds.), *Model-Theoretic Logics*, Perspectives in Mathematical Logic, pp. 509–556. New York Berlin Heidelberg Tokyo: Springer-Verlag.

Knijnenburg, P. M. W. (1988). On axiomatisations for propositional logics of programs. Technical report, University of Utrecht. RUU-CS-88-34.

Knijnenburg, P. M. W. and J. van Leeuwen (1991). On models for propositional dynamic logic. *Theoretical Computer Science 91*, 181–203.

Knuth, D. (1976-1977). The computer as master mind. *Journal of Recreational Mathematics 9*(1), 1–6.

Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability* (second ed.). New York: Chelsea Publishing Company. Translation edited by Nathan Morrison with an added bibliography by A.T. Bharucha-Reid.

Kooi, B. P. (2003). Probabilistic dynamic epistemic logic. Accepted for the *Journal of Logic Language and Information*.

Koyoma, K. and T. Lai (1993). An optimal Mastermind strategy. *Journal of Recreational Mathematics 25*(4), 251 – 256.

Kozen, D. and R. Parikh (1981). An elementary proof of the completeness of PDL. *Theoretical Computer Science 14*, 113–118.

Kraitchik, M. (1943). *Mathematical Recreations*. London: George Allen & Unwin, Ltd.

Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics 22*, 79–86.

Larsen, K. and A. Skou (1991). Bisimulation through probabilistic testing. *Information and Computation 94*(1), 1–28.

Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.

Li, M. and P. M. B. Vitányi (1993). *An introduction to Kolmogorov complexity and its applications* (second ed.). Graduate texts in computer science. New York etc.: Springer.

Linsky, L. (Ed.) (1971). *Reference and modality*. Oxford readings in philosophy. London: Oxford University Press.

Linzer, E. (1994). The two envelope paradox. *The American Mathematical Monthly 101*(5), 417–419.

McCarthy, J. (1990). Formalization of two puzzles involving knowledge. In V. Lifschitz (Ed.), *Formalizing Common Sense : Papers by John McCarthy*, Ablex series in artificial intelligence. Norwood, N.J.: Ablex Publishing Corporation. Available online at `http://www-formal.stanford.edu/jmc/`.

McGrew, T. J., D. Shier, and H. S. Silverstein (1997). The two-envelope paradox resolved. *Analysis 57*(1), 28–33.

Meyer, J.-J. and W. van der Hoek (1995). *Epistemic Logic for AI and Computer Science*. Cambridge: Cambridge University Press.

Mirkowska, G. (1981). PAL-propositional algorithmic logic. In E. Engeler (Ed.), *Logic of Programs*, Volume 125 of *Lecture Notes in Computer Science*, pp. 23 – 101. Berlin Heidelberg New York: Springer-Verlag.

Mosteller, F. (1965). *Fifty Challenging Problems in Probability with Solutions*. Reading, Massachusetts: Addison-Wesley.

Nalebuff, B. (1988). Cider in your ear, continuing dilemma, the last shall be the first, and more. *Journal of Economic Perspectives 2*(2), 149–156.

Nalebuff, B. (1989). The other person's envelope is always greener. *Journal of Economic Perspectives 3*(1), 171–181.

Neuwirth, E. (1982). Some strategies for Mastermind. *Zeitschrift für Operations Research 26*, B257 –B278.

Nilsson, N. J. (1986). Probabilistic logic. *Artificial Intelligence 28*, 71–87.

Pratt, V. R. (1976). Semantical considerations on Floyd-Hoare logic. In *Proceedings of the 17th IEEE Symposium on the Foundations of Computer Science*, pp. 109–121.

Prohorov, Y. V. and Y. A. Rozanov (1969). *Probability Theory, basic concepts, limit theorems, random processes*, Volume 157 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete*. Berlin, Heidelberg, New York: Springer-Verlag. Translation of *Teorija Verojatnostej* by K. Krickeberg and H. Urmitzer.

Quine, W. V. O. (1956). Quantifiers and propositional attitudes. *Journal of Philosophy 53*, 177–187. Reprinted in (Linsky 1971).

Quine, W. V. O. (1980). *From a logical point of view : 9 logico-philosophical essays* (second, revised ed.). Cambridge, Massachusetts: Harvard University Press.

Rada, R. (1984). Mastermind in SIGART. *SIGART Newsletter 89*, 24 –25.

Rao, A. S. and M. P. Georgeff (1991). Modeling rational agents within a BDI-architecture. In R. Fikes and E. Sandewall (Eds.), *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*, pp. 473–484. Morgan Kaufmann Publishers: San Mateo, CA.

Rao, T., G. Kazin, and D. O'Brien (1986). Algorithms to play Mastermind. *SIGART Newsletter 95*, 33 − 35.

Reichenbach, H. (1949). *The Theory of Probability, an Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability* (second ed.). California library reprint series. Berkeley, Los Angeles, London: University of California Press. English translation by Ernest H. Rutten and Maria Reichenbach.

