

SEMANTIC BELIEF CHANGE

by

THOMAS ANDREAS MEYER

submitted in accordance with the requirements for
the degree of

DOCTOR OF PHILOSOPHY

in the subject

COMPUTER SCIENCE

at the

UNIVERSITY OF SOUTH AFRICA

PROMOTER: PROF J HEIDEMA

JOINT PROMOTER: DR WA LABUSCHAGNE

MARCH 1999

Abstract

The ability to change one’s beliefs in a rational manner is one of many facets of the abilities of an intelligent agent. Central to any investigation of belief change is the notion of an *epistemic state*. This dissertation is mainly concerned with three issues involving epistemic states:

1. How should an epistemic state be represented?
2. How does an agent use an epistemic state to perform belief change?
3. How does an agent arrive at a particular epistemic state?

With regard to the first question, note that there are many different methods for constructing belief change operations. We argue that semantic constructions involving ordered pairs, each consisting of a set of beliefs and an ordering on the set of “possible worlds” (or equivalently, on the set of basic independent bits of information) are, in an important sense, more fundamental.

Our answer to the second question provides indirect support for the use of semantic structures. We show how well-known belief change operations and related structures can be modelled semantically. Furthermore, we introduce new forms of belief change related operations and structures which are all defined, and motivated, in terms of such semantic representational formalisms. These include a framework for unifying belief revision and nonmonotonic reasoning, new versions of entrenchment orderings on beliefs, novel approaches to withdrawal operations, and an expanded view of iterated belief change.

The third question is one which has not received much attention in the belief change literature. We propose to extract extra-logical information from the formal representation of an agent’s set of beliefs, which can then be used in the construction of epistemic

states. This proposal is just a first approximation, although it seems to have the potential for developing into a full-fledged theory.

Keywords: Belief change, theory change, theory revision, belief revision, epistemic state, contraction, nonmonotonic reasoning, withdrawal, epistemic entrenchment, base change, base revision, base contraction.

Acknowledgements

On completion of such a time-consuming endeavour, there are always numerous people to thank. Notwithstanding the length of this dissertation, I would like to believe that I am a man of few words. As a result, and in keeping with the spirit of the principle of Informational Economy, I aim to keep the acknowledgements short and sweet.

Financial assistance by the Centre for Science Development for this research is hereby acknowledged. Opinions expressed, or conclusions reached in this publication, are those of the author and should not necessarily be ascribed to the Centre for Science Development.

To my supervisors, mentors, and good friends, Willem Labuschagne and Johannes Heidema, I would like to express my gratitude for the hours spent sitting in the tea room and listening to me. A special word of thanks to Arina Britz for many chats, and her insightful comments and thoughts. Thank you, also, to Horacio Arlo-Costa, Sven Ove Hansson, Daniel Lehmann, Isaac Levi, Maurice Pagnucco, Odile Papini, Pavlos Peppas and Hans Rott, for being kind enough to supply me with some of their research papers.

On a more personal note, to my best friend Louise, who also happens to be my wife, thank you for putting up with me for the past eighteen years, and for too many other things to mention here. To my family and in-laws who kept asking “Aren’t you finished yet?”, the answer is finally, “Yes, I am!”. And lastly, I cannot overemphasise the enormous influence that my two year old son Thomas has had on my work. Without his constant support and encouragement, this dissertation would have been completed at least a year earlier.

Contents

Abstract	i
Acknowledgements	iii
1 Introduction	1
1.1 A brief history of belief change	6
1.2 A reader’s guide	8
1.3 Formal preliminaries	10
2 AGM theory change	17
2.1 Postulates for AGM theory change	18
2.1.1 Connections between contraction and revision	20
2.2 Partial meet contraction	22
2.3 Epistemic entrenchment	24
2.3.1 Plausibility orderings	26
2.4 Safe contraction	27
3 Semantic belief change	31
3.1 Semantic content and infatoms	32
3.2 A semantics for theory change	39
3.2.1 The propositional finite case	45
3.2.2 Semantic AGM revision without smoothness	46
3.3 Orderings as epistemic states	47
3.3.1 Semantic epistemic entrenchment	48
3.3.2 The connection with relational partial meet contraction	50
3.3.3 Safe contraction	53
3.3.4 Summary	54

4	Nonmonotonic reasoning	59
4.1	KLM nonmonotonic reasoning	61
4.2	Preferential consequence relations	62
4.2.1	A semantics for preferential consequence relations	64
4.3	Rational consequence relations	66
4.4	Nonmonotonic reasoning as theory revision	71
4.4.1	Expectation based consequence relations	71
4.4.2	Expectations, beliefs and epistemic states	76
4.5	A dynamic view of nonmonotonic reasoning	79
4.6	Representing default information	81
4.7	Unifying cautious and bold reasoning	84
4.8	Conclusion	87
5	Epistemic entrenchment	89
5.1	AGM contraction via the EE-orderings	90
5.2	EE-orderings and minimality	93
5.3	Ordinal conditional functions	96
5.4	Generalised epistemic entrenchment	100
5.4.1	LR-entrenchment	101
5.4.2	GEE-entrenchment	103
5.5	Refined entrenchment	106
5.5.1	Refined entrenchment and the EE-orderings	110
5.5.2	Postulates for refined entrenchment	115
5.5.3	Refined entrenchment and AGM contraction	117
5.5.4	A comparison with generalised entrenchment	119
5.5.5	Refined G-plausibility	124
5.6	Other alternatives	127
5.7	Unifying epistemic and refined entrenchment	130
5.8	Summary	134
6	Withdrawal	139
6.1	To recover or not to recover	140
6.2	Basic withdrawal	143
6.2.1	Saturatable withdrawal	144
6.2.2	Sensible withdrawal	146

6.3	Principled withdrawal	150
6.3.1	Severe withdrawal	151
6.3.2	Systematic withdrawal	153
6.3.3	Revision-equivalence	154
6.3.4	Reasonable withdrawal	160
6.3.5	Systematic withdrawal vs. severe withdrawal	163
6.3.6	Informational value	171
6.4	Withdrawal and entrenchment	173
6.5	Systematic withdrawal and entrenchment	178
6.5.1	Systematic withdrawal and the EE-orderings	179
6.5.2	Systematic withdrawal and the RE-orderings	183
6.5.3	Representing systematic withdrawal graphically	190
6.6	Summary	197
7	Iterated belief change	201
7.1	Transmutation	202
7.2	AGM and iterated belief change	205
7.3	Iterated DP-revision	207
7.3.1	Minimal change	211
7.3.2	Conditional beliefs	212
7.3.3	Is iterated DP-revision rational?	214
7.3.4	Iterated DP-withdrawal	216
7.4	Iterated L-revision	222
7.5	Observation-based revision	225
7.5.1	P_{\triangleright} -revision	225
7.5.2	P_{\triangleleft} -revision	229
7.6	Merging epistemic states	231
7.6.1	Basic properties of merge operations	232
7.6.2	Constructing merge operations	233
7.7	Conclusion	238
8	Infobase change	239
8.1	Base change	240
8.2	Constructing infobase change	243
8.2.1	Infobase contraction	245

8.2.2	Properties of basic infobase contraction	253
8.2.3	Infobase contraction and reason maintenance	255
8.2.4	Infobase revision	256
8.3	Related approaches	258
8.3.1	Nebel's approach	259
8.3.2	Nayak's approach	261
8.4	Iterated infobase change	263
8.4.1	DP-revision	265
8.4.2	L-revision	270
8.5	Future research	271
9	Conclusion	275
9.1	Future research	280
A	Proofs of some results in chapter 3	283
A.1	Theorems 3.2.3 and 3.3.1	283
A.2	Results used in the proof of theorem 3.2.6	287
B	Proofs of some results in chapter 6	293
B.1	Results used in the proof of theorem 6.3.4	293
B.2	Theorems 6.5.12 and 6.5.14	295
C	List of identities	299
	Bibliography	303
	Index	318

Calvin and Hobbes c1984 Watterson. Reprinted with permission of Universal Press Syndicate. All rights reserved.

Chapter 1

Introduction

Of course not. After all, I may be wrong.

Bertrand Russell, on being asked whether
he would be prepared to die for his beliefs.

The comic strip on the opposite page concisely captures the central topic of this dissertation: that a rational intelligent agent is sometimes forced to adjust its current beliefs in some appropriate fashion when confronted with new information. The investigation of the reasoning patterns involved in such a task is known as the study of *belief revision* or *belief change*.

The ability to change one's beliefs in a manner that can be described as rational is one of many facets of the abilities of an intelligent agent. Central to the analysis of reasoning is the (somewhat nebulous) notion of an *epistemic state*. An epistemic state contains, in one form or another, the knowledge and beliefs of an agent, together with the information needed for coherent reasoning. This includes, in particular, the strategies for performing belief change. Our aim in this dissertation is to obtain a clear picture of the part of an epistemic state involving belief change. In doing so, it is necessary to draw a clear distinction between an agent's knowledge and its beliefs. We consider the beliefs of an agent to be the information that it is willing to act on, while knowledge comprises the beliefs that the agent refuses to retract; at least until some state change takes place. Belief, then, is defeasible knowledge, a view that is compatible with that of Moses and Shoham [1993].

This difference between belief and knowledge is also the basis of the difference between belief change and *belief update*. The former is concerned with changes to the epistemic state of an agent resulting from new information in a *static* world. In contrast, the latter deals with changes to epistemic states when the world described by it changes; changes to epistemic states in a *dynamic* world, if you will.¹ In our view, the knowledge of an agent can only be substantially altered once a state change has taken place (although knowledge can increase monotonically without any change in the current state). By concentrating on belief change, we operate under the assumption that the knowledge of an agent is fixed.

In the course of research into the area of belief change, two different (but not necessarily incompatible) approaches have begun to emerge; the *foundationalist* and *coherentist* approaches. The distinguishing feature of the foundationalist approach is that it assumes the existence of a set of *basic beliefs* which need no justification. All other beliefs in a foundational system have a justificatory pedigree. Every such a belief can be justified in terms of other beliefs, which in turn, can be justified in terms of other beliefs, until we eventually encounter the set of basic beliefs on which the original belief is ultimately based. The best known examples of foundational systems are Doyle's [1979, 1992] Truth Maintenance Systems and their successors, Reiter and de Kleer's [1987] Assumption Based Truth Maintenance Systems. Approaches to base change [Fuhrmann, 1991, Hansson, 1989, 1992b, 1993c, 1996] are also motivated by foundationalist ideas. The coherentist approach, on the other hand, sees the justification for beliefs in terms of the way they interact or "cohere" with other beliefs. In determining whether a belief is justified, one should thus look at its relationship with other beliefs.

A proper description of belief change demands that we specify an appropriate representational formalism. For our purposes, a certain family of logic languages with a propositional structure will be sufficient. More details can be found in section 1.3. For concreteness, the reader may think of a propositional language generated by a (possibly infinite) number of atoms, and equipped with a classical semantics. (See, for instance, Enderton [1972] or Fitting [1996].) The beliefs of an agent, as well as any information obtained, will be expressed in this language. The knowledge of an agent is equated with the sentences whose models establish the semantic framework within which belief

¹Keller and Wilkins [1985] first pointed out the distinction between belief change and belief update. Subsequently, Katsuno and Mendelzon [1992] formalised this distinction and proposed an abstract framework for belief update.

change occurs. This implements the view of knowledge as those beliefs that an agent refuses to retract.

Having chosen an appropriate language, we now turn to the three basic issues concerning belief change that we shall be addressing:

1. How should an epistemic state (or at least the part pertaining to belief change) be represented?

When addressing this question, observe that we are only concerned with that part of an epistemic state which involves belief change. When we talk about representing an epistemic state in a certain manner, it should be understood that such a representation can be *extracted* from the epistemic state.

We shall primarily be concerned with three representations of epistemic states; the second two being richer in structure than the first. The first representation is as a *belief set*, a set of sentences closed under logical entailment. Although such a representation contains too little information to be appropriate, it plays an important role in the establishment of abstract patterns and properties. As such, it is an extremely useful first approximation. The next representation is as an ordered pair, consisting of a belief set and an ordering on a set of “possible worlds” associated with the logic language under consideration (or equivalently, as a belief set and an ordering on the basic bits of information from which any set of beliefs is built up). As we shall see, such a view of epistemic states has proved to be a significant step forward in the study of belief change. Finally, epistemic states are often represented as ordered pairs consisting of a belief set and an epistemic entrenchment ordering on the sentences of the language under consideration. While such an ordering is, in a sense, equivalent to an ordering on possible worlds, we shall argue that the latter is a more fundamental construction. In doing so, we rely on the following principle:

(Reductionism) Complex objects are built up from simpler objects.

A consequence of the fact that these representations make use of belief sets, is the assumption that agents believe all the logical consequences of their beliefs. Levi [1991] refers to this as the agent’s epistemic commitment, and such agents are referred to as logically omniscient. In this sense, we provide an analysis of belief change on Newell’s [1982] *knowledge level*. Newell postulated the existence of a knowledge level above the

symbol level, on which there is no distinction between explicit information and derived information. This implies the satisfaction of Dalal's [1988] principle of the Irrelevance of Syntax:

(Irrelevance of Syntax) A belief change operation is independent of the form of the belief set involved.

The assumption of logical omniscience is clearly an idealisation, though, and future research on belief change will, no doubt, incorporate results on the resource-boundedness of agents.

Regardless of the way in which epistemic states are represented, however, it is important that the following principle is adhered to:

(Categorical Matching) A belief change operation performed on epistemic states should produce an epistemic state.

While this principle is almost too obvious to mention explicitly, much of the research on belief change has concentrated on operations that produce sets of beliefs, and not epistemic states. We now turn to the second issue.

2. How does an agent use an epistemic state to perform belief change?

Let us first make it clear that, although there are psychological studies which focus on the way human agents perform belief change and similar kinds of reasoning [Edwards, 1968, Einhorn and Hogarth, 1978, Ross and Lepper, 1980, Hoenkamp, 1988, Pelletier and Elio, 1997], our interest lies in the development of a normative account of belief change. That is, we are not (necessarily) concerned with the way in which human agents reason, but with the ways in which all rational agents ought to reason.

An answer to the question of how to use an epistemic state to perform belief change will, of course, depend on the specific belief change operation to be performed. Nevertheless, there are some basic principles underlying the appropriate use of information contained in epistemic states. The most important of these is the principle of Minimal Change [Harman, 1986].

(Minimal Change) Keep loss and addition to a minimum.

The basic idea is that the current epistemic state possesses a kind of inertia, and that any changes made to it ought to be only those that *have* to be made. We shall

also encounter two more specific versions of this principle, known as the principles of Informational Economy and Conservatism:

(Informational Economy) Keep the loss of information to a minimum.

(Conservatism) Keep the set of beliefs as large as possible.

One of the main obstacles to be encountered when attempting to satisfy principles such as these, is that there may not be a unique way in which to effect minimal change. In such an event, the use of the principle of Indifference is frequently advocated.

(Indifference) Objects held in equal regard should be treated equally.

A related principle, the principle of Preference, will play an important role in the analysis of *withdrawal*, an important form of belief change.

(Preference) Objects held in higher regard should be afforded a more favourable treatment.

This brings us to the third issue.

3. How does an agent arrive at a particular epistemic state?

This is a question which has not received much attention in the belief change literature, and there are, most probably, quite a number of angles from which it can be approached. We investigate one particular proposal in this regard. Our idea is to use finite ordered sequences of sentences to represent the beliefs of an agent. The structure of this representational formalism is then exploited to aid in the construction of epistemic states. On one level this is a violation of the principle of the Irrelevance of Syntax, but on other levels, this principle is still being respected. This proposal is not intended as a broad investigation into the question posed in point (3). It is just a first approximation, although it seems to have the potential for developing into a full-fledged theory.

Having outlined the three questions which we intend to address, it is perhaps necessary to mention that this dissertation does not contain a description of the computational aspects of belief change. This is not because we regard it as unimportant. On the contrary, the algorithmic and complexity-theoretic aspects of belief change is

perhaps *the* most important issue to be dealt with in future research. But, although some results dealing with these issues have recently begun to appear [Lehmann and Magidor, 1992, Gärdenfors and Rott, 1995, Goldszmidt and Pearl, 1996, Greiner, 1999], a more general picture has yet to emerge.

1.1 A brief history of belief change

The quest for a detailed theory of belief change is an old one. In 1907, for example, James [1907,p. 59] already gave a detailed description of a process by which we acquire new beliefs. Among contemporary researchers, Isaac Levi [1973, 1980, 1991, 1996] has been active in research related to belief change for three decades, and many of the ideas being developed today can be traced back to Levi's writings.

A major advance in the development of a detailed theory of belief change occurred during the first half of the 1980s. Known as the AGM approach to belief change, and named after its three originators, Carlos Alchourrón, Peter Gärdenfors and David Makinson, it was developed in a number of papers published in the late seventies and the beginning of the eighties [Gärdenfors, 1978, 1982, 1984, Alchourrón and Makinson, 1981, 1985, Alchourrón et al., 1985]. It forms the basis of most current research on belief change, including this dissertation.

AGM belief change is primarily concerned with three types of operations:

- A *removal* occurs when information is removed from the current set of beliefs of an agent.
- A *revision* occurs when new information is incorporated into the current set of beliefs in a way that ensures consistency.
- An *expansion* occurs when new information is simply added to the information currently in the set of beliefs, regardless of the consequences.

Expansion turns out to be non-problematic, and is defined by adding the new information to the agent's current set of beliefs, and then closing under logical consequence. It is the AGM proposals for removal and revision operations which have proved to be so influential. Actually, AGM belief change is mostly concerned with methods for constructing removal operations. This can be translated into the construction of revision operations by the application of one of Isaac Levi's ideas. Levi claims that the only

legitimate ways of transforming an epistemic state are expansion and removal [Levi, 1977], a view known as the *commensurability thesis* [Levi, 1991]. In this view, a revision by a sentence α consists of a removal of the negation of α , followed by an expansion with α .

AGM belief change is coherentist in nature. It subscribes to the principle of Minimal Change in that it strives to make the minimal changes necessary to an agent's set of beliefs following a change operation. Ironically, their version of revision, which is defined in terms of removal, has been accepted enthusiastically, while their proposal for removal, known as *contraction*, has met with some resistance [Makinson, 1987, Fuhrmann, 1991, Lindström and Rabinowicz, 1991, Niederee, 1991, Hansson, 1991, 1992a, 1993c, 1996]. In recent years, there have been a number of proposals for constructing removal operations that retain the advantages of AGM contraction without suffering from its disadvantages [Levi, 1991, 1998, Hansson and Olsson, 1995, Rott and Pagnucco, 1999, Cantwell, 1999, Fermé, 1998, Fermé and Rodriguez, 1998].

Since AGM belief change is coherentist in nature, it is concerned with sets of beliefs closed under logical consequence without any justificatory structure. It has been argued though [Alchourrón and Makinson, 1982, Makinson, 1985, Hansson, 1989, 1992b, Fuhrmann, 1991], that some of our beliefs have no independent standing, but arise only as inferences from our basic beliefs. This foundationist view has led to the development of a generalisation of AGM belief change, known as *base change*, in which the emphasis is placed on changes made to the set of basic beliefs of an agent. While such an approach accommodates the idea of basic beliefs, it is, to a large extent, in violation of the principle of the Irrelevance of Syntax. It can thus be seen as operating on the symbol level, thereby forfeiting an important characteristic of AGM belief change: an analysis of belief change on the knowledge level. Interestingly enough though, it turns out that by relaxing some widely held assumptions about base change, it is indeed possible to provide a knowledge level description of base change [Nebel, 1989, 1990, 1991, 1992].

AGM belief change places the emphasis on sets of beliefs closed under logical consequence. As such, it is in violation of the principle of Categorical Matching. For, although it needs more than just a set of beliefs to perform revision and contraction, it concentrates only on the sets of beliefs obtained when performing such change operations. This is one of the reasons why AGM belief change is not able to provide a proper account of iterated belief change, the description of the process of performing

sequences of changes. One of the most important contributions to the enterprise of belief change in recent years concerns the realisation that belief change ought to be described on the level of epistemic states, and not on the level of belief sets. This has led to an influential proposal by Darwiche and Pearl [1994, 1997] for a framework for iterated belief change. Their proposal has served to highlight the semantic methods for constructing AGM-style belief change operations, for it presupposes the existence of the semantic structures used for constructing AGM belief change.

Finally, in this brief discussion we have concentrated on revision and removal, but AGM belief change has also been the inspiration for a number of other types of belief change such as relational change [Lindström and Rabinowicz, 1991], multiple change [Fuhrmann and Hansson, 1994, Peppas and Sprakis, 1999] and multi-agent belief change [Kfir-Dahav and Tennenholtz, 1996], with merging [Borgida and Imielinski, 1984, Baral et al., 1991, 1992, Subrahmanian, 1994, Liberatore and Schaerf, 1998, Konieczny and Pino-Pérez, 1998] as a special case of the latter.

1.2 A reader's guide

The next two chapters are mainly concerned with classic AGM belief change. Chapter 2 is a survey of AGM belief change, containing sets of rationality postulates for revision and contraction, as well as a description of the primary methods used in the construction of such operations. The one aspect which is missing from this chapter is a discussion of the semantic modellings of AGM belief change. We regard the latter as important enough to devote the whole of chapter 3 to it. The semantic construction methods discussed in chapter 3 form the cornerstone of the results presented in the rest of the dissertation. Besides the well-known semantic modellings in terms of orderings on the interpretations of the logic language under consideration, we propose that an information-theoretic semantics be used, with orderings on the basic bits of information available to an agent. While such a semantics has very strong formal links with the traditional possible-worlds semantics (they are dual to each other in a sense that will be made precise in propositions 3.1.5 and 3.1.6), we contend that the information-theoretic view is of use in the intuitive justification of such semantic constructive methods. By summarising well-known results about AGM belief change, we show that the methods discussed in chapter 2, for constructing AGM belief change operations, can all be seen as being obtained from the semantic method of construction. Although this is to be

expected, given the various representation results linking these methods to AGM belief change, it is difficult to escape the conclusion that the semantic modelling is more “basic”, in a sense.

Chapter 4 explores the relationship between belief change and nonmonotonic reasoning. As has been noted, the influential semantic approaches of Kraus et al. [1990] and Lehmann and Magidor [1992] to nonmonotonic reasoning have much in common with the semantic construction of belief change. So much so, in fact, that it has been claimed that the processes involved in belief revision and nonmonotonic reasoning are the same, although used for different purposes [Gärdenfors and Makinson, 1994]. We show that these two areas can be unified into a more general theory of bold and cautious reasoning. Furthermore, applying the view of belief change as a dynamic process to nonmonotonic reasoning, we argue that most approaches to nonmonotonic reasoning operate under the implicit assumption that obtaining new pieces of evidence sequentially is equivalent to obtaining them simultaneously; a view that seems too strong to be appropriate for a general theory of nonmonotonic reasoning.

Chapter 5 is concerned with one of the standard methods for constructing AGM belief change; in terms of epistemic entrenchment orderings on sentences of the logic language to be used. We provide a survey of the field, and present a new form of entrenchment — termed *refined entrenchment* — which does not suffer from the same drawbacks as the best known form of epistemic entrenchment [Gärdenfors and Makinson, 1988, Gärdenfors, 1988]. Refined entrenchment is defined semantically, but it can also be characterised in terms of a set of rationality postulates. Chapter 6 is devoted to the study of a family of removal operations which are intended as alternatives to AGM contraction; the withdrawal operations. We survey the field, and propose the addition of a new member of this family, known as *systematic withdrawal*. It is defined semantically (in terms of the same set of orderings used to define refined entrenchment), but can also be characterised in terms of a set of rationality postulates. Systematic withdrawal seems to retain the advantages of most forms of withdrawal, while not being subject to their disadvantages. Some of the results in chapters 5 and 6 suggest the use of a general set of orderings on interpretations which can be used to construct a wide variety of entrenchment orderings and withdrawal operations.

In chapter 7 we investigate the issue of iterated belief change. We discuss the frameworks recently proposed by Darwiche and Pearl [1994, 1997] and Lehmann [1995], as well as the work of Spohn [1988] (which had served as inspiration for Darwiche

and Pearl), and its generalised version in the form of the transmutations proposed by Williams [1994]. Both of the proposed frameworks for iterated belief change operate on the level of epistemic states, and both concentrate on revision. We show that the extension of the proposal of Darwiche and Pearl to various forms of withdrawal casts some doubt on the desirability of some of the restrictions they impose, and we measure a recently proposed version of revision [Papini, 1998, 1999] against these frameworks. Inspired by the semantic approach of Darwiche and Pearl, as well as by the work of Nayak [1994b], Nayak et al. [1996] and Liberatore and Schaerf [1998], we regard the *merging* of epistemic states as a fruitful area for future research.

Chapter 8 is an attempt at solving a problem that has not received its fair share of attention in the belief change literature; determining how an agent arrives at a particular epistemic state. Our proposal is to represent the information obtained by an agent as ordered sequences of sentences, with each one being seen as a piece of information (or an observation) obtained from an independent source. Such a sequence of sentences is referred to as an infobase. The structure of infobases is used to induce the semantic structures necessary for performing belief change. This process determines the appropriate belief set resulting from a change operation, thereby operating on the knowledge level. A second phase then determines the infobase resulting from the change operation by weakening the sentences contained in the original infobase. While infobase change can be compared with traditional approaches to base change [Fuhrmann, 1991, Hansson, 1989, 1992b, 1993c, 1996], it has more in common with the pseudo-contraction operations [Hansson, 1999, p. 334] of Nebel [1989, 1990, 1991, 1992].

Finally, chapter 9 summarises the results presented and points to open problems and future research. As the title suggests, the main thesis defended in this dissertation is that semantic approaches to belief change have proved to be most fruitful in the past, and will continue to play such a role in future.

1.3 Formal preliminaries

In our investigation of belief change we assume a formal object language L in which the beliefs of an agent are expressed. We take L to be closed under the usual propositional connectives \neg , \wedge , \vee , \rightarrow , \leftrightarrow , and to contain the symbols \top and \perp . The well-formed formulas (wffs) of L will be denoted by lower-case Greek letters. Furthermore, we assume L to be equipped with a two-valued model-theoretic semantics defining truth

and falsity. A (*possible-worlds*) *semantics* for L is thus an ordered pair (U, \models) , where U is a (non-empty) set of *interpretations* of L , and \models is a *satisfaction* relation for L . That is, for the relation \models from U to L , $u \models \alpha$ means that α is true in u , or that u satisfies α . Elements of U will be denoted by lower-case Latin letters. The satisfaction relation \models is required to behave classically with respect to the propositional connectives. We use \top and \perp as canonical representatives for the logically valid and logically invalid wffs respectively. The set of *models* $M(A)$ of any set of wffs A is the set of interpretations satisfying all the wffs in A . That is $M(A) = \{u \in U \mid \forall \alpha \in A, u \models \alpha\}$. For $\alpha \in L$ we write $M(\alpha)$ instead of $M(\{\alpha\})$. We refer to the set $U \setminus M(A)$ as the *countermodels* of A .

Such a semantics allows us to define the notion of *semantic entailment* in the standard manner. Formally, semantic entailment is a binary relation from $\wp L$ (the powerset of L) to L , and is defined as follows: $A \models \beta$ iff $M(A) \subseteq M(\beta)$. For $\alpha \in L$ we write $\alpha \models \beta$ instead of $\{\alpha\} \models \beta$, and we abbreviate $\emptyset \models \beta$ as $\models \beta$. Intuitively, $A \models \alpha$ means that α follows logically from A . The only requirement that we place on \models is that it satisfies *compactness*: $A \models \alpha$ iff $A_F \models \alpha$ for some finite subset A_F of A . The entailment relation \models is associated with a consequence operation Cn . Formally, Cn is a unary consequence operation on $\wp L$, and is defined in terms of \models as follows: $Cn(A) = \{\alpha \mid A \models \alpha\}$. So, intuitively, $Cn(A)$ consists of all the beliefs that follow logically from A . Whether one uses Cn or \models is a matter of preference and convenience, since it is clear that \models can also be defined in terms of Cn as follows: $A \models \alpha$ iff $\alpha \in Cn(A)$.

It should be obvious that the logics we consider include all classical propositional logics and classical first-order logics (with the first-order languages restricted to closed wffs). In fact, every logic we consider can be “converted” into a propositional logic, based on a propositional language PL , in the following sense. Define the set of atoms A_{PL} of the propositional language PL in terms of L as $A_{PL} = L \setminus N_{PL}$, where

$$N_{PL} = \left\{ \alpha \in L \left| \begin{array}{l} \alpha = \perp, \alpha = \top, \alpha = \neg\beta, \text{ or } \alpha = \beta \diamond \gamma, \\ \text{where } \diamond \in \{\vee, \wedge, \leftarrow, \leftrightarrow\} \text{ and } \beta, \gamma \in L \end{array} \right. \right\}.$$

So A_{PL} is the set of wffs of L *not* having one of the propositional connectives as its main connective. We refer to A_{PL} as the *propositional atoms* of L . Now, for every interpretation $u \in U$ we define a *valuation* $val_u : A_{PL} \rightarrow \{F, T\}$ as: $val_u(\alpha) = T$ iff $u \models \alpha$, and we let the set of valuations V of PL be $V = \{val_u \mid u \in U\}$. A satisfaction relation \models_V from V to PL is then obtained recursively from V in the standard way:

1. For every $v \in V$, $v \Vdash_V \top$ and $v \nVdash_V \perp$.
2. If $\alpha \in A_{PL}$, then $v \Vdash_V \alpha$ iff $v(\alpha) = T$.
3. If $\alpha = \neg\beta$ then $v \Vdash_V \alpha$ iff $v \nVdash_V \beta$.
4. If $\alpha = \beta \vee \gamma$ then $v \Vdash_V \alpha$ iff either $v \Vdash_V \beta$, or $v \Vdash_V \gamma$, or both.
5. If $\alpha = \beta \wedge \gamma$ then $v \Vdash_V \alpha$ iff both $v \Vdash_V \beta$ and $v \Vdash_V \gamma$.
6. If $\alpha = \beta \rightarrow \gamma$ then $v \Vdash_V \alpha$ iff either $v \nVdash_V \beta$, or $v \Vdash_V \gamma$, or both.
7. If $\alpha = \beta \leftrightarrow \gamma$ then $v \Vdash_V \alpha$ iff either both $v \Vdash_V \beta$ and $v \Vdash_V \gamma$, or both $v \nVdash_V \beta$ and $v \nVdash_V \gamma$.

Since the languages PL and L are identical, the set of valuations V also provides an acceptable semantics for L . This is easily verified by observing that the semantics (V, \Vdash_V) for L generates exactly the same entailment relation \models as the semantics (U, \Vdash) for L . (Actually, the inclusion of \Vdash_V is redundant, since it can be obtained from V .) In fact, in any equivalence class containing every semantics for L that generates the same entailment relation \models , (V, \Vdash_V) occupies a unique position, since it is the only semantics (up to isomorphism) without elementarily equivalent interpretations. (Two interpretations x, y are *elementarily equivalent* iff they satisfy exactly the same wffs of L . That is, $x \Vdash \alpha$ iff $y \Vdash \alpha$, for every $\alpha \in L$.) We shall refer to (V, \Vdash_V) as the \models -*valuation semantics* for L . In general, we refer to a semantics (V, \Vdash) for L in which V is a set of valuations, and in which \Vdash is obtained from V in the manner described above, as a *valuation semantics* for L .

In the belief change literature it is not standard practice to start with a semantic description. Instead, L usually comes equipped with an abstract consequence relation, denoted by the single turnstile \vdash , in place of the semantic entailment relation. The consequence operation Cn is defined in terms of \vdash , and Cn is assumed to satisfy the following properties:

(Inclusion) $A \subseteq Cn(A)$

(Idempotence) $Cn(Cn(A)) \subseteq Cn(A)$

(Monotonicity) If $A \subseteq B$ then $Cn(A) \subseteq Cn(B)$

(Supraclassicality) $Cn(A)$ includes every truth-functional tautology and satisfies Modus Ponens

(Deduction theorem) $\beta \in Cn(A \cup \{\alpha\})$ iff $\alpha \rightarrow \beta \in Cn(A)$

(Compactness) $\alpha \in Cn(A)$ iff $\alpha \in Cn(A_F)$ for some finite subset A_F of A

It is easy to show that, with the exception of a single pathological case, the entailment relations we consider are precisely the semantic versions of these consequence relations. To see why, note firstly that the consequence operation Cn associated with every entailment relation \models we consider, clearly satisfies the six properties outlined above. And conversely, from every consequence relation \vdash whose associated consequence operation Cn satisfies these six properties (except the trivial one for which $\vdash = \wp L \times L$), we can construct an appropriate semantics for L that will satisfy all the requirements set out above.² Simply take U , the set of interpretations of L , to be the set of maximally consistent subsets of L . That is, let

$$U = \{A \subseteq L \mid A \not\vdash \perp \text{ and } \forall B \subseteq L \text{ such that } A \subset B, B \vdash \perp\}.$$

The satisfaction relation \models is then defined as follows: $A \models \alpha$ iff $\alpha \in A$. It is readily verified that the semantic entailment relation \models obtained from \models behaves exactly like \vdash . Note also that the semantics obtained in this way is isomorphic to the \models -valuation semantics for L .³

A *theory* or a *belief set* is a set $K \subseteq L$ closed under entailment, i.e. for which $K = Cn(K)$. A set $X \subseteq L$ *axiomatises* a belief set K iff $Cn(X) = K$, and X *finitely axiomatises* K iff X is finite. For every $W \subseteq U$, the *theory determined by* W is

$$Th(W) = \{\alpha \in L \mid W \subseteq M(\alpha)\},$$

and for $u \in U$ we write $Th(u)$ instead of $Th(\{u\})$. A set $A \subseteq L$ *axiomatises* a set of interpretations W iff $M(A) = W$. A set $A \subseteq L$ is *satisfiable* iff $M(A) \neq \emptyset$, iff

²The trivial consequence relation $\vdash = \wp L \times L$ can be obtained from a possible-worlds semantics (U, \vdash) with $U = \emptyset$.

³The single consequence relation \vdash for which we cannot obtain a corresponding semantics is the trivial one for which every wff follows from every set of wffs, defined as $\vdash = \wp L \times L$. It is easily verified that such a \vdash satisfies the six properties above, but allows for no maximally consistent subsets. Consequently, the set U of interpretations obtained from \vdash will be empty, something that is not permitted by our definition of a semantics.

$Cn(A) \neq Cn(\perp)$. For every satisfiable subset A of L , $\alpha \in L$ is *A-established* (or *A-believed*) iff $A \models \alpha$, α is *A-undecided* (or *A-neutral*) iff $A \not\models \alpha$ and $A \not\models \neg\alpha$, and α is *A-refuted* (or *A-disbelieved*) iff $A \models \neg\alpha$. For an unsatisfiable subset A of L , all the wffs of L are *A-established*, while none are *A-undecided* or *A-refuted*.

The use of the following abbreviations will be convenient. By $\alpha \equiv \beta$ we understand that α and β are logically equivalent, i.e. $\alpha \models \beta$ and $\beta \models \alpha$. For every finite $A, B \in \wp L$ we write $A \diamond B$ as an abbreviation for $\{\alpha \diamond \beta \mid \alpha \in A \text{ and } \beta \in B\}$ where $\diamond \in \{\vee, \wedge\}$, $\neg A$ as an abbreviation for $\{\neg\alpha \mid \alpha \in A\}$, $\bigwedge A$ as an abbreviation for the conjunction of all elements in A , with $\bigwedge \emptyset = \top$, and $\bigvee A$ as an abbreviation for the disjunction of all elements in A , with $\bigvee \emptyset = \perp$. For a belief set K and a wff $\alpha \in L$, the *expansion* of K by α is defined as $K + \alpha = Cn(K \cup \{\alpha\})$.

A binary relation R on any set X is *connected* iff xRy or yRx for every $x, y \in X$. A preorder \sqsubseteq (i.e. a reflexive and transitive binary relation) on a set X that is also connected is called a *total preorder*. For any preorder \sqsubseteq on a set X , we write $x \sqsubset y$ iff $x \sqsubseteq y$ and $y \not\sqsubseteq x$, $x \equiv_{\sqsubseteq} y$ iff $x \sqsubseteq y$ and $y \sqsubseteq x$, $x \parallel_{\sqsubseteq} y$ iff $x \not\sqsubseteq y$ and $y \not\sqsubseteq x$, and we let $[x]_{\sqsubseteq} = \{y \mid x \equiv_{\sqsubseteq} y\}$. For every non-empty $Y, Z \subseteq X$, we write $Y \sqsubseteq Z$ iff $y \sqsubseteq z$ for every $y \in Y$ and $z \in Z$, $Y \sqsubset Z$ iff $y \sqsubset z$ for every $y \in Y$ and $z \in Z$, and $Y \equiv_{\sqsubseteq} Z$ iff $y \equiv_{\sqsubseteq} z$ for every $y \in Y$ and $z \in Z$. And as a limiting case, we set $\emptyset \sqsubset Y$ for every non-empty $Y \subseteq X$.

Our examples are usually phrased in propositional languages (containing the usual propositional connectives) that are generated by at most three atoms. We use the letters p, q and r to denote these atoms, and interpretations (or rather valuations) of the languages will be represented by appropriate sequences of 0s and 1s, 0 representing falsity and 1 representing truth. The convention is that the first digit in the sequence represents the truth value of p , the second the truth value of q and the third the truth value of r .

Sometimes it will be convenient to use *transparent propositional languages* in our examples. These are restricted versions of first-order languages containing no variables and no quantifiers. The propositional atoms of such a language are then simply the first-order atoms that can be formed from the available predicate symbols and terms. For example, suppose that L is a transparent propositional language containing the predicate symbols b and f and the two constant symbols t and c . Then L is generated from the four propositional atoms $b(t)$, $b(c)$, $f(t)$ and $f(c)$.

For the reader's convenience, we provide (without proof) the following well-known

model-theoretic results about the kind of semantic setup we consider.

Proposition 1.3.1 *Let (U, \models) be a possible-worlds semantics for L .*

1. *If $\alpha \equiv \beta$ then $M(\alpha) = M(\beta)$.*
2. *$M(\top) = U$.*
3. *$M(\perp) = \emptyset$.*
4. *$u \in M(\alpha)$ iff $u \notin M(\neg\alpha)$.*
5. *$M(\alpha \wedge \beta) = M(\alpha) \cap M(\beta)$.*
6. *$M(\alpha \vee \beta) = M(\alpha) \cup M(\beta)$.*
7. *For every $W \subseteq U$, $W \subseteq M(Th(W))$.*

Proposition 1.3.2 *Let L be a finitely generated propositional language and let (V, \models) be a valuation semantics for L .*

1. *For every $W \subseteq V$, $W = M(Th(W))$.*
2. *For every $W \subseteq V$ there is an $\alpha_W \in L$ such that $M(\alpha_W) = W$. That is, every set of valuations can be axiomatised by a wff of L .*

The following model-theoretic results will also prove to be most useful.

Lemma 1.3.3 *Suppose that K is a belief set and that $W \subseteq M(\neg\alpha)$. Then*

$$(M(Th(M(K) \cup W)) \setminus M(K)) \subseteq M(\neg\alpha).$$

Proof We only consider the case where $W \neq \emptyset$, and there is thus a $w \in W$ such that $w \in M(\neg\alpha)$. Let $X = M(Th(M(K) \cup W)) \setminus M(K)$, suppose that $x \in X$ and assume that $x \in M(\alpha)$. So there is a $\beta \in K$ such that $x \notin M(\beta)$, and then $\alpha \rightarrow \beta \in Th(M(K) \cup W)$ (since $\alpha \rightarrow \beta \in K$ and $W \subseteq M(\neg\alpha)$). But this means that $x \in M(\alpha \rightarrow \beta)$, contradicting the fact that $x \in M(\alpha)$ and $x \notin M(\beta)$. \square

Lemma 1.3.4 *Suppose that K is a belief set, $\alpha \in K$, $W \subseteq M(\neg\alpha)$, and*

$$X = M(Th(M(K) \cup W)) \setminus M(K).$$

Then $Th(W) = Th(X)$.

Proof Since $W \subseteq X$, it suffices to show that for every $\beta \in L$, if $W \subseteq M(\beta)$ then $X \subseteq M(\beta)$. So pick a $\beta \in L$ and suppose that $W \subseteq M(\beta)$. Then $\neg\beta \rightarrow \alpha \in Th(M(K) \cup W)$, since $\neg\beta \rightarrow \alpha \in K$ and $W \subseteq M(\beta)$. Therefore $X \subseteq M(\neg\beta \rightarrow \alpha)$, and then $X \subseteq M(\beta)$ since $X \subseteq M(\neg\alpha)$ by lemma 1.3.3. \square

Lemma 1.3.5 *Let K be a belief set, and suppose that $X \subseteq M(\neg\alpha)$ and $W \subseteq M(\alpha)$. Then $M(Th(M(K) \cup X \cup W)) \cap M(\neg\alpha) = M(Th(M(K) \cup X)) \cap M(\neg\alpha)$.*

Proof We only consider the case where $\neq \alpha$. Assume that the left-to-right inclusion does not hold. So there is a $u \in M(Th(M(K) \cup X \cup W)) \cap M(\neg\alpha)$ such that $u \notin M(Th(M(K) \cup X)) \cap M(\neg\alpha)$. There is thus a β such that $(M(K) \cup X) \subseteq M(\beta)$, but $u \notin M(\beta)$. Now observe that $\neg\beta \rightarrow \alpha \in Th(M(K) \cup X \cup W)$. But this means that $u \in M(\neg\beta \rightarrow \alpha)$, contradicting the fact that $u \in M(\neg\beta)$ and $u \in M(\neg\alpha)$. The proof for the right-to-left inclusion is similar. \square

Chapter 2

AGM theory change

*But O the heavy change, now thou art gone,
Now thou art gone, and never must return!*

John Milton, Lycidas, 37

One of the most influential contributions to the study of belief change is that of Alchourrón, Gärdenfors and Makinson — the so-called AGM approach to theory change — developed in a number of papers in the late 1970s and 1980s [see Gärdenfors, 1978, 1982, 1984, Alchourrón and Makinson, 1981, 1985, Alchourrón et al., 1985]. Even though it is mainly concerned with belief sets, it has become a benchmark against which to test and compare (whether directly or indirectly) a wide variety of belief change operations. AGM theory change takes the epistemic state of an agent to be a belief set [Gärdenfors, 1988, p. 47], and aims to give a description of the permissible changes to a belief set resulting from the revision by, or the removal of, a single wff.¹ This is accomplished in terms of two sets of rationality postulates. Formally, we assume a fixed belief set K , defining (belief) removal and revision pertaining to K as functions from L to $\wp L$. Where there is no ambiguity, we shall drop the references to K . (In later chapters it will be necessary to view change operations differently, as functions from $Bel \times L$ to $\wp L$, where Bel is the set of all belief sets.) By an α -removal, α -revision, α -contraction, and so on, we mean a removal of α from K , revision of K by α , a contraction of K by α , and so forth.

¹Although the original AGM papers are not exclusively concerned with belief sets, the major results in [Alchourrón et al., 1985] only hold for belief sets.

By now a whole array of methods have been developed for constructing AGM theory change. In this chapter we briefly discuss three classical ways of doing so. When removing a wff α from a belief set K , *partial meet contraction* does so using the maximal subsets of a belief set K not entailing a wff α , *safe contraction* employs minimal subsets of K that entail α , and *epistemic entrenchment* makes use of an ordering of relative entrenchment on wffs. Our treatment of AGM theory change in this chapter cannot be regarded as complete, primarily because it does not contain a discussion of semantic approaches to theory change. We regard the latter as important enough to devote the whole of chapter 3 to it.

2.1 Postulates for AGM theory change

AGM theory change is concerned with a whole spectrum of rational ways to perform belief change, and does not provide unique definitions for revision and removal. Instead, a number of postulates are provided with which all removals and revisions are required to comply. The idea is that these are the rational choices to be made. As we have seen in chapter 1 on page 7, Levi's commensurability thesis views removal as more primitive than revision, and it is thus appropriate that we start with the AGM postulates for belief removal.

$$(K-1) \quad K - \alpha = Cn(K - \alpha)$$

$$(K-2) \quad K - \alpha \subseteq K$$

$$(K-3) \quad \text{If } \alpha \notin K \text{ then } K - \alpha = K$$

$$(K-4) \quad \text{If } \not\models \alpha \text{ then } \alpha \notin K - \alpha$$

$$(K-5) \quad \text{If } \alpha \equiv \beta \text{ then } K - \alpha = K - \beta$$

$$(K-6) \quad \text{If } \alpha \in K \text{ then } (K - \alpha) + \alpha = K$$

Definition 2.1.1 A removal is a *basic AGM contraction* iff it satisfies (K-1) to (K-6). We refer to these six postulates as the *basic AGM contraction postulates*. \square

The first five contraction postulates together constitute little more than an obvious expression of the intuition that the AGM trio associate with belief removal. (K-1)

is the requirement that AGM contraction operate on belief sets, while (K-2) ensures that the contraction of a belief set actually results in a contracted belief set. (K-3) is a straightforward appeal to the principle of Informational Economy in the pathological case of contraction by a wff that is not in the belief set to begin with. (K-4) ensures that contraction by any wff other than a logically valid one is successful, and (K-5) is a formalisation of the principle of the Irrelevance of Syntax. This brings us to (K-6), the postulate also known as Recovery. It was originally phrased as follows:

$$(K-6') \quad K \subseteq (K - \alpha) + \alpha$$

but it is easily verified that these two formulations are equivalent in the presence of (K-1), (K-2), and (K-3).

The Recovery postulate is an expression of the principle of Informational Economy. It requires of a contraction of K by a wff $\alpha \in K$ to retain so much of K , that it is possible to recover the whole of K from an α -expansion of the resulting belief set. The desirability of the Recovery postulate is a contentious issue and has evoked a vigorous debate [see Makinson, 1987, 1997, Hansson, 1991, 1993c, 1996, Levi, 1991, Lindström and Rabinowicz, 1991, Niederee, 1991]. We take up the matter in chapter 6, where we discuss belief removals that satisfy all the basic AGM contraction postulates except for Recovery. In accordance with a suggestion by Makinson [1987], we refer to such removals as withdrawals.

Definition 2.1.2 A removal is a *withdrawal* iff it satisfies (K-1) to (K-5). \square

With the exception of (K-5), which involves two logically equivalent wffs, the basic AGM contraction postulates all refer to a fixed wff by which to contract a belief set. Basic AGM contraction can thus be seen as a description of how to contract a fixed belief set K by a fixed wff α . The addition of the two *supplementary AGM contraction postulates* below, enforces a connection between the belief sets obtained during a contraction of a (fixed) belief set by different wffs.²

$$(K-7) \quad (K - \alpha) \cap (K - \beta) \subseteq K - (\alpha \wedge \beta)$$

$$(K-8) \quad \text{If } \beta \notin K - (\alpha \wedge \beta) \text{ then } K - (\alpha \wedge \beta) \subseteq K - \beta$$

Definition 2.1.3 A removal is an *AGM contraction* iff it satisfies (K-1) to (K-8). \square

²This matter is discussed in more detail in section 2.2.

The postulates for revision follow the same pattern as for contraction. There are six *basic AGM revision postulates*.

$$(K*1) \quad K * \alpha = Cn(K * \alpha)$$

$$(K*2) \quad K * \alpha \subseteq K + \alpha$$

$$(K*3) \quad \text{If } \neg\alpha \notin K \text{ then } K * \alpha = K + \alpha$$

$$(K*4) \quad \alpha \in K * \alpha$$

$$(K*5) \quad \text{If } \alpha \equiv \beta \text{ then } K * \alpha = K * \beta$$

$$(K*6) \quad \perp \in K * \alpha \text{ iff } \models \neg\alpha$$

Definition 2.1.4 A revision is a *basic AGM revision* iff it satisfies (K*1) to (K*6). \square

(K*1) is the requirement that revision operate on belief sets, while (K*2) places an appropriate upper bound on the belief set obtained from a revision. (K*3) invokes the principle of Informational Economy for the case where the wff with which to revise is consistent with the current belief set. (K*4) ensures that revision is always successful, and (K*5) expresses the principle of the Irrelevance of Syntax. Finally, (K*6) highlights the difference between expansion and revision.

Like (basic AGM) contraction, basic AGM revision can be seen as a description of how to revise a fixed belief set by a fixed wff. To ensure that there is a connection between the revision by different wffs of the same belief set, it is necessary to add the *supplementary AGM revision postulates*.

$$(K*7) \quad K * (\alpha \wedge \beta) \subseteq (K * \alpha) + \beta$$

$$(K*8) \quad \text{If } \neg\beta \notin K * \alpha \text{ then } (K * \alpha) + \beta \subseteq K * (\alpha \wedge \beta)$$

Definition 2.1.5 A revision is an *AGM revision* iff it satisfies (K*1) to (K*8). \square

2.1.1 Connections between contraction and revision

A quick perusal of all the AGM postulates shows that, with the exception of (K–6) and (K*6), there are obvious similarities between the AGM contraction postulates and their revision counterparts. Gärdenfors [1988] shows that AGM contraction and AGM revision are interdefinable by courtesy of the two identities given below.

(Def * from \sim) $K * \alpha = (K \sim \neg\alpha) + \alpha$

(Def – from *) $K - \alpha = (K * \neg\alpha) \cap K$

These identities are known respectively as the *Levi identity* and the *Harper identity*. Observe that the Levi identity is a formal expression of Levi's commensurability thesis.

Theorem 2.1.6 1. A revision defined in terms of a (basic AGM) contraction using (Def * from \sim) is a basic AGM revision.³

2. A removal defined in terms of a basic AGM revision using (Def – from *) is a (basic AGM) contraction.

3. A revision defined in terms of an AGM contraction using (Def * from \sim) is an AGM revision.

4. A removal defined in terms of an AGM revision using (Def – from *) is an AGM contraction.

What is more, these two identities are also *interchangeable*. That is, if we start with a theory change operation (satisfying either the six basic AGM contraction postulates or the six basic AGM revision postulates), and then apply one of these identities, followed by an application of the other, we'll be back at the theory change operation that we started with.

Theorem 2.1.7 [Gärdenfors, 1988]

1. Let – be a (basic AGM) contraction, let * be obtained from – using (Def * from \sim) and let \sim be obtained from * using (Def – from *). Then – and \sim are identical.

2. Let * be a basic AGM revision, let – be obtained from * using (Def – from *) and let \ast be obtained from – using (Def * from \sim). Then * and \ast are identical.

So the Levi and Harper identities provide us with a strong form of interdefinability between contraction and revision. The significance of this result will become apparent when we discuss the methods for constructing AGM theory change.

³The proof of this part of the theorem does not make use of the Recovery postulate. This is a significant result that will be discussed and exploited in chapter 6.

2.2 Partial meet contraction

The first method proposed for constructing contraction operations [Alchourrón et al., 1985] is known as partial meet contraction.⁴ In this approach, the construction of an α -contraction uses as building blocks the maximal subsets of K that do not contain α .

Definition 2.2.1 A belief set K' is an α -remainder (of K) iff $K' \subseteq K$, $\alpha \notin K'$ and $\alpha \in K' + \beta$ for every $\beta \in K \setminus K'$. The set of α -remainders of K is denoted by $K \perp \alpha$. \square

It is easily seen that $K \perp \alpha = \{K\}$ iff $\alpha \notin K$, and that $K \perp \alpha = \emptyset$ if $\models \alpha$. Compactness further ensures that $K \perp \alpha = \emptyset$ only if $\models \alpha$. The partial meet contractions are obtained by picking out a set of α -remainders, and taking their intersection. Intuitively, we pick the best α -remainders, and then retain those wffs that occur in every one of them.

Definition 2.2.2 A *selection* function is a function $s_K : \{K \perp \alpha \mid \alpha \in L\} \rightarrow \wp \wp K$ such that $\emptyset \subset s_K(\mathcal{A}) \subseteq \mathcal{A}$ for every $\mathcal{A} \neq \emptyset$, and $s_K(\emptyset) = \{K\}$. \square

Selection functions are used to define the partial meet contractions.

(Def \sim from s_K) $K \sim \alpha = \bigcap s_K(K \perp \alpha)$

Definition 2.2.3 A removal is a *partial meet* contraction iff it is defined in terms of a selection function s_K using (Def \sim from s_K). \square

Theorem 2.2.4 [Alchourrón et al., 1985] *Every removal defined in terms of a selection function using (Def \sim from s_K) is a (basic AGM) contraction. Conversely, every (basic AGM) contraction can be defined in terms of a selection function using (Def \sim from s_K).*

The two limiting cases of partial meet contraction, in which $s_K(\alpha)$ is either taken as the set of all α -remainders, or as a single α -remainder, are known as *full meet contraction* and *maxichoice contraction* respectively. Clearly there is only one full meet contraction, but many maxichoice contractions. In fact, it is easily verified that every basic AGM contraction – can be defined in terms of a set M of maxichoice contractions, as follows:

(Def – from M) $K - \alpha = \bigcap_{\sim \in M} K \sim \alpha$

⁴Partial meet contraction is directly concerned with contraction, but the corresponding revisions can, of course, be obtained in terms of (Def $*$ from \sim).

and that full meet contraction is obtained when M contains *all* the maxichoice contractions. Full meet contraction thus provides a lower bound on basic AGM contraction in the sense that, for any basic AGM contraction, the belief set obtained from the α -contraction of a wff α includes the one obtained from the full meet α -contraction. It is also easily verified that full meet contraction is an AGM contraction (satisfying the supplementary postulates as well), but that not all of the maxichoice contractions are.

For the construction of AGM contraction in terms of $(\text{Def } \sim \text{ from } s_K)$, a selection function has to be more principled in its choice of α -remainders. This is attained by imposing a suitable binary relation \subseteq on the set of *remainders*

$$K \perp L = \bigcup \{A \in K \perp \alpha \mid \alpha \in L \setminus Cn(\top)\}$$

and defining a selection function from it as follows:

$$(\text{Def } s_K \text{ from } \subseteq) \quad s_K(K \perp \alpha) = \begin{cases} \{A \in K \perp \alpha \mid B \subseteq A, \forall B \in K \perp \alpha\} & \text{if } \not\models \alpha, \\ \{K\} & \text{otherwise} \end{cases}$$

Intuitively, \subseteq is used to obtain the maximal or “best” α -remainders (higher up being better), and these are the ones picked out by the selection function.

Definition 2.2.5 A partial meet contraction is called *relational* iff it is defined in terms of a selection function s_K (using $(\text{Def } \sim \text{ from } s_K)$), where s_K is defined in terms of a relation \subseteq using $(\text{Def } s_K \text{ from } \subseteq)$. If \subseteq is transitive, the partial meet contraction is called *transitively relational*, and if \subseteq is connected as well as transitive (which means that it is a total preorder), it is called *connectively relational*. \square

It turns out that all relational partial meet contractions satisfy (K-7), and that the transitively relational partial meet contractions, the connectively relational partial meet contractions, and the AGM contractions coincide exactly.

Theorem 2.2.6 [Gärdenfors, 1988] *A removal is an AGM contraction iff it is a transitively relational partial meet contraction, iff it is a connectively relational partial meet contraction.*

It is worth noting that not every relation \subseteq on $K \perp L$ can succeed in producing a selection function using $(\text{Def } s_K \text{ from } \subseteq)$. By definition, a selection function has to produce non-empty sets of α -remainders for every $K \perp \alpha$. So $(\text{Def } s_K \text{ from } \subseteq)$ will yield a selection function only if, for every α that is not logically valid, there are elements of

$K \perp \alpha$ that are at least as good, in terms of \subseteq , as *all* the elements of $K \perp \alpha$. And it is easy to verify that not every relation on $K \perp L$, nor even every transitive relation on $K \perp L$, has this property. Any irreflexive relation serves as an obvious counterexample. In fact, not even all the total preorders have this property. In this case, a counterexample is provided by considering a total preorder \subseteq that contains an infinitely ascending chain of elements of $K \perp \alpha$. This restriction of the application of (Def s_K from \subseteq) to well-behaved relations also explains the (seemingly) surprising result that the set of transitively relational partial meet contractions and the set of connectively relational partial meet contractions are identical. For it is a consequence of this result that the selection functions defined in terms of the total preorders using (Def s_K from \subseteq) coincide with the selection functions defined in terms of the transitive relations using (Def s_K from \subseteq). And this is the case because both the ill-behaved transitive relations and the ill-behaved total preorders are simply not taken into consideration in the definition of the selection functions. The obvious question to consider is whether it is possible to give a direct description of a set of transitive relations on $K \perp L$ that are well-behaved, in the sense that they induce selection functions when using (Def s_K from \subseteq), and can be used to construct all the AGM contractions. Such a description would provide a sharper version of theorem 2.2.6. In section 3.2 we shall see that this can be done.

2.3 Epistemic entrenchment

The basic idea behind epistemic entrenchment is that some of our beliefs are more firmly entrenched than others, and we would thus be more willing to give up the latter wffs than the former if we are forced to choose. In the view of Gärdenfors and Makinson [1988] and Gärdenfors [1988], an epistemic entrenchment ordering should be subject to the following set of postulates (with wffs higher up in the ordering being more entrenched):

(EE1) \subseteq_{EE} is transitive.

(EE2) If $\alpha \models \beta$ then $\alpha \subseteq_{EE} \beta$

(EE3) For all $\alpha, \beta \in K$, $\alpha \subseteq_{EE} \alpha \wedge \beta$ or $\beta \subseteq_{EE} \alpha \wedge \beta$

(EE4) If $K \neq Cn(\perp)$ then $\alpha \notin K$ iff $\alpha \subseteq_{EE} \beta$ for all β

(EE5) If $\alpha \sqsubseteq_{EE} \beta$ for all α then $\models \beta$

Definition 2.3.1 A binary relation \sqsubseteq_{EE} on L is an *EE-ordering* (an epistemic entrenchment ordering) with respect to a belief set K iff it satisfies (EE1) to (EE5). \square

(EE1) seems to be a reasonable condition to impose on a relation that qualifies as an ordering. (EE2) requires that logically weaker wffs be more entrenched, which makes perfect sense once we realise that it is impossible to remove a wff from a belief set without removing all the logically stronger wffs as well. The innocent-looking postulate (EE3) turns out to be very powerful indeed. It is the cornerstone of the controversial property that every EE-ordering is a total preorder. In chapter 5 we consider entrenchment orderings that are not total preorders. Finally, (EE4) and (EE5) are minimality and maximality conditions respectively. (EE4) states that all the wffs not in K are equally entrenched, but less entrenched than the wffs in K . And (EE5) (together with (EE2)) requires the logically valid wffs to be equally entrenched, but more entrenched than all the other wffs.

From results in [Gärdenfors and Makinson, 1988], AGM contraction and epistemic entrenchment are interdefinable by means of the following two identities:

$$(\text{Def} - \text{from } \sqsubseteq_{EE}) \quad K - \alpha = \begin{cases} K \cap \{\beta \mid \alpha \sqsubseteq_{EE} \alpha \vee \beta\} & \text{if } \alpha \in K, \text{ and } \not\models \alpha, \\ K & \text{otherwise} \end{cases}$$

$$(\text{Def } \sqsubseteq_{EE} \text{ from } \sim) \quad \alpha \sqsubseteq_{EE} \beta \text{ iff } \alpha \notin K \sim (\alpha \wedge \beta) \text{ or } \models \alpha \wedge \beta$$

Theorem 2.3.2 1. A removal is an AGM contraction iff it is defined in terms of an EE-ordering using (Def – from \sqsubseteq_{EE}).

2. A binary relation on L is an EE-ordering iff it is defined in terms of an AGM contraction using (Def \sqsubseteq_{EE} from \sim).

In fact, as we shall see in chapter 3, these identities are interchangeable in the sense that moving from an EE-ordering to an AGM contraction and back (or vice versa), brings us back to where we started.

2.3.1 Plausibility orderings

Grove [1988] presents a class of plausibility orderings on wffs. The set of postulates he uses to describe these orderings bears some resemblance to that for the EE-orderings.

(GE1) \sqsubseteq_{GE} is connected

(GE2) \sqsubseteq_{GE} is transitive

(GE3) If $\alpha \models \beta \vee \gamma$ then $\beta \sqsubseteq_{GE} \alpha$ or $\gamma \sqsubseteq_{GE} \alpha$

(GE4) If $K \neq Cn(\perp)$ then $\neg\alpha \notin K$ iff $\alpha \sqsubseteq_{GE} \beta$ for all $\beta \in L^5$

(GE5) $\models \neg\alpha$ iff $\beta \sqsubseteq_{GE} \alpha$ for all $\beta \in L$

Definition 2.3.3 A binary relation \sqsubseteq_{GE} on L is a *GE-ordering* (with respect to a belief set K) iff it satisfies (GE1) to (GE5). \square

Grove then defines AGM revision in terms of the GE-orderings as follows:⁶

(Def * from \sqsubseteq_{GE}) $\beta \in K * \alpha$ iff $\begin{cases} (\alpha \wedge \beta) \sqsubseteq_{GE} (\alpha \wedge \neg\beta) \text{ if } \not\models \neg\alpha, \\ \beta \in L \text{ otherwise} \end{cases}$

Theorem 2.3.4 [Grove, 1988] *Every GE-ordering defines an AGM revision using (Def * from \sqsubseteq_{GE}). Conversely, every AGM revision can be defined in terms of a GE-ordering using (Def * from \sqsubseteq_{GE}).*

⁵Grove [1988] does not include the condition that $K \neq Cn(\perp)$ in (GE4), but without it some of his results (Theorem 4, p. 164) do not hold for an unsatisfiable K . Gärdenfors [1988] gives the same formulation as Grove, but his result about the relation between epistemic entrenchment orderings and the Grove orderings (Lemma 4.27, p. 96) only holds if the above condition is included. The proposal of Boutilier [1992, 1994] to rectify the formulation of (GE4) is to exclude the condition that $K \neq Cn(\perp)$, as well as the reverse direction of our version of (GE4). But this is too weak, and it can be shown that it destroys the desired relationship between the Grove orderings and the epistemic entrenchment orderings of Gärdenfors and Makinson.

⁶Grove's definition of revision in [1988] in terms of G-orderings does not include the case where $\models \neg\alpha$, and neither does the definition of Gärdenfors [1988], but it is clearly a necessary part of the definition. For if α is logically invalid, then $K * \alpha = Cn(\perp)$, by (K*6). But, since both $\alpha \wedge \beta$ and $\alpha \wedge \neg\beta$ are then also logically invalid for every β , it follows from (GE5) that $(\alpha \wedge \beta) \not\sqsubseteq_{GE} (\alpha \wedge \neg\beta)$.

Grove describes the GE-orderings as measures of relative importance, and they have also been described as orderings of plausibility [Boutilier, 1992]. However, (GE4) seems to be at odds with both these epistemic interpretations of the GE-orderings, for in both these cases, one would expect the wffs in K to be more important, or more plausible than, all the wffs not in K . And yet (GE4) requires of every wff whose negation is not in K to be as important, or as plausible, as the wffs in K . We return to this issue in chapter 5.

Gärdenfors [1988] shows that the resemblance between the postulates for the EE-orderings and the GE-orderings is not just coincidental, and that the GE-orderings are dual to the EE-orderings in the following sense:

(Def \sqsubseteq_E from \sqsubseteq_G) $\alpha \sqsubseteq_E \beta$ iff $\neg\alpha \sqsubseteq_G \neg\beta$

Theorem 2.3.5 [Gärdenfors, 1988] *A relation on wff is an EE-ordering iff it can be defined in terms of a GE-ordering using (Def \sqsubseteq_E from \sqsubseteq_G).*

From (GE2) it follows that logically equivalent wffs are equally plausible, and the GE-orderings can thus be defined in terms of the EE-orderings in a manner analogous to that in (Def \sqsubseteq_E from \sqsubseteq_G):

(Def \sqsubseteq_G from \sqsubseteq_E) $\alpha \sqsubseteq_G \beta$ iff $\neg\alpha \sqsubseteq_E \neg\beta$

2.4 Safe contraction

Safe contraction was originally introduced by Alchourrón and Makinson [1981, 1985]. Intuitively, the idea is to identify wffs in the belief set K that cannot be blamed for a wff α being in K . When contracting K by α , these wffs should all be retained, i.e., they are *safe* with respect to a contraction by α . The belief set resulting from an α -contraction is then taken to be the belief set generated by the wffs that are safe with respect to α . To determine the wffs that are safe with respect to α , we first need to consider the minimal subsets of K that entail α , dubbed the entailment sets for α .

Definition 2.4.1 B is an *entailment set* for α (with respect to K) iff $B \subseteq K$ and $B \models \alpha$, but $B \setminus \{\beta\} \not\models \alpha$ for every $\beta \in B$. We denote the set of entailment sets of α by $K \top \alpha$. \square

Note that $K \top \alpha = \emptyset$ iff $\models \alpha$ or $\alpha \notin K$. We also need to introduce a binary relation on the wffs in K , subject to certain conditions.

Definition 2.4.2 A binary relation \sqsubset_H on K is a *hierarchy* (over K) iff for every finite sequence $\alpha_1, \dots, \alpha_n$ of wffs in K , if $\alpha_i \sqsubset_H \alpha_{i+1}$ for $1 \leq i < n$, then $\alpha_n \not\sqsubset_H \alpha_1$. \square

A hierarchy over K can be seen as an indication of the reliability of the wffs in K , with those higher up being more reliable. As such, it is not unlike an epistemic entrenchment ordering with respect to K . Wffs in K that are safe with respect to α are taken to be those that are not minimal elements (with respect to \sqsubset_H) of any of the entailment sets for α . In other words, the wffs in K that are *not* safe with respect to α , are those that occur as the least reliable members of some entailment set for α . We denote the wffs that are safe with respect to α (and a hierarchy \sqsubset_H) by K/α . That is:

$$(\text{Def } K/\alpha \text{ from } \sqsubset_H) \quad K/\alpha = \left\{ \beta \in K \left| \begin{array}{l} \forall B \in K \top \alpha \text{ such that } \beta \in B, \\ \exists \gamma \in B \text{ such that } \gamma \sqsubset_H \beta \end{array} \right. \right\}$$

K/α is then used to define safe contraction.

$$(\text{Def } - \text{ from } \sqsubset_H) \quad K - \alpha = \left\{ \begin{array}{l} Cn(K/\alpha) \text{ where } K/\alpha \text{ is defined in terms of } \sqsubset_H \\ \text{using (Def } K/\alpha \text{ from } \sqsubset_H) \end{array} \right.$$

Definition 2.4.3 A removal $-$ is a *safe contraction* iff it is defined in terms of a hierarchy \sqsubset_H using (Def $-$ from \sqsubset_H). \square

Alchourrón and Makinson [1985] show that every safe contraction is a (basic AGM) contraction. In [1986], they also provide a connection with AGM contraction for the case where K is finitely axiomatisable. To do so, they impose stricter conditions on hierarchies.

Definition 2.4.4 A hierarchy over K

1. *continues up* iff the following holds for every $\alpha, \beta, \gamma \in K$: if $\alpha \sqsubset_H \beta$ and $\beta \models \gamma$ then $\alpha \sqsubset_H \gamma$,
2. *continues down* iff the following holds for every $\alpha, \beta, \gamma \in K$: if $\alpha \models \beta$ and $\beta \sqsubset_H \gamma$ then $\alpha \sqsubset_H \gamma$,
3. is *regular* iff it continues up and down, and

4. is *virtually connected* iff the following holds for every $\alpha, \beta, \gamma \in K$: if $\alpha \sqsubset_H \beta$ then $\alpha \sqsubset_H \gamma$ or $\gamma \sqsubset_H \beta$.

□

They show that if K is finitely axiomatisable, the removals defined in terms of the regular virtually connected hierarchies using (Def – from \sqsubset_H) are precisely the AGM contractions. Rott [1992b] extends this to the general case as well.

Theorem 2.4.5 *A removal – is a safe contraction defined in terms of a regular virtually connected hierarchy using (Def – from \sqsubset_H) iff it is an AGM contraction.*

In chapter 3 we delve deeper into the connection between safe contraction and other methods for constructing AGM contraction.

Chapter 3

Semantic belief change

*‘When I use a word,’ Humpty Dumpty said
in rather a scornful tone, ‘it means just what
I choose it to mean — neither more nor less.’*

Lewis Carrol, *Alice’s Adventures in Wonderland*

One of the central themes of this dissertation is to emphasise the fundamental role that semantic approaches play in belief change. In this chapter we take the initial steps in the justification of such a claim. We commence with the introduction of a notion of semantic information and its relation to a possible-worlds semantics for L . This is followed by a discussion of semantic approaches to AGM theory change, in which the central idea is that of a preorder on the interpretations of L . We point out the strong links between such semantic approaches and the methods for constructing AGM theory change that were discussed in chapter 2. With the aid of our theory of semantic information, we argue that the preorders on interpretations can be transformed into preorders on the basic units of belief of an agent, and that it is appropriate to use these orderings as representations of the epistemic states of an agent. In this and in later chapters, such a representation of epistemic states will prove to be most fruitful.

3.1 Semantic content and infatoms

One of the assumptions encountered in chapter 2 is the representation of an epistemic state as a belief set. This decision has a number of associated problems, one of the most basic objections being that the elements of a belief set are linguistic in nature. In our view, an epistemic state ought to consist of non-linguistic entities from which the beliefs associated with the epistemic state can be determined. And since it is the semantics of the language that determines the meaning of the wffs in the language, the interpretations of the language (the elements of U) are usually used as the basic building blocks of epistemic states. This is the basis for the representations used by many authors [Harper, 1977, Grove, 1988, Katsuno and Mendelzon, 1991, Morreau, 1992, Peppas and Williams, 1995, Darwiche and Pearl, 1997]. Such representations have proved to be very useful in a wide variety of situations, and much of the work discussed in this and later chapters are based on the idea of an epistemic state as a set of interpretations. But if we think of the elements of an epistemic state as objects from which (linguistic) beliefs are built up, the use of interpretations does not seem to be quite satisfactory. For it is difficult to see how an interpretation can be considered as a basic part of a belief expressed as a wff in L .¹

It is with this objection in mind that we propose the use of *infatoms* as the basic units of an epistemic state. Intuitively, infatoms are the basic independent pieces of information from which the beliefs of an agent (expressed as wffs of L) are built up. In this view, the information contained in a belief, and in a belief set, is a set of infatoms. More infatoms thus correspond to a set of beliefs that contains more information and is logically stronger. Infatoms are *independent* in the sense that it is only the set of all infatoms that contains too much information, leading an agent to include all wffs in its set of beliefs. Any strict subset of the set of all infatoms corresponds to a satisfiable set of beliefs.

Since the notion of an epistemic state is so central to the study of belief change, it seems more appropriate to use a semantics based on infatoms when constructing belief change operations. Although we give a formal description of infatoms and an infatom semantics below, we shall express most of the formal work on semantic belief

¹In fact, it makes more sense to do it the other way round and think of an interpretation (or rather, a valuation) as being built up from a set of wffs. As we have seen in section 1.3, this is a standard way of constructing a semantics that is isomorphic to the \models -valuation semantics for L .

change in this, and indeed in later chapters, in terms of a possible-worlds semantics. There are two reasons for this. Firstly, semantic descriptions of belief change have thus far concentrated on the use of a possible-worlds semantics. (In fact, with a few exceptions, such as [Lindström, 1991], the emphasis has been placed on a semantics that is isomorphic (or identical) to a valuation semantics for L .) And secondly, we'll show that there is such a close technical association between infatoms and interpretations (and valuations in particular), that a switch from interpretations to infatoms is merely a matter of regarding an interpretation as its associated infatom. In order to formalise this relationship, we now proceed with a formal explication of a semantics for L based on infatoms.

Infatoms are generalised semantic versions of the *content elements* of Carnap and Bar-Hillel [1952, 1953], and as such, are quite different from Keith Devlin's [1991] *infons*, although the latter is also described as basic bits of information. Formally, an infatom is a function i from A_{PL} , the set of propositional atoms of L (see section 1.3), to the set $\{I, E\}$. The intuition is that infatoms are independent bits of information from which the information contained in the wffs of L are built up. An infatom i sends a propositional atom α to I if i is *Included* in the information contained in α , and i sends α to E if i is *Excluded* from the information contained in α .

Definition 3.1.1 Given a set Inf of infatoms, the *content relation* \Vdash from Inf to L is then defined recursively as follows:

1. for every $i \in Inf$, $i \not\Vdash \top$ and $i \Vdash \perp$,
2. if $\alpha \in A_{PL}$, then $i \Vdash \alpha$ iff $i(\alpha) = I$,
3. if $\alpha = \neg\beta$ then $i \Vdash \alpha$ iff $i \not\Vdash \beta$,
4. if $\alpha = \beta \vee \gamma$ then $i \Vdash \alpha$ iff $i \Vdash \beta$ and $i \Vdash \gamma$,
5. if $\alpha = \beta \wedge \gamma$ then $i \Vdash \alpha$ iff either $i \Vdash \beta$ or $i \Vdash \gamma$, or both,
6. if $\alpha = \beta \rightarrow \gamma$ then $i \Vdash \alpha$ iff $i \not\Vdash \beta$ and $i \Vdash \gamma$, and
7. if $\alpha = \beta \leftrightarrow \gamma$ then $i \Vdash \alpha$ iff either both $i \not\Vdash \beta$ and $i \Vdash \gamma$, or both $i \Vdash \beta$ and $i \not\Vdash \gamma$.

□

The *semantic content* of a set of wffs A , denoted by $C(A)$, is defined as

$$C(A) = \{i \in \text{Inf} \mid \exists \alpha \in A \text{ such that } i \Vdash \alpha\}.$$

For a wff $\alpha \in L$, we write $C(\alpha)$ instead of $C(\{\alpha\})$. So the semantic content of A consists of all the infatoms that are part of the information contained in at least one of the wffs in A . We shall refer to such infatoms as the *content bits* of A . Conversely, for an infatom i , A is said to be *i -containing* iff i is a content bit of A . An *infatom semantics* for L is an ordered pair (Inf, \Vdash) , where Inf is a set of infatoms and \Vdash is the content relation of definition 3.1.1. The *theory generated* by a set of infatoms $I \subseteq \text{Inf}$ is defined as $Th(I) = \{\alpha \mid C(\alpha) \subseteq I\}$. That is, $Th(I)$ contains all the wffs whose contents bits are included in I . Our first result about infatoms is given without proof.

Proposition 3.1.2 *Let (Inf, \Vdash) be an infatom semantics for L .*

1. $C(\alpha \wedge \beta) = C(\alpha) \cup C(\beta)$.
2. $C(\alpha \vee \beta) = C(\alpha) \cap C(\beta)$.

It turns out that there is a natural way to associate a unique infatom semantics with every possible-worlds semantics, and to associate a unique valuation semantics with every infatom semantics.

Definition 3.1.3 1. Given a possible-worlds semantics (U, \Vdash) for L , the *associated* infatom semantics for L is defined as (Inf, \Vdash) , where $\text{Inf} = \{i_u \mid u \in U\}$, \Vdash is obtained as in definition 3.1.1, and for every $u \in U$, the *associated* infatom i_u is defined as follows: for every $\alpha \in A_{PL}$, $i_u(\alpha) = I$ iff $u \Vdash \alpha$.

2. Given an infatom semantics (Inf, \Vdash) for L , the *associated* valuation semantics (V, \Vdash) based on valuations is defined as (V, \Vdash) , where $V = \{v_i \mid i \in \text{Inf}\}$, \Vdash is obtained in the standard way (see section 1.3), and for every $i \in \text{Inf}$, the *associated* valuation v_i is defined as follows: for every $\alpha \in A_{PL}$, $v_i(\alpha) = T$ iff $i \Vdash \alpha$.

□

Definition 3.1.3 is justified by propositions 3.1.5 and 3.1.6 below. They, in turn, rely heavily on the following lemma.

Lemma 3.1.4 1. Let (U, \models) be a possible-worlds semantics and let (Inf, \Vdash) be the associated infatom semantics for L . For every $\alpha \in L$ and every $u \in U$, $i_u \in C(\alpha)$ iff $u \notin M(\alpha)$.

2. Let (Inf, \Vdash) be an infatom semantics and let (V, \models) be the valuation semantics associated with (Inf, \Vdash) . For every $\alpha \in L$ and every $i \in \text{Inf}$, $v_i \in M(\alpha)$ iff $i \notin C(\alpha)$.

Proof Both proofs follow by induction on the structure of the wffs of L , and applications of definition 3.1.3. \square

Proposition 3.1.5 establishes some connections between interpretations and infatoms.

Proposition 3.1.5 Let (U, \models) be a possible-worlds semantics and let (Inf, \Vdash) be the associated infatom semantics for L .

1. $A \models \beta$ iff $M(A) \subseteq M(\beta)$ iff $C(A) \supseteq C(\beta)$.
2. $\text{Th}(C(A)) = \text{Th}(M(A))$.
3. $\models \alpha$ iff $M(\alpha) = U$ iff $C(\alpha) = \emptyset$.
4. $C(A) = \text{Inf} \setminus \{i_u \mid u \in M(A)\}$.
5. If $W \subseteq U$ and $I = \{i_w \mid w \in W\}$ then $\text{Th}(W) = \text{Th}(\text{Inf} \setminus I)$.
6. $\text{Th}(M(A) \cup \{u\}) = \text{Th}(C(A) \setminus \{i_u\})$.
7. If $u \in U$ then

$$\text{Th}(M(A) \setminus \{w \mid \text{is elementarily equivalent to } u\}) = \text{Th}(C(A) \cup \{i_u\}).$$

Proof 1. Suppose that $M(A) \subseteq M(\beta)$ and pick any $i_u \in C(\beta)$. Now assume that $i_u \notin C(A)$. That is, for every $\alpha \in A$, $i_u \notin C(\alpha)$. Then, by lemma 3.1.4, $u \in M(\alpha)$ for every $\alpha \in A$, and therefore $u \in M(A)$. But by supposition, $u \in M(\beta)$, and by lemma 3.1.4, $i_u \notin C(\beta)$; a contradiction. Conversely, suppose that $C(\beta) \subseteq C(A)$ and pick any $u \in M(A)$. Now assume that $u \notin M(\beta)$. By lemma 3.1.4, $i_u \in C(\beta)$. So $i_u \in C(A)$, and there is thus an $\alpha \in A$ such that $i_u \in C(\alpha)$. But by lemma 3.1.4, $u \notin M(\alpha)$, contradicting the supposition that $u \in M(A)$.

2. $\alpha \in Th(C(A))$ iff $C(\alpha) \subseteq C(A)$ iff $M(A) \subseteq M(\alpha)$ (by part (1) above) iff $\alpha \in Th(M(A))$.
3. It follows easily from the definitions of $M(\alpha)$ and $C(\alpha)$ that $M(\alpha) = U$ iff $\models \alpha$ and that $C(\alpha) = \emptyset$ iff $\models \alpha$.
4. Pick an $i_u \in C(A)$. So, there is an $\alpha \in A$ such that $i_u \in C(\alpha)$. By lemma 3.1.4, $u \notin M(\alpha)$. So $u \notin M(A)$, and thus $i_u \in \text{Inf} \setminus \{i_v \mid v \in M(A)\}$. Conversely, suppose that $i_u \in \text{Inf} \setminus \{i_v \mid v \in M(A)\}$. Then $u \notin M(A)$, and there is thus an $\alpha \in A$ such that $u \notin M(\alpha)$. Hence, by lemma 3.1.4, $i_u \in C(\alpha)$. Therefore $i \in C(A)$.
5. Suppose that $W \subseteq U$ and $I = \{i_w \mid w \in W\}$, and pick an $\alpha \in Th(W)$. So $W \subseteq M(\alpha)$. Now assume that $\alpha \notin Th(\text{Inf} \setminus I)$. That is, $C(\alpha) \not\subseteq \text{Inf} \setminus I$. In other words, there is an $i_u \in C(\alpha)$ such that $i_u \in I$. So $u \in W$ and by lemma 3.1.4, $u \notin M(\alpha)$, contradicting the fact that $W \subseteq M(\alpha)$. Conversely, suppose that $\alpha \in Th(\text{Inf} \setminus I)$. So $C(\alpha) \subseteq \text{Inf} \setminus I$. Now assume that $\alpha \notin Th(W)$. Then $W \not\subseteq M(\alpha)$, and there is thus a $w \in W$ such that $w \notin M(\alpha)$. So $i_w \in I$, and by lemma 3.1.4, $i_w \in C(\alpha)$, thus contradicting the fact that $C(\alpha) \subseteq \text{Inf} \setminus I$.
6. Pick any $\beta \in Th(M(A) \cup \{u\})$. That is, $M(A) \cup \{u\} \subseteq M(\beta)$. It suffices to show that $C(\beta) \subseteq C(A) \setminus \{i_u\}$. So pick any $i_v \in C(\beta)$. By lemma 3.1.4, $v \notin M(\beta)$, and this means that $v \notin M(A) \cup \{u\}$. So $u \neq v$ (and hence $i_u \neq i_v$) and there is an $\alpha \in A$ such that $v \notin M(\alpha)$. By lemma 3.1.4 it then follows that $i_v \in C(\alpha)$, and thus that $i_v \in C(A)$. The required result then follows from the fact that $i_u \neq i_v$. Conversely, pick any $\beta \in Th(C(A) \setminus \{i_u\})$. That is $C(\beta) \subseteq C(A) \setminus \{i_u\}$. It suffices to show that $M(A) \cup \{u\} \subseteq M(\beta)$. So pick any $v \in M(A) \cup \{u\}$. Then either $v = u$ (and hence $i_u = i_v$), or $v \in M(\alpha)$ for every $\alpha \in A$. In the former case $i_v \notin C(\beta)$, and in the latter case, it follows by lemma 3.1.4 that $i_v \notin C(\alpha)$ for every $\alpha \in A$, and thus that $i_v \notin C(A)$. So either way, $i_v \notin C(\beta)$, and hence, by lemma 3.1.4, $v \in M(\beta)$.
7. Pick any $\beta \in Th(M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\})$. That is, $M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\} \subseteq M(\beta)$. It suffices to show that $C(\beta) \subseteq C(A) \cup \{i_u\}$. So pick any $i_v \in C(\beta)$. By lemma 3.1.4, $v \notin M(\beta)$. And this means that $v \notin M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\}$. So either

v is elementarily equivalent to u (and hence $i_u = i_v$), or there is an $\alpha \in A$ such that $v \notin M(\alpha)$. In the former case it obviously follows that $i_v \in C(A) \cup \{i_u\}$, and in the latter case it follows by lemma 3.1.4 that $i_v \in C(\alpha)$. But then $i_v \in C(A)$, which means we are done. Conversely, pick any $\beta \in Th(C(A) \cup \{i_u\})$. That is, $C(\beta) \subseteq C(A) \cup \{i_u\}$. It suffices to show that

$$M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\} \subseteq M(\beta).$$

So pick any $v \in M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\}$. Then $v \in M(\alpha)$ for every $\alpha \in A$. By lemma 3.1.4, $i_v \notin C(\alpha)$ for every $\alpha \in A$. Therefore $i_v \notin C(A)$. Furthermore $i_v \neq i_u$, for if $i_v = i_u$, it would have meant that v is elementarily equivalent to u , contradicting the fact that $v \in M(A) \setminus \{w \mid w \text{ is elementarily equivalent to } u\}$. And thus $i_v \notin C(\beta)$, which means, by lemma 3.1.4, that $v \in M(\beta)$. □

Part (1) of proposition 3.1.5 shows us that semantic entailment can also be defined in terms of infatoms. A wff α is semantically entailed by a set of wffs A iff the content bits of A includes all the content bits of α . This enables us to define a notion of axiomatisability for sets of infatoms. A set of wffs A *axiomatises* a set of infatoms I iff $C(A) = I$. The intuition is along the same lines as the axiomatisability of sets of interpretations; both provide syntactic descriptions of a semantic concept.

Of particular interest in the proposition above are the last two parts. Part (6) shows that adding an interpretation to the models of a set of wffs A is the same as removing its associated infatom from the semantic content of A . Part (7) shows that the removal, from the models of A , of all interpretations that are elementarily equivalent to an interpretation u is the same as adding u 's associated infatom to the semantic content of A .

The next proposition draws parallels between valuations and infatoms.

Proposition 3.1.6 *Let (Inf, \Vdash) be an infatom semantics and let (V, \Vdash) be the valuation semantics associated with (Inf, \Vdash) .*

1. $M(A) = V \setminus \{v_i \mid i \in C(A)\}$.
2. If $I \subseteq Inf$ and $W = \{w_i \mid i \in I\}$ then $Th(I) = Th(V \setminus W)$.

$$3. Th(C(A) \cup \{i\}) = Th(M(A) \setminus \{v_i\}).$$

$$4. Th(C(A) \setminus \{i\}) = Th(M(A) \cup \{v_i\}).$$

Proof 1. $v_i \in M(A)$ iff $v_i \in M(\alpha)$ for every $\alpha \in A$, iff $i \notin C(\alpha)$ for every $\alpha \in A$ (by lemma 3.1.4), iff $i \notin C(A)$, iff $v_i \in V \setminus \{w_i \mid i \in C(A)\}$.

2. Suppose that $I \subseteq \text{Inf}$ and $W = \{w_i \mid i \in I\}$. Now pick any $\alpha \in Th(I)$. That is, $C(\alpha) \subseteq I$. It suffices to show that $V \setminus W \subseteq M(\alpha)$. So pick any $w_i \in V \setminus W$. By the definition of W it follows that $i \notin I$. But this means that $i \notin C(\alpha)$, and by lemma 3.1.4, that $w_i \in M(\alpha)$. Conversely, pick any $\alpha \in Th(V \setminus W)$. That is, $V \setminus W \subseteq M(\alpha)$. It suffices to show that $C(\alpha) \subseteq I$. So pick an $i \notin I$. By the definition of W it follows that $w_i \in V \setminus W$. But then $w_i \in M(\alpha)$, and by lemma 3.1.4, $i \notin C(\alpha)$.

3. The proof is very similar to the proof of part (7) of proposition 3.1.5 and is omitted.

4. The proof is very similar to the proof of part (6) of proposition 3.1.5 and is omitted.

□

These results clearly show that there is a strong connection between interpretations and valuations on the one hand, and infatoms on the other.

The connection between infatoms and the contents elements of Carnap and Bar-Hillel [1952, 1953] is easily established as follows. Let L be the language of a finitely generated propositional logic, generated by the atoms p_1, \dots, p_n and let (V, \models_V) be the classical valuation semantics for L (i.e., V contains all possible valuations). Now, let (Inf, \models) be the associated infatom semantics for L . It is well known that a valuation $v \in V$ can be axiomatised by a conjunction of literals $\bigwedge_{i=1}^n l_i$, where $l_i \in \{p_i, \neg p_i\}$. That is, $M(\bigwedge_{i=1}^n l_i) = \{v\}$. These conjunctions are called *state descriptions*. From part (4) of proposition 3.1.5 it then follows that $C(\neg(\bigwedge_{i=1}^n l_i)) = \{i_v\}$. That is, the negation of the state description of v axiomatises the infatom associated with v . And it is precisely these negations of the state descriptions that are the content elements of Carnap and Bar-Hillel.

From the discussion above it is clear that every entailment relation \models for L can be obtained from a unique valuation semantics (V, \models) and a unique infatom semantics (Inf, \models) , and that (V, \models) is the valuation semantics associated with (Inf, \models) and (Inf, \models) is the infatom semantics associated with (V, \models) . It is therefore appropriate to consider (V, \models) and (Inf, \models) as dual to each other. It is also clear that, like valuations and unlike interpretations, there can never be two distinct infatoms i and j that are elementarily equivalent in the sense that they are content bits of exactly the same wffs (i.e. $\text{Th}(i) = \text{Th}(j)$). So valuations and infatoms have the same grainsize, with interpretations (possibly) being finer grained than either valuations or infatoms. From an information-theoretic point of view, it seems reasonable to appeal to the Principle of the Identity of Indiscernibles, thereby disallowing the elementary equivalence of distinct infatoms:

(Identity of Indiscernibles) Objects that cannot be distinguished from one another should be regarded as identical.

Given the close connection between interpretations and valuations on the one hand, and infatoms on the other, we shall frequently find it convenient to refer to interpretations as infatoms. In particular, when referring to an interpretation $u \in U$ as an infatom, we actually have in mind the infatom i_u associated with u . While this convention introduces some ambiguity, it should cause no confusion, and will aid in brevity.

3.2 A semantics for theory change

The construction of basic AGM contraction in terms of partial meet contraction can easily be converted into a semantic description of basic AGM theory change. Grove [1988] pointed out that it is just a matter of realising that the α -remainders are obtained by adding single models of $\neg\alpha$ to the models of K .

Proposition 3.2.1 *Let $\not\models \alpha$ and $\alpha \in K$.*

1. *If $u \in M(\neg\alpha)$ then $\text{Th}(M(K) \cup \{u\}) \in K \perp \alpha$.*
2. *If $A \in K \perp \alpha$ then $A = \text{Th}(M(K) \cup \{u\})$ for some $u \in M(\neg\alpha)$.*

Proof For (1) pick any $u \in M(\neg\alpha)$. Clearly, $\text{Th}(M(K) \cup \{u\}) \subseteq K$ and $\text{Th}(M(K) \cup \{x\}) \not\models \alpha$. Now pick any $\beta \in K$ such that $\text{Th}(M(K) \cup \{u\}) \not\models \beta$, and thus $u \notin M(\beta)$.

By lemma 1.3.3, $Th(M(K) \cup \{u\}) + \beta = Th(M(K) \cap M(\beta)) = K$, and so $\alpha \in K = Th(M(K) \cup \{u\}) + \beta$.

For (2) pick any $A \in K \perp \alpha$. Because $A \not\models \alpha$, there is a $u \in M(A)$ such that $u \in M(\neg\alpha)$, and there is thus a $W \subseteq U$ such that $W \cap M(K) = \emptyset$, $u \in W$ and $A = Th(M(K) \cup W) \subseteq Th(M(K) \cup \{u\})$. If $A \subset Th(M(K) \cup \{u\})$ then there is a $\beta \in Th(M(K) \cup \{u\})$, and therefore $\beta \in K$, such that $\beta \notin A$. But then $A + \beta = Th(M(K) \cup W) + \beta \subseteq Th(M(K) \cup \{u\}) + \beta = Th(M(K) \cup \{u\})$. And since $\alpha \notin Th(M(K) \cup \{u\})$, it also follows that $\alpha \notin A + \beta$, contradicting the fact that $A \in K \perp \alpha$. \square

If L has a valuation semantics, then there is a one-to-one correspondence between the elements of $M(\neg\alpha)$ and the elements of $K \perp \alpha$. In general however, different elements of $M(\neg\alpha)$ may determine the same element of $K \perp \alpha$. From an information-theoretic viewpoint, an α -remainder is obtained by removing one of the content bits of α from the semantic content of K .

Proposition 3.2.1 gives us a way to characterise the partial meet contractions semantically. Instead of a function selecting a subset of the remainders of K after removing α , we have a function selecting a subset of the models of $\neg\alpha$ to be added to the models of K . We call such a function a semantic selection function.