Renardel de Lavalette, G. (2002). Modal structures and changing modalities. Available online at `http://www.cs.rug.nl/~grl/`.

Renardel de Lavalette, G. R., B. P. Kooi, and R. Verbrugge (2002). Strong completeness for propositional dynamic logic. In P. Balbiani, N.-Y. Suzuki, and F. Wolter (Eds.), *AiML2002 — Advances in Modal Logic (conference proceedings)*, pp. 377–393. Institut de Recherche en Informatique de Toulouse IRIT.

Segerberg, K. (1977). A completeness theorem in the modal logic of programs. *Notices of the American Mathematical Society 24*(6), A–552.

Segerberg, K. (1994). A model existence program in infinitary propositional modal logic. *Journal of Philosophical Logic 23*, 337–367.

Segerberg, K. (1999). Two traditions in the logic of belief: bringing them together. In H. J. Ohlbach and U. Reyle (Eds.), *Logic, Language and Reasoning: essays in honour of Dov Gabbay*, Volume 5 of *Trends in Logic*, pp. 135–147. Dordrecht etc.: Kluwer Academic Publishers.

Selvin, S. (1975). A problem in probability. *The American Statistician 29*(1), 67.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal 27*, 379–423, 623–656.

Shapiro, E. (1983). Playing Mastermind logically. *SIGART Newsletter 85*, 28 – 29.

Smullyan, R. (1997). *The Riddle of Scheherazade, and other amazing puzzles, ancient and modern*. New York: Knopf.

Sterling, L. and E. Shapiro (1994). *The Art of Prolog: advanced programming techniques* (Second ed.). Cambridge, Massachusetts: MIT Press.

Sundholm, G. (1977). A completeness proof for an infinitary tense-logic. *Theoria 43*, 47–51.

Tanaka, Y. (2001). Model existence in non-compact modal logic. *Studia Logica 67*, 61–73.

ten Cate, B. D. (2002). Internalizing epistemic actions. In M. Martinez (Ed.), *Proceedings of the NASSLLI 2002 student session*, Stanford University, pp. 109 – 123.

van Benthem, J. F. A. K. (1984). Correspondence theory. In D. Gabbay and F. Guenther (Eds.), *Handbook of Philosophical Logic Volume II: extensions of classical logic*, Volume 165 of *Synthese Library Studies in epistemology, logic, methodology, and philosophy of science*, Chapter 4, pp. 167–247. Dordrecht Boston Lancaster: D. Reidel Publishing Company.

van Benthem, J. F. A. K. (1994). A note on dynamic arrow logic. In J. van Eijk and A. Visser (Eds.), *Logic and Information Flow*, Foundations of Computing, pp. 15–29. Cambridge, Massachusetts: MIT Press.

van Benthem, J. F. A. K. (2001). Games in dynamic epistemic logic. *Bulletin of Economic Research 53*(4), 219–248. Proceedings of LOFT-4, Torino.

van Benthem, J. F. A. K. (2002a). Conditional probability is update logic. unpublished.

van Benthem, J. F. A. K. (2002b). 'One is a lonely number': on the logic of communication. To be published in the proceedings of the *Logic Colloquium 2002*.

van Ditmarsch, H. P. (2000). *Knowledge Games*. Amsterdam: ILLC Dissertation Series DS-2000-06. Available online at `http://www.illc.uva.nl/Publications/`.

van Ditmarsch, H. P. (2002). Descriptions of game actions. *Journal of Logic, Language and Information 11*(3), 349–365.

van Ditmarsch, H. P., W. van der Hoek, and B. P. Kooi (2003). Concurrent dynamic epistemic logic. In V. F. Hendricks, K. F. Jørgensen, and S. A. Pedersen (Eds.), *Knowledge - Foundations and Applications*, Synthese Library Series. Kluwer Academic Publishers. To appear.

van Heijenoort, J. (1967). Logical paradoxes. In P. Edwards (Ed.), *The Encyclopedia of Philosophy*, Volume 5, pp. 45–51. New York: The Macmillan Company & The Free Press.

van Rooy, R. (2003). Quality and quantity of information exchange. *Journal of Logic Language and Information*. To appear.

Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic 25*, 221–261.

Vos Savant, M. (1990, September 9,). Ask Marilyn. *Parade Magazine*, 15.

Weatherford, R. (1982). *Philosophical Foundations of Probability Theory*. International Library of Philosophy. London, Boston, Melbourne and Henley: Routledge & Kegan Paul.

Wouters, P. (1991). Filosofische puzzel. *Parados 29*, 28 − 43.

Zabell, S. L. (1988a). Loss and gain: the exchange paradox. See Bernardo, DeGroot, Lindley, and Smith (1988), pp. 233–236. held June 1–5 1987.