Definition 3.2.2 A function $sm_K : L \rightarrow \wp U$ is a *semantic selection function* iff the following holds:

1. If $\alpha \equiv \beta$ then $sm_K(\alpha) = sm_K(\beta)$ and
2. if $\alpha \notin K$ or $\models \alpha$ then $sm_K(\alpha) = \emptyset$, otherwise $\emptyset \subset sm_K(\alpha) \subseteq M(\neg\alpha)$.

\square

Basic AGM theory change can then be defined in terms of semantic selection functions as follows:

(Def \sim from sm_K) $K \sim \alpha = Th(M(K) \cup sm_K(\alpha))$

(Def $*$ from sm_K) $K * \alpha = \begin{cases} Th(sm_K(\neg\alpha)) & \text{if } \neg\alpha \in K \text{ and } \not\models \neg\alpha, \\ K + \alpha & \text{otherwise} \end{cases}$

Theorem 3.2.3 1. *A removal defined in terms of a semantic selection function using $(Def \sim \text{from } sm_K)$ is a basic AGM contraction. Conversely, every basic AGM contraction can be defined in terms of a semantic selection function using $(Def \sim \text{from } sm_K)$.*

2. *A revision defined in terms of a semantic selection function using $(Def * \text{from } sm_K)$ is a basic AGM revision. Conversely, every basic AGM revision can be defined in terms of a semantic selection function using $(Def * \text{from } sm_K)$.*

Proof The proof can be found in appendix A. □

Information-theoretically, theorem 3.2.3 tells us that if the semantic content of K contains all the content bits of α and α is not logically valid, then α -contraction is obtained by removing some of the content bits of α from the semantic content of K . Similarly, if the semantic content of K contains all the content bits of $\neg\alpha$, and $\neg\alpha$ is not logically valid, then an α -revision is obtained by adding all content bits of α to the semantic content of K , and removing some of content bits of $\neg\alpha$. So basic AGM contraction involves the removal of content bits of α , while basic AGM revision means adding all the content bits of α , and removing some content bits of $\neg\alpha$.

For a semantic construction of AGM theory change, it is necessary to approach matters in a more systematic fashion. It turns out that the use of preorders on the interpretations of L , subject to certain restrictions, will do the trick. The first explicitly semantic method for constructing AGM revision (satisfying all eight of the AGM revision postulates) is due to Grove [1988], who uses a generalised version of Lewis' [1973] *sphere-semantics* for counterfactuals. Grove's systems of spheres are based on the maximally satisfiable subsets of L . By considering these sets as interpretations of L , we obtain a (possible-worlds) semantics for L that is isomorphic to the \models -valuation semantics for L . Let us denote by $[A]$ the set of maximally satisfiable extensions of a set of wffs $A \subseteq L$, and that of a single wff $\alpha \in L$ by $[\alpha]$. When viewed as interpretations, the elements of $[A]$ are thus the models of A . A *system of spheres* (centred on K) is a collection \mathcal{S} of subsets of $[\top]$, the set of all maximally satisfiable subsets of L , that satisfy the following conditions:

(S1) \mathcal{S} is totally ordered by set-inclusion

(S2) $[K]$ is the \subseteq -minimum of \mathcal{S}

(S3) $[\top] \in S$

(S4) If any element of S intersects $[\alpha]$, then there is a smallest element of S doing so

Letting $S_{\min}(\alpha)$ be the smallest element of S intersecting $[\alpha]$, AGM revision can then be defined in terms of S as follows:

$$(\text{Def } * \text{ from } S) \quad K * \alpha = \begin{cases} \bigcap([\alpha] \cap S_{\min}) & \text{if } [\alpha] \neq \emptyset, \\ L & \text{otherwise} \end{cases}$$

Theorem 3.2.4 [Grove, 1988] *Every system of spheres defines an AGM revision using (Def * from S). Conversely, every AGM revision can be defined in terms of a system of spheres using (Def * from S).*

With $[\top]$ viewed as the set of interpretations of L , it is not difficult to see that a system of spheres corresponds to a preorder on U , subject to a number of conditions. (S1) ensures that the preorder is total, (S2) requires that the models of K all be equally comparable and strictly below the countermodels of K , and (S3) ensures that the preorder is defined on the whole of U . The purpose of (S4) is to retain only those total preorders for which the set of minimal models of every wff (that is not logically invalid) is non-empty. From a suggestion by Katsuno and Mendelzon [1991], we refer to such preorders as *faithful*. For reasons that will become clear, our definition applies to all the preorders on U and not just the total preorders.

Definition 3.2.5 Let \preceq be any preorder on U .

1. If $W \subseteq U$ then any $v \in W$ is \preceq -minimal in W iff for every $w \in W$, $w \not\prec v$. We denote the set of \preceq -minimal elements of $M(\alpha)$ by $Min_{\preceq}(\alpha)$.
2. For a $W \subseteq U$, \preceq is W -smooth iff for every $w \in W$ there is a $v \preceq w$ such that v is \preceq -minimal in W .
3. \preceq is smooth iff it is $M(\alpha)$ -smooth for every α .²
4. A preorder \preceq on U is *faithful* (with respect to a belief set K) iff \preceq is smooth, $x \prec y$ for every $x \in M(K)$ and $y \notin M(K)$, and $x \not\prec y$ for every $x, y \in M(K)$. For an $X \subseteq L$, we say that \preceq is X -faithful iff \preceq is faithful with respect to $Cn(X)$.

²Smoothness is also known as stopperedness [Makinson, 1989] and the limit assumption [Lewis, 1973].

□

Grove regards such preorders as measures of the compatibility of an interpretation with the current beliefs of an agent, whilst interpretations lower down in the preorder are regarded as more compatible. Revision is then defined in terms of a faithful preorder by letting the minimal models of a wff α (the models of α that are most compatible with the current beliefs of the agent) generate the belief set resulting from a revision by α .

(Def * from \preceq) $K * \alpha = Th(Min_{\preceq}(\alpha))$

This approach is a bit more general than Grove's sphere-semantics, since faithful total preorders can be imposed on the interpretations of any (possible-worlds) semantics (U, \models) for L , and not just the interpretations obtained from a system of spheres. In fact, Grove's result can be seen as the special case in which elementarily equivalent interpretations form part of the same equivalence class (modulo the faithful total preorder). With the aid of (Def * from \preceq) and (Def $-$ from $*$), obtaining a definition of contraction in terms of faithful preorders is also a straightforward matter:

(Def \sim from \preceq) $K \sim \alpha = Th(M(K) \cup Min_{\preceq}(\neg\alpha))$

And as expected, the use of faithful *total* preorders characterises AGM theory change.

Theorem 3.2.6 *1. Every faithful total preorder defines an AGM contraction using (Def \sim from \preceq). Conversely, every AGM contraction can be defined in terms of a faithful total preorder using (Def \sim from \preceq).*

*2. Every faithful total preorder defines an AGM revision using (Def * from \preceq). Conversely, every AGM revision can be defined in terms of a faithful total preorder using (Def * from \preceq).*

Proof This result is essentially the same as a result of Peppas and Williams [1995]. The proof draws heavily on similar results in [Gärdenfors, 1988] and [Grove, 1988]. For the reader's convenience, we provide the complete proof in appendix A. □

A recurring theme throughout this dissertation is the advocacy of orderings on infatoms as an adequate representation of epistemic states in many contexts. One of the reasons for advancing this claim is that, in many respects, such orderings seem

to lie at the heart of the construction of belief change operations. Here is the first formal argument in support of such a claim. We show that the AGM contraction and revision defined in terms of the same faithful total preorder can also be defined in terms of each other using the Levi and Harper identities.

Definition 3.2.7 An AGM contraction $-$ and an AGM revision $*$ are *semantically related* iff they can be defined in terms of the same faithful total preorder using (Def \sim from \preceq) and (Def $*$ from \preceq). \square

The notion of semantic relatedness will be extended, as we proceed, to various constructions involving faithful preorders.

Proposition 3.2.8 *Let $-$ be an AGM contraction and $*$ an AGM revision that are semantically related.*

1. $-$ can also be defined in terms of $*$ using (Def $-$ from $*$).
2. $*$ can also be defined in terms of $-$ using (Def $*$ from \sim).

Proof Let \preceq be a faithful total preorder in terms of which $-$ and $*$ can be defined using (Def \sim from \preceq) and (Def $*$ from \preceq). The proof of (1) is trivial and is omitted. For the proof of (2), it suffices to show that $Th(Min_{\preceq}(\alpha)) = Th(M(K) \cup Min_{\preceq}(\alpha)) + \alpha$. If $\neg\alpha \in K$, it follows from lemma 1.3.4, and if $\neg\alpha \notin K$, it follows from the fact that $Min_{\preceq}(\alpha) \subseteq M(K)$. \square

When viewed as orderings on infatoms, a faithful total preorder can be seen as a way of ordering the basic units of information according to their entrenchment (or importance, or credibility), with an infatom higher up in the ordering considered as more entrenched. Recall from part (1) of proposition 3.1.6 that the models of K correspond to the infatoms that do not form part of the semantic content of K , and from part (4) of proposition 3.1.5 that the countermodels of K correspond exactly to the content bits of K . So a faithful total preorder places the content bits of K strictly above the remaining infatoms, which are all placed on the same level. The accompanying intuition is clear. The content bits of K are more entrenched than the infatoms not forming part of the semantic content of K .

3.2.1 The propositional finite case

In the context of theory change, Katsuno and Mendelzon [1991] seem to have been the first to make the transition from Grove's sphere-semantics to faithful total preorders. They investigate theory revision for the simplified case of the finitely generated classical propositional logics (for which (U, \models) is the \models -valuation semantics for L). This simplification ensures that all belief sets can be axiomatised by single wffs, and accordingly, this is the way they choose to represent belief sets. That is, a belief set K is represented by any wff α such that $Cn(\alpha) = K$. For them, a revision is thus a function from L to L . They provide four basic revision postulates, and two supplementary ones.

$$(KM*1) \quad \alpha \in Cn(\phi * \alpha)$$

$$(KM*2) \quad \text{If } \neg\alpha \notin Cn(\phi) \text{ then } Cn(\phi * \alpha) = Cn(\phi \wedge \alpha)$$

$$(KM*3) \quad \text{If } Cn(\phi) = Cn(\psi) \text{ and } \alpha \equiv \beta \text{ then } Cn(\phi * \alpha) = Cn(\psi * \beta)$$

$$(KM*4) \quad \text{If } \not\models \neg\alpha \text{ then } \perp \notin Cn(\phi * \alpha)$$

$$(KM*5) \quad Cn(\phi * (\alpha \wedge \beta)) \subseteq Cn((\phi * \alpha) \wedge \beta)$$

$$(KM*6) \quad \text{If } \neg\beta \notin Cn(\phi * \alpha) \text{ then } Cn((\phi * \alpha) \wedge \beta) \subseteq Cn(\phi * (\alpha \wedge \beta))$$

It is easy to see that these postulates are just the AGM revision postulates phrased to fit in with their representation of belief sets. (KM*1) corresponds to (K*2), (KM*2) is a combination of (K*3) and (K*4), (KM*3) corresponds to (K*5), and (KM*4) combined with (KM*1) give (K*6). Furthermore, (KM*5) and (KM*6) respectively correspond to (K*7) and (K*8).

The method that Katsuno and Mendelzon employ to construct revisions involves the faithful total preorders on U . They use the term *faithful* to refer to an assignment of total preorders to every wff ϕ (representing the belief set $Cn(\phi)$), with \preceq_ϕ satisfying the following three conditions:

1. If $u, v \in M(\phi)$ then $u \not\prec_\phi v$.
2. If $u \in M(\phi)$ and $v \notin M(\phi)$ then $u \prec_\phi v$.
3. If $\phi \equiv \psi$ then $\preceq_\phi = \preceq_\psi$.

We shall refer to these as the KM-faithful total preorders. Since U is finite, every total preorder on U is smooth, and so \preceq_ϕ is clearly a faithful total preorder (with respect to $Cn(\phi)$). In their view, a KM-faithful total preorder is an indication of minimal change (of some sort) on interpretations, a suggestion that is more or less in line with Grove's idea of a measure of compatability. They proceed to show that the postulates (KM*1) to (KM*6) characterise AGM revision.

Theorem 3.2.9 [*Katsuno and Mendelzon, 1991*] *A revision satisfies the postulates (KM*1) to (KM*6) iff there is a KM-faithful total preorder \preceq_ϕ such that $M(\phi * \alpha) = \text{Min}_{\preceq_\phi}(\alpha)$.*

3.2.2 Semantic AGM revision without smoothness

The reason for including smoothness as one of the properties of the faithful preorders is that the lack thereof opens the door for the possibility that a wff α (which is not logically invalid) need not have any minimal models. In such cases, the use of (Def * from \preceq) to define revision will result in the violation of (K*2) and (K*6). Apparently Boutilier [1990, 1994] first noticed that it is possible to do away with smoothness. His idea can be explained as follows. When dealing with total preorders, a lack of smoothness only causes problems for an α -revision if α is not logically invalid and α doesn't have minimal models. And this can only occur if there is an infinite descending chain of models of α . In such situations it makes sense to obtain the belief set resulting from an α -revision by a simple extension of minimality. Instead of taking a wff β to be in $K * \alpha$ iff β is true in all the minimal models of α , we allow β into $K * \alpha$ iff there is some level in the total preorder, below which all models of α are also models of β . Boutilier's setup differs from ours in a number of aspects. He casts L , the language in which an agent expresses his beliefs, into a propositional modal framework, and his construction for defining revision is phrased in terms of modal operators. But it is easy to see that, in effect, he considers the same logics as we do, and that his definition of revision corresponds to (Def * from B) below. Let us refer to a preorder on U (with respect to a belief set K) as *B-faithful* iff the following two conditions hold:

1. If $u, v \in M(K)$ then $u \not\prec v$.
2. If $u \in M(K)$ and $v \notin M(K)$ then $u \prec v$.

So a preorder is faithful iff it is B-faithful and smooth. Boutilier's definition of revision in terms of a B-faithful total preorder \preceq then looks as follows:

$$(\mathbf{Def} * \mathbf{from} \mathbf{B}) \quad \beta \in K * \alpha \text{ iff } \left(\begin{array}{l} \forall w \in M(\alpha), \exists v \preceq w \text{ such that } v \in M(\alpha) \cap M(\beta), \\ \text{and } \forall u \in M(\alpha) \text{ such that } u \preceq v, u \in M(\beta) \end{array} \right)$$

Boutilier shows that this construction can be used to define AGM revision.

Theorem 3.2.10 [Boutilier, 1994] *Every B-faithful total preorder defines an AGM revision using (Def * from B). Conversely, every AGM revision can be defined in terms of a B-faithful total preorder using (Def * from B).*

It is easily verified that, for the faithful total preorders, the identities (Def * from \preceq) and (Def * from B) are equivalent, and Boutilier's construction is thus clearly an extension of the minimal model semantics for revision.

3.3 Orderings as epistemic states

Recall from chapter 1 that the *epistemic state* of an agent has to be represented in a way that, at the very least, ensures the extraction of the beliefs of the agent, as well as the information needed to perform reasoning in a coherent fashion. In the context of AGM theory change, the latter includes the information to decide which of the permissible AGM theory change operations to use. Semantically, it is thus sufficient to represent an epistemic state as an ordered pair (K, \preceq) , where K is a belief set and \preceq is a faithful total preorder. We shall see that such a representation becomes particularly apt when we adopt an information-theoretic view of the faithful preorders, where an infatom higher up in the ordering is regarded as more entrenched. For the moment though, we concentrate on matters more formal, and discuss the connection between semantic AGM theory change and the three construction methods discussed in chapter 2. It turns out that the use of faithful total preorders is already implicitly contained in transitively (and connectively) relational partial meet contraction, safe contraction and epistemic entrenchment. In fact, there is a very strong connection between the faithful total preorders, the orderings used for the construction of the transitively (and connectively) relational partial meet contractions, the epistemic entrenchment orderings, and the hierarchies used by safe contraction. Coupled with the principle of Reductionism, these results provide support for the proposal to represent epistemic states semantically.

3.3.1 Semantic epistemic entrenchment

From an information-theoretic point of view it seems natural to be able to extend the faithful total preorders to orderings on the wffs of L . The basic idea is simply to lift a faithful total preorder (on infatoms) in a sensible way to a *power order* (an ordering on *sets* of infatoms). Because every wff is associated with a particular set of infatoms — its semantic content — we can view the ordering on sets of infatoms as an ordering on wffs. The question of deciding what constitutes a *sensible* way of lifting a faithful total preorder is, of course, largely dependent on the stated purpose of such an ordering on wffs. Recall from section 2.3 that the intuition associated with an epistemic entrenchment ordering is that wffs lower down are less entrenched, and should be given up more easily. So epistemic entrenchment places the emphasis on what should be *discarded* rather than on what should be retained. We can thus think of the level of entrenchment of a wff as being determined by its least entrenched content bits. Accordingly, it seems reasonable to regard β as *at least as entrenched* as α iff every content bit of β is at least as entrenched as some content bit of α . It is in this spirit that we define the *power order* \sqsubseteq in terms of a preorder \preceq on the infatoms of L as follows:

$$\alpha \sqsubseteq \beta \text{ iff for every } j \in C(\beta) \text{ there is an } i \in C(\alpha) \text{ such that } i \preceq j.$$

It turns out that the model-theoretic version of this definition applied to the faithful total preorders yields precisely the EE-orderings of section 2.3.

(Def \sqsubseteq_E from \preceq) $\alpha \sqsubseteq_E \beta$ iff $\forall y \in M(\neg\beta) \exists x \in M(\neg\alpha)$ such that $x \preceq y$

This follows from the relationship between the GE-orderings and the EE-orderings discussed in section 2.3.1, and results in [Grove, 1988, Gärdenfors, 1988, Boutilier, 1992, 1994], showing that the GE-orderings can be defined in terms of the faithful total preorders as follows:

(Def \sqsubseteq_G from \preceq) $\alpha \sqsubseteq_G \beta$ iff $\forall y \in M(\beta) \exists x \in M(\alpha)$ such that $x \preceq y$

Theorem 3.3.1 1. *Every faithful total preorder defines a GE-ordering using (Def \sqsubseteq_G from \preceq). Conversely, every GE-ordering can be defined in terms of a faithful total preorder using (Def \sqsubseteq_G from \preceq).*

2. *Every faithful total preorder defines an EE-ordering using (Def \sqsubseteq_E from \preceq). Conversely, every EE-ordering can be defined in terms of a faithful total preorder using (Def \sqsubseteq_E from \preceq).*

Proof 1. The proof draws heavily on results of Grove [1988], Gärdenfors [1988], Boutilier [1992, 1994]. For the reader's convenience, we provide a complete proof in appendix A.

2. Follows from part (1) and theorem 2.3.5.

□

With the help of part (1) of theorem 3.3.1, the GE-orderings can be defined in terms of AGM revision as follows:

(Def \sqsubseteq_{GE} from $*$) $\alpha \sqsubseteq_{GE} \beta$ iff $\neg\alpha \notin K * (\alpha \vee \beta)$ or $\neg\alpha \notin K$ or $\models \neg\beta$

Proposition 3.3.2 *Let $*$ be an AGM revision. The relation defined in terms of $*$ using (Def \sqsubseteq_{GE} from $*$) is a GE-ordering.*

Proof Let \preceq be a faithful total preorder from which $*$ is obtained using (Def $*$ from \preceq), and consider the GE-ordering \sqsubseteq_{GE} defined in terms of \preceq using (Def \sqsubseteq_{GE} from \preceq). We show that $\alpha \sqsubseteq_{GE} \beta$ iff $\neg\alpha \notin K * (\alpha \vee \beta)$ or $\neg\alpha \notin K$ or $\models \neg\beta$. We only consider the case where $\neg\alpha \in K$ and $\not\models \beta$. Suppose that $\alpha \sqsubseteq_{GE} \beta$. So, for every $y \in M(\beta)$ there is an $x \in M(\alpha)$ such that $x \preceq y$. And hence, for every $y \in \text{Min}_{\preceq}(\beta)$ there is an $x \in \text{Min}_{\preceq}(\alpha)$ such that $x \preceq y$. It thus follows that $\text{Min}_{\preceq}(\alpha) \subseteq \text{Min}_{\preceq}(\alpha \vee \beta)$. So $\text{Min}_{\preceq}(\alpha \vee \beta) \not\subseteq M(\neg\alpha)$ and therefore $\neg\alpha \notin K * (\alpha \vee \beta)$. Conversely, suppose that $\neg\alpha \notin K * (\alpha \vee \beta)$. Then there is an $x \in \text{Min}_{\preceq}(\alpha \vee \beta)$ such that $x \in M(\alpha)$. So, $x \preceq y$ for every $y \in M(\beta)$ and therefore $\alpha \sqsubseteq_{GE} \beta$. □

A reasonable interpretation of part (2) of theorem 3.3.1 is that one should think of the EE-orderings as being derived from the faithful total preorders. This view is also supported by an appeal to the principle of Reductionism, since every EE-ordering is built up from an ordering on infatoms in much the same way that the entailment relation \models is built up from the interpretations of L (or from the infatoms of L). And it is in line with the claim that orderings on infatoms are adequate representations of the epistemic states of an agent. What is more, there is a strong connection between the AGM contraction and the EE-ordering defined in terms of the same faithful total preorder, much along the lines of proposition 3.2.8.

Definition 3.3.3 An AGM contraction and an EE-ordering are *semantically related* iff they can be defined in terms of the same faithful total preorder using (Def \sim from \preceq) and (Def \sqsubseteq_E from \preceq). \square

Proposition 3.3.4 Let $-$ be an AGM contraction and \sqsubseteq_{EE} an EE-ordering that are semantically related.

1. $-$ can also be defined in terms of \sqsubseteq_{EE} using (Def $-$ from \sqsubseteq_{EE}).
2. \sqsubseteq_{EE} can also be defined in terms of $-$ using (Def \sqsubseteq_{EE} from \sim).

Proof Let \preceq be a faithful total preorder in terms of which $-$ and \sqsubseteq_{EE} are defined using (Def \sim from \preceq) and (Def \sqsubseteq_E from \preceq).

1. We need to show that if $\alpha \in K \setminus Cn(\top)$ then $\beta \in K - \alpha$ iff $\beta \in K$ and $\alpha \sqsubseteq_{EE} (\alpha \vee \beta)$ (the remaining case is trivial). It suffices to show that $Min_{\preceq}(\neg\alpha) \subseteq M(\beta)$ iff $\alpha \sqsubseteq_{EE} \alpha \vee \beta$. Now, $Min_{\preceq}(\neg\alpha) \subseteq M(\beta)$ iff $y \in M(\beta)$ for every $y \in Min_{\preceq}(\neg\alpha)$, iff there is a $y \in M(\neg\alpha)$ such that $x \in M(\alpha \vee \beta)$ for every $x \preceq y$, iff $\alpha \vee \beta \not\sqsubseteq_{EE} \alpha$, iff $\alpha \sqsubseteq_{EE} \alpha \vee \beta$.
2. We need to show that if $\not\models \beta$ then $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \notin K - (\alpha \wedge \beta)$ (the remaining case is trivial). Suppose that $\alpha \sqsubseteq_{EE} \beta$. So, for every $y \in M(\neg\beta)$ there is an $x \in M(\neg\alpha)$ such that $x \preceq y$. In particular, for every $y \in Min_{\preceq}(\neg\beta)$ there is an $x \in M(\neg\alpha)$ such that $x \preceq y$. So $Min_{\preceq}(\neg\alpha) \subseteq Min_{\preceq}(\neg(\alpha \wedge \beta))$ and thus $\alpha \notin K - (\alpha \wedge \beta)$. Conversely, suppose that $\alpha \notin K - (\alpha \wedge \beta)$. Then there is a $z \in M(K) \cup Min_{\preceq}(\neg(\alpha \wedge \beta))$ such that $z \in M(\neg\alpha)$. And since $z \preceq y$ for every $y \in M(\neg\beta)$, it follows that $\alpha \sqsubseteq_{EE} \beta$.

\square

3.3.2 The connection with relational partial meet contraction

With proposition 3.2.1 at our disposal, it becomes clear that the use of faithful total preorders can be traced back to the construction of relational partial meet contractions (see section 2.2). Recall that the relational partial meet contractions are constructed with the aid of a binary relation \subseteq on the set of all remainders

$$K \perp L = \{A \in K \perp \alpha \mid \alpha \in L \setminus Cn(\top)\}.$$

Intuitively, \subseteq is seen as an ordering, with elements “higher up” in the relation being regarded as “better”. To obtain a related ordering on interpretations, we reinterpret \subseteq as (the inverse of) a relation, not on remainders, but on the corresponding interpretations, in the sense of proposition 3.2.1. Since proposition 3.2.1 just applies to α -remainders where $\alpha \in K \setminus Cn(\top)$, we use \subseteq restricted to $(K \perp L) \setminus \{K\}$. (Recall that $K \perp \alpha = \{K\}$ iff $\alpha \notin K$.) It is easily verified from proposition 3.2.1 that the interpretations corresponding to the elements of $(K \perp L) \setminus \{K\}$ are precisely the countermodels of K . The corresponding relation \ll on U is then defined as follows:

$$(\mathbf{Def} \ll \mathbf{from} \subseteq) \quad u \ll v \text{ iff } \begin{cases} Th(M(K) \cup \{v\}) \subseteq Th(M(K) \cup \{u\}) \\ \text{if } u, v \notin M(K), \\ u \in M(K) \text{ otherwise.} \end{cases}$$

So \ll orders the countermodels of K inversely to the way \subseteq orders the corresponding elements of $K \perp L$, puts the models of K strictly below the countermodels of K , and places all the models of K equally low down in the ordering. Now define a removal – in terms of \ll as follows:

$$(\mathbf{Def} - \mathbf{from} \ll) \quad K - \alpha = Th(M(K) \cup \{u \in M(\neg\alpha) \mid u \ll v \ \forall v \in M(\neg\alpha)\})$$

That is, instead of taking the intersection of the “best” α -remainders (in terms of \subseteq) to obtain an α -contraction, we add the “best” models of $\neg\alpha$ (in terms of \ll) to $M(K)$ and take $K - \alpha$ to be the corresponding theory. Under the proviso that the function s_K defined in terms of \subseteq using (Def s_K from \subseteq) is indeed a selection function, it is easily verified that – is identical to the partial meet contraction defined in terms of s_K using (Def \sim from s_K). Furthermore, the set of faithful total preorders is clearly a strict subset of the transitive relations (and indeed of the total preorders) on U defined in terms of the transitive relations (and the total preorders respectively) on $K \perp L$ using (Def \ll from \subseteq). And most importantly, every faithful total preorder \preceq is well-behaved in the sense that the removal defined in terms of \preceq using (Def – from \ll) is an AGM contraction. This observation enables us to answer a question posed in section 2.2. To obtain a set of relations on $K \perp L$ that are well-behaved in the sense that the functions they induce using (Def s_K from \subseteq) are selection functions, and for which these selection functions define *all* the AGM contractions when using (Def \sim from s_K), we simply need to obtain the relations on $K \perp L$ corresponding to the faithful total preorders. They are obtained as follows. First we consider the set containing every faithful total preorder \preceq

in which all elementarily equivalent interpretations form part of the same equivalence class (modulo \preceq). Then we show how to obtain the appropriate corresponding relation \Subset on $K \perp L$ from such a faithful total preorder \preceq :

$$(\text{Def } \Subset \text{ from } \preceq) \quad A \Subset B \text{ iff } \begin{cases} w \preceq v \ \forall v, w \notin M(K) \text{ such that} \\ Th(M(K) \cup \{v\}) = A \text{ and } Th(M(K) \cup \{w\}) = B \\ \text{if } A, B \neq K, \\ B = K \text{ otherwise.} \end{cases}$$

It is easily verified that \Subset is a total preorder, and that the function s_K defined in terms of \Subset using (Def s_K from \Subset) is a selection function. So the contraction – defined in terms of s_K using (Def \sim from s_K) is an AGM contraction. In fact, it is easily verified that – is the same contraction as the one defined in terms of \preceq using (Def \sim from \preceq). So this set of total preorders on $K \perp L$ is the set of well-behaved relations on $K \perp L$ referred to in section 2.2. They are all well-behaved in the sense that the functions induced from them using (Def s_K from \Subset) are all selection functions. Furthermore, it follows indirectly from theorem 3.2.4 that all the AGM contractions can be defined in terms of these selection functions using (Def \sim from s_K). And analogous to the situation with the faithful total preorders and the EE-orderings, an appeal to the principle of Reductionism provides support for the claim that the faithful total preorders are more fundamental than the corresponding total preorders on $K \perp L$.

We conclude with a semantic view of full meet contraction and maxichoice contraction, the two limiting cases of partial meet contraction mentioned in section 2.2. From the discussion above it is clear that full meet contraction is obtained semantically (using (Def \sim from \preceq)) from the faithful total preorder on interpretations in which the countermodels of K are all equally comparable. Intuitively, this corresponds to the most cautious form of contraction in which all content bits of K are equally entrenched. On the other hand, those maxichoice contractions that are also AGM contractions, are obtained from the faithful total preorders in which the ordering restricted to the countermodels of K is linear. The intuitive reading of these orderings corresponds to the boldest forms of contraction, in that we are able to distinguish between the entrenchment of *all* the content bits of K .

3.3.3 Safe contraction

Rott [1992b] describes a very strong connection between the EE-orderings and the regular virtually connected hierarchies (see definition 2.4.4). Recall from theorem 2.4.5 that the AGM contractions can be defined in terms of the regular virtually connected hierarchies (over K) using (Def – from \sqsubset_H). A closer look at virtual connectivity shows that when it is applied to a hierarchy, it yields the strict version of a total preorder on K . So the strict version of every EE-ordering, restricted to K , is thus a virtually connected hierarchy. What is more, it is easily verified that every strict version of an EE-ordering restricted to K is also regular. So every strict version of an EE-ordering can also be used as a regular virtually connected hierarchy to define an AGM contraction using (Def – from \sqsubset_H). In a slight abuse of notation we sometimes use the term EE-ordering to refer to the strict version \sqsubset_{EE} of an EE-ordering \sqsubseteq_{EE} . Of course, \sqsubseteq_{EE} can easily be obtained from \sqsubset_{EE} as follows:

$$\sqsubseteq_{EE} = \sqsubset_{EE} \cup \{(\alpha, \beta) \mid \alpha \not\sqsubset_{EE} \beta \text{ and } \beta \not\sqsubset_{EE} \alpha\}.$$

Rott shows the following remarkable connection between the EE-orderings, the regular virtually connected hierarchies, and AGM contraction. Every regular virtually connected hierarchy \sqsubset_H defines an EE-ordering as follows:

(Def \sqsubset_{EE} from \sqsubset_H) $\alpha \sqsubset_{EE} \beta$ iff there is a $B \subseteq K$ such that $B \models \beta$, and for every $A \subseteq K$ such that $A \models \alpha$, it is the case that $A \neq \emptyset$, and for every $\delta \in B$ there is a $\gamma \in A$ such that $\gamma \sqsubset_H \delta$

Furthermore, the regular virtually connected hierarchies defining the same EE-ordering \sqsubset_{EE} includes \sqsubset_{EE} itself, and are precisely those that define the same AGM contraction as well. And finally, every EE-ordering yields the same AGM contraction, whether used as an EE-ordering, or as a regular virtually connected hierarchy. These results from Rott [1992b] are summarised in the following theorem.

Theorem 3.3.5 1. *Let \sqsubset_H be a regular virtually connected hierarchy. The relation defined in terms of \sqsubset_H using (Def \sqsubset_{EE} from \sqsubset_H) is the strict version of an EE-ordering.*

2. *Two regular virtually connected hierarchies define the same AGM contraction using (Def – from \sqsubset_H) iff they also define the same strict version of an EE-ordering using (Def \sqsubset_{EE} from \sqsubset_H).*

3. Let \sqsubset_{EE} be the strict version of an EE-ordering \sqsubseteq_{EE} . If the regular virtually connected hierarchy, obtained by restricting \sqsubset_{EE} to K , is applied to $(\text{Def } \sqsubset_{EE} \text{ from } \sqsubset_H)$, the resulting relation is identical to \sqsubset_{EE} .
4. Let \sqsubset_H be a regular virtually connected hierarchy, and let \sqsubset_{EE} be the strict version of the EE-ordering \sqsubseteq_{EE} , where the former is defined in terms of \sqsubset_H using $(\text{Def } \sqsubset_{EE} \text{ from } \sqsubset_H)$. Then the AGM contractions defined in terms of \sqsubset_H using $(\text{Def } - \text{ from } \sqsubset_H)$ is identical to the AGM contraction defined in terms of \sqsubseteq_{EE} using $(\text{Def } - \text{ from } \sqsubseteq_{EE})$.

So every AGM contraction $-$ can be defined in terms of an equivalence class \mathcal{H} of regular virtually connected hierarchies using $(\text{Def } - \text{ from } \sqsubset_H)$, with \mathcal{H} containing a unique EE-ordering \sqsubset_{EE} . Given these results, it seems reasonable to regard \sqsubset_{EE} as the canonical hierarchy from which $-$ is obtained, especially since \sqsubset_{EE} is also the EE-ordering defined in terms of every element of \mathcal{H} using $(\text{Def } \sqsubset_{EE} \text{ from } \sqsubset_H)$.

3.3.4 Summary

We conclude this discussion with a summary of the semantic connections between AGM contraction and revision, the EE-orderings, the GE-orderings and the regular virtually connected hierarchies.³ Centre stage is occupied by the faithful total preorders, from which all these belief change related operations and orderings can be obtained. To be able to draw the connections properly, it is necessary to work with equivalence classes of faithful total preorders.

Definition 3.3.6 Two faithful preorders \preceq and \preccurlyeq are said to be *minimal-equivalent* iff $Th(\text{Min}_{\preceq}(\alpha)) = Th(\text{Min}_{\preccurlyeq}(\alpha))$ for every $\alpha \in L$. \square

For the finitely generated propositional logics, no two different faithful total preorders will be minimal-equivalent, but as the next example shows, this is not so in the general case.

Example 3.3.7 Let L be the propositional language generated by the set of propositional atoms $\{p_i \mid i \geq 0\}$, and with the standard valuation semantics (V, \models) in which V contains all possible valuations. Furthermore, let

³The orderings on remainders encountered in section 3.3.2 are too closely related to the faithful total preorders to be mentioned separately.

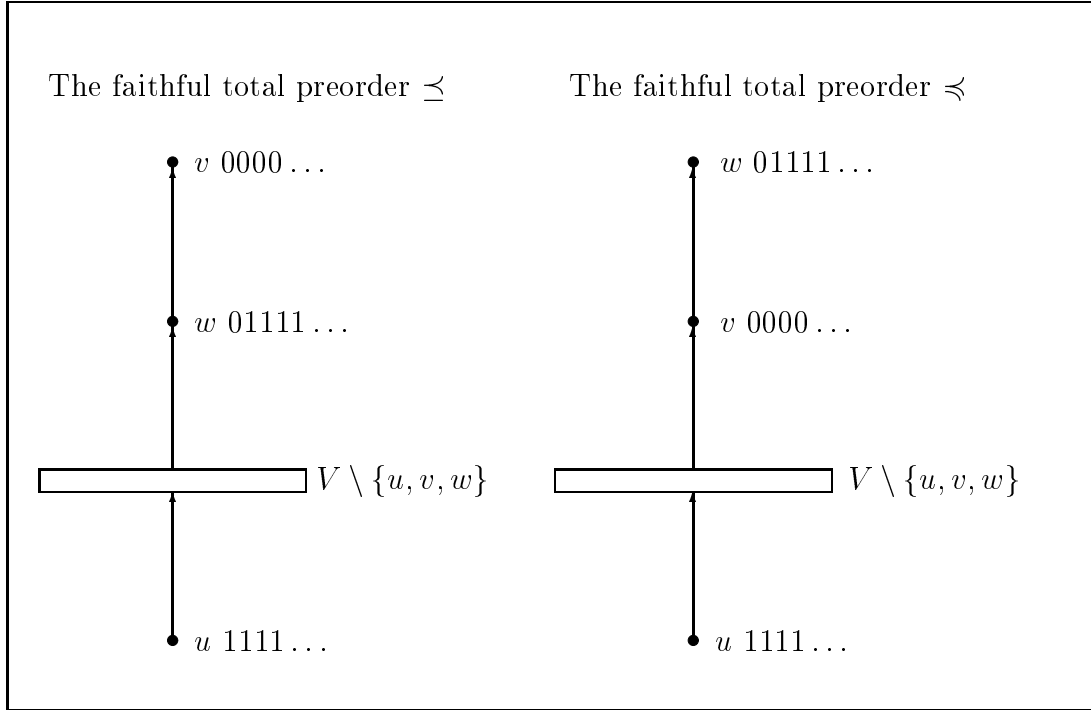


Figure 3.1: The faithful total preorders used in example 3.3.7. The faithful total preorders \preceq and \preccurlyeq are obtained from the reflexive transitive closures of the relations determined by the arrows.

1. u denote the valuation that assigns the value T to all atoms, i.e. $u(p_i) = T$ for every $i \geq 0$,
2. v denote the valuation that assigns the value F to all atoms, i.e. $v(p_i) = F$ for every $i \geq 0$,
3. w denote the valuation that assigns the value T to all atoms except p_0 , i.e. $w(p_i) = T$ for every $i > 0$, and $w(p_0) = F$.

Now let \preceq be the total preorder that places u on its own on the lowest level, followed by all the remaining valuations, except v and w , on the next level, followed by v on the next level, and followed by w on the highest level. Also, let \preccurlyeq be the total preorder that is identical to \preceq , except that v and w exchange positions. Figure 3.1 contains a graphical representation of these two total preorders. Clearly, both \preceq and \preccurlyeq are K -faithful total preorders, where $K = Cn(\{p_i \mid i \geq 0\})$. Furthermore, it is also easily

verified that neither v nor w are minimal models of any wff $\alpha \in L$. And it thus follows that \preceq and \preccurlyeq are minimal-equivalent. \square

Although an equivalence class of minimal-equivalent faithful total preorders may contain a large number of different total preorders, they all have the same relative ordering of the minimal models of every wff in L . For if the minimal models of a wff α are at least as low as the minimal models of β in terms of one member \preceq of such an equivalence class, but not in terms of some other member \preccurlyeq of the same equivalence class, the minimal models of $\alpha \wedge \beta$ cannot be the same in terms of both \preceq and \preccurlyeq . It is therefore easy to see that any two minimal-equivalent faithful total preorders define the same

$$\left\{ \begin{array}{l} \text{AGM contraction} \\ \text{AGM revision} \\ \text{EE-ordering} \\ \text{GE-ordering} \end{array} \right\} \text{ in terms of } \left\{ \begin{array}{l} (\text{Def } \sim \text{ from } \preceq) \\ (\text{Def } * \text{ from } \preceq) \\ (\text{Def } \sqsubseteq_E \text{ from } \preceq) \\ (\text{Def } \sqsubseteq_G \text{ from } \preceq) \end{array} \right\}.$$

In view of these results, it makes sense to generalise definitions 3.2.7 and 3.3.3, and extend the notion of semantic relatedness as follows.

Definition 3.3.8 An AGM contraction, an AGM revision, an EE-ordering and a GE-ordering are *semantically related* iff they can be defined in terms of the same faithful total preorder using (Def \sim from \preceq), (Def $*$ from \preceq), (Def \sqsubseteq_E from \preceq), and (Def \sqsubseteq_G from \preceq). \square

It follows, either directly or indirectly, from theorems 2.3.5, 3.2.6, and 3.3.1, as well as propositions 3.2.8 and 3.3.4, that an

$$\left\{ \begin{array}{l} \text{AGM contraction} \\ \text{AGM contraction} \\ \text{AGM revision} \\ \text{EE-ordering} \end{array} \right\} \text{ and a(n) } \left\{ \begin{array}{l} \text{AGM revision} \\ \text{EE-ordering} \\ \text{GE-ordering} \\ \text{GE-ordering} \end{array} \right\}$$

that are semantically related are interchangeable using

$$\left\{ \begin{array}{l} (\text{Def } - \text{ from } *) \\ (\text{Def } - \text{ from } \sqsubseteq_{EE}) \\ (\text{Def } * \text{ from } \sqsubseteq_{GE}) \\ (\text{Def } \sqsubseteq_E \text{ from } \sqsubseteq_G) \end{array} \right\} \text{ and } \left\{ \begin{array}{l} (\text{Def } * \text{ from } \sim) \\ (\text{Def } \sqsubseteq_{EE} \text{ from } \sim) \\ (\text{Def } \sqsubseteq_{GE} \text{ from } *) \\ (\text{Def } \sqsubseteq_G \text{ from } \sqsubseteq_E) \end{array} \right\}.$$

Figure 3.2 contains a summary of these results, together with the results of theorem 3.3.5.

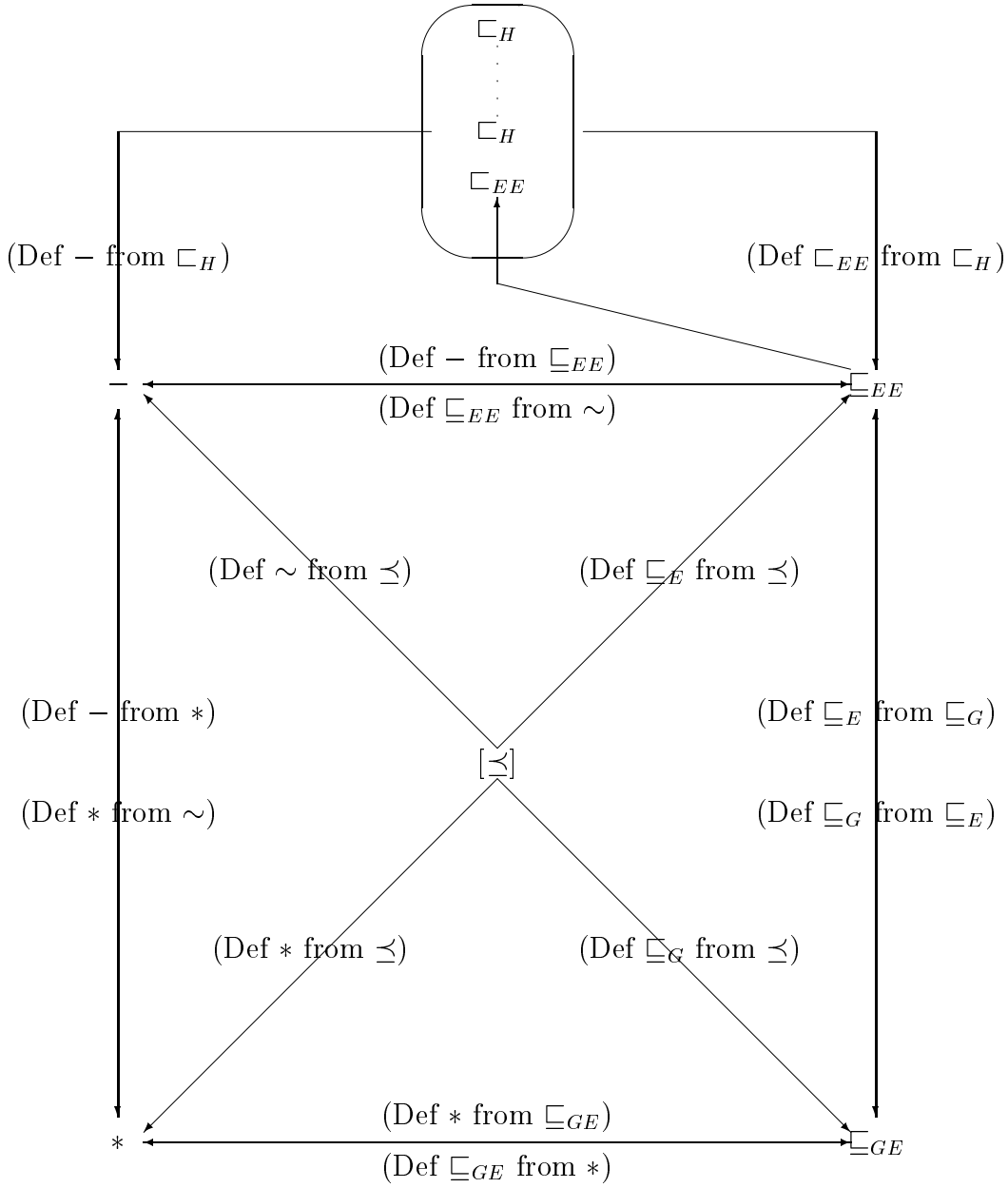


Figure 3.2: The relationship between minimal-equivalent faithful total preorders, and semantically related AGM contractions, AGM revisions, EE-orderings, GE-orderings, as well as safe contractions defined in terms of regular virtually connected hierarchies.

Chapter 4

Nonmonotonic reasoning

We demand guaranteed rigidly defined areas of doubt and uncertainty.

Douglas Adams

The phrase “logical reasoning” is usually associated with the kind of arguments found in mathematical proofs. Perhaps the most essential ingredient of such arguments is truth preservation, which ensures that the truth of the conclusions drawn from a set of assumptions are guaranteed by the truth of the assumptions. Although useful in many areas, an agent equipped solely with reasoning abilities of this kind will soon find itself paralysed and unable to draw almost any conclusion. For, as Benjamin Franklin is so aptly quoted by Matthew Ginsberg [1987] in his introduction to nonmonotonic reasoning, “Nothing is certain but death and taxes.”. To be able to operate at all in a world filled with uncertainties, it is frequently necessary to be able to jump to conclusions of which the truth is *not* sanctioned by the evidence at our disposal. Of course, for this to be seen as *some* kind of reasoning, it will have to be a rational and systematic method of determining what is *plausible*, and not just an arbitrary drawing of inferences in a seemingly random fashion.

Nonmonotonic reasoning is part of the study of such forms of *defeasible* reasoning. A logic is said to be *nonmonotonic* if its associated entailment relation \vdash need not always satisfy the following monotonicity property: if $A \vdash \beta$ then $A \cup \{\alpha\} \vdash \beta$. With \vdash seen as a relation of plausible consequence, there are many examples to show that monotonicity is an undesirable property. Perhaps the one most deeply entrenched in the nonmonotonic reasoning literature is the Tweety example. Given that Tweety is a

bird, it seems plausible to infer that Tweety can fly. But given the additional evidence that Tweety is an ostrich, we should abandon our conclusion about Tweety’s flying capabilities.

To be able to draw plausible conclusions, nonmonotonic reasoning formalisms are usually concerned (whether implicitly or explicitly) with three types of information. Firstly, we have *fixed information*. This includes information such as “ostriches are birds”. Secondly, we have *default information* which consists of information such as “birds normally fly”, and “ostriches normally don’t fly”. Together the fixed and default information provide a *background context* [Geffner, 1992,p. 25]. And thirdly, we have *evidence* such as “Tweety is a bird” and “Chirpy is an ostrich”, containing information specific to the situation at hand. The difference between fixed and default information is that the conclusions drawn from the system may *defeat* default information, but not fixed information. For example, any nonmonotonic reasoning system worth its salt should be able to conclude from the background context and the evidence given above that “Chirpy doesn’t fly”, thus defeating the information that “birds normally fly” (combined with the information that ostriches are birds). But adding the evidence that “Chirpy is an ostrich but not a bird” should render the system inconsistent, since the evidence now conflicts with the fixed information.

A cursory comparison of belief change and nonmonotonic reasoning might create the impression that they have very little in common. After all, the former is concerned with the dynamic process of changing one’s beliefs, while the latter deals with the seemingly static process of jumping to conclusions on the basis of new evidence. As we shall see however, these two fields of research just provide different views of what are essentially identical processes of reasoning. The suggestion of identifying nonmonotonic reasoning with theory change can already be found in [Glymour and Thomason, 1984], but it was only with the subsequent development of general frameworks for both theory change and nonmonotonic reasoning that such a suggestion was properly investigated. In the case of theory change, the relevant framework is that of AGM theory change. For nonmonotonic reasoning the appropriate setting is provided by the nonmonotonic consequence relations of Kraus et al. [1990], and the subsequent extensions proposed by Lehmann and Magidor [1992], and Gärdenfors and Makinson [1994]. We shall therefore focus our attention on these approaches to nonmonotonic reasoning.

4.1 KLM nonmonotonic reasoning

During the 1980s a host of nonmonotonic logics made their appearance, of which the modal systems of McDermott and Doyle [1980, 1982], Moore’s [1984, 1985] autoepistemic logic, Reiter’s [1980] default logic, McCarthy’s [1980, 1986] circumscription [Lifschitz, 1986, 1987], and Poole’s [1988] system for default reasoning are probably the best known. While these systems all have interesting properties when looked at individually, the lack of a general framework for nonmonotonic reasoning made it difficult to compare and evaluate them.

One of the most influential attempts to establish such a general nonmonotonic setting is the KLM approach, named after its three originators Sarit Kraus, Daniel Lehmann and Menachem Magidor [1990]. The success of their approach is largely attributable to their decision to focus on the consequence relations associated with nonmonotonic logics, an idea that seems to have originated with Gabbay [1985]. From a semantic point of view, the work of the KLM trio is an extension of Shoham’s [1987a, 1987b] proposed model theory for nonmonotonic reasoning. As a formal study of consequence relations, it grew out of the work of Gabbay [1985], and has much in common with Makinson’s [1989] theory of cumulative inference, which was developed independently and more or less at the same time. Kraus et al. concern themselves with binary relations, denoted by \sim , on a propositional language L closed under the usual propositional connectives. The semantics for L is assumed to be a valuation semantics (V, \models) as defined in section 1.3, with \models denoting the standard notion of semantic entailment associated with it. As we have shown in section 1.3, every one of the logics we consider can be “converted” into such a propositional logic, which means that the logics permitted by Kraus et al. are precisely those that we consider as well. One of their primary aims is to demarcate those binary relations on L that are deserving of the name “nonmonotonic consequence relation”. Elements of such relations are denoted by expressions of the form $\alpha \sim \beta$ (where α and β are wffs of L), and should be read as “ β is a plausible consequence of α ”, or “if α holds then I am willing to (defeasibly) jump to the conclusion that β holds”.

Of the three types of information used in nonmonotonic reasoning systems, only the evidence is explicitly represented in the KLM setup. In an expression such as $\alpha \sim \beta$, α is the available evidence from which the plausible conclusion β is drawn. The fixed information is coded into the semantics for L , and is represented on the object level

by the logically valid wffs. Default information, on the other hand, should be seen as somehow being encoded into the consequence relation \sim . For example, let $b(t)$ and $f(t)$ be atoms of a transparent propositional language L (see section 1.3), with $b(t)$ representing the assertion that Tweety is a bird, and $f(t)$ representing the claim that Tweety can fly. Then $b(t) \sim f(t)$ is read as “I am willing to jump to the conclusion that Tweety can fly, given that Tweety is a bird”. The fact that such a conclusion seems reasonable can be attributed to the existence of a default rule stating that birds normally fly. But it would be a mistake to think that $b(t) \sim f(t)$ is, or forms part of, such a default rule. Rather, it is the fact that such a default rule is built *into* \sim that allows us to plausibly conclude that Tweety can fly from the evidence that Tweety is a bird. In section 4.6 we discuss these matters in more detail.

4.2 Preferential consequence relations

Formally, the KLM approach to nonmonotonic reasoning mirrors the AGM approach to theory change in many ways. The KLM nonmonotonic consequence relations are defined in terms of sets of postulates. This is followed by a description of semantic methods for constructing these relations, and the statement of representation theorems, proving that the construction methods do indeed yield precisely the set of consequence relations described by the appropriate set of postulates. Four families of consequence relations are studied by Kraus et al. [1990] and Lehmann and Magidor [1992]: cumulative consequence relations, loop-cumulative consequence relations, preferential consequence relations and rational consequence relations. We shall restrict our attention to the preferential consequence relations in this section and to the rational consequence relations in section 4.3.

Definition 4.2.1 A *preferential* consequence relation \sim is a binary relation on L satisfying the postulates Ref, LLE, RW, And, Or and CM given below.¹ \square

(Ref) For every $\alpha \in L$, $\alpha \sim \alpha$ (Reflexivity)

(LLE) If $\alpha \equiv \beta$ and $\alpha \sim \gamma$ then $\beta \sim \gamma$ (Left Logical Equivalence)

(RW) If $\beta \models \gamma$ and $\alpha \sim \beta$ then $\alpha \sim \gamma$ (Right Weakening)

¹This description of the preferential consequence relations is given by Lehmann and Magidor [1992].

(And) If $\alpha \sim \beta$ and $\alpha \sim \gamma$ then $\alpha \sim \beta \wedge \gamma$

(Or) If $\alpha \sim \gamma$ and $\beta \sim \gamma$ then $\alpha \vee \beta \sim \gamma$

(CM) If $\alpha \sim \beta$ and $\alpha \sim \gamma$ then $\alpha \wedge \beta \sim \gamma$ (Cautious Monotonicity)

Reflexivity ensures that α itself is a plausible consequence of α , while Left Logical Equivalence requires different bits of evidence, which happen to be logically equivalent, to have the same plausible consequences. Right Weakening expresses the intuition that anything logically weaker than some plausible consequence of α should also be a plausible consequence of α . The And postulate requires the conjunction of two plausible consequences to be a plausible consequence, while Or stipulates that the same plausible consequence of two different pieces of evidence should also be a plausible consequence of their disjunction. As the name suggests, Cautious Monotonicity is a weakened form of the monotonicity property. In the context of binary consequence relations, the latter can be phrased as follows:

(Mon) If $\alpha \sim \gamma$ then $\alpha \wedge \beta \sim \gamma$ (Monotonicity)

While Monotonicity ensures that a consequence γ of α will also be a consequence of a wff obtained by adding *any* wff β to α , Cautious Monotonicity requires that the wff β added to α has to be a plausible consequence of α . In other words, all the plausible consequences of α are also plausible consequences of $\alpha \wedge \beta$, as long as β is a plausible consequence of α . Kraus et al. mention a number of other properties satisfied by the preferential consequence relations, and it is not that difficult to come up with even more. We limit ourselves below to some intuitively desirable ones, mainly to give the reader a flavour of the characteristics of these consequence relations.

(SC) If $\alpha \models \beta$ then $\alpha \sim \beta$ (Supraclassicality)

(Cut) If $\alpha \wedge \beta \sim \gamma$ and $\alpha \sim \beta$ then $\alpha \sim \gamma$

(Cum) If $\alpha \sim \beta$ then $\alpha \sim \gamma$ iff $\alpha \wedge \beta \sim \gamma$ (Cumulativity)

(Rec) If $\alpha \sim \beta$ and $\beta \sim \alpha$ then $\alpha \sim \gamma$ iff $\beta \sim \gamma$ (Reciprocity)

(Cond) If $\alpha \wedge \beta \sim \gamma$ then $\alpha \sim \beta