Zabell, S. L. (1988b). Symmetry and its discontents. In B. Skyrms and W. L. Harper (Eds.), *Causation, Chance and Crecedence*. Kluwer Academic Publishers.

Notation

$\bot$: absurdity (generally equivalent to $p \wedge \neg p$)

$@_\mathsf{i}$: at operator

$\Box$: denotes the end of a proof or definition or the alethic modality

$\Box_a$: individual epistemic operator

$\Diamond_a$: dual of $\Box_a$ ($\Diamond_a \varphi$ is equivalent to $\neg\Box_a\neg\varphi$)

$[\![\cdot]\!]$: interpretation function

$[\cdot]$: dynamic modality

$\langle\cdot\rangle$: dual of $[\cdot]$

$\neg$: negation

$\wedge$: conjunction

$\bigwedge$: conjunction over a set

$\vee$: disjunction

$\bigvee$: disjunction over a set

$\rightarrow$: material implication, or to indicate the domain and range of a function

$\leftrightarrow$: double implication

$\mapsto$: in assigments: $g[x \mapsto d]$ means the assignment $(g \setminus (\{x\} \times D)) \cup \{(x, d)\}$.

$\underline{\leftrightarrow}$: bisimulation relation between models

$\cup$: union, or nondeterministic choice

$\cap$: intersection

$\setminus$: difference

$2^{\cdot}$: powerset

$\cdot \times \cdot$: cartesian product, or multiplication

$\#(\cdot)$: cardinality

$\restriction$: restriction of a function

$\mid$: separator in BNF, or conditional probability, or the extension of a set

$cert_a$: certainty operator

$\triangleleft$: dual of $\triangleright$

$\triangleright$: dyadic modality for DHEL

$*$: indicates the set of pointed models or frames

pre : precondition

a : agent

b : agent

c : agent, or individual constant

d : element of the domain

f : function (lean modal structure)

g : function (lean modal structure) (assignment)

h : function (lean modal structure)

i : index

j : index

k : index

n : index

p : proposition

q : rational number

r : run

s : state (local)

t : translation function, or type of an action

u : world or possibility

v : world or possibility

w : world or possibility

x : variable

y : variable

z : variable

A : a set

B : a set

C : common knowledge operator

D : domain

E : everybody knows operator

F : frame

I : interpretation

K : class of Kripke frames or models

L : learn operator

M : model

P : probability function

R : accessibility relation, or predicate symbol

S : set of (local) states, or sample space assignment, or predicate symbol

U : update operator

α : action

β : action

γ : action

λ : lambda abstractor

μ : probability measure

π : sequence of formulas or modalities

τ : term

φ : sentence or formula

χ : sentence or formula

ψ : sentence or formula

ω : the first transfinite ordinal

Γ : set of sentences or formulas

Δ : set of sentences or formulas

Λ : set of sentences or formulas

Σ : summation

Φ : sentence or formula

Ψ : sentence or formula

$\mathcal{A}$: set of agents

$\mathcal{B}$: group of agents

F : set of function symbols

$\mathcal{G}$: set of global states

$\mathcal{I}$: interpreted multi-agent system

$\mathcal{P}$: set of propositions

$\mathcal{R}$: multi-agent system

i : nominal

j : nominal

k : nominal

u : actionworld

v : actionworld

w : actionworld

A : action modality

C : set of constants

I : set of nominals

K : the minimal modal logic

R : action model relation, or set of predicate symbols

S : proof system

W : set of action worlds

X : set of variables

$\mathbf{D}$: nonempty access axiom

$\mathbf{E}$: everybody axiom

$\mathbf{K}$: the distribution axiom

$\mathbf{P}$: probability operator

$\mathbb{N}$: set of natural numbers

$\mathbb{Q}$: set of rational numbers

$\mathbb{R}$: set of real numbers

$\mathscr{L}$: a language

$\mathfrak{R}$: a bisimulation relation between worlds

Index

Titles in the ILLC Dissertation Series:

ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke

ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation

ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz

ILLC DS-1999-04: **Barbara Terhal**
Quantum Algorithms and Quantum Entanglement

ILLC DS-2000-01: **Renata Wassermann**
Resource Bounded Belief Revision

ILLC DS-2000-02: **Jaap Kamps**
A Logical Approach to Computational Theory Building (with applications to sociology)

ILLC DS-2000-03: **Marco Vervoort**
Games, Walks and Grammars: Problems I've Worked On

ILLC DS-2000-04: **Paul van Ulsen**
E.W. Beth als logicus

ILLC DS-2000-05: **Carlos Areces**
Logic Engineering. The Case of Description and Hybrid Logics

ILLC DS-2000-06: **Hans van Ditmarsch**
Knowledge Games

ILLC DS-2000-07: **Egbert L.J. Fortuin**
Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries

ILLC DS-2003-01: **Barteld Kooi**
Knowledge, chance, and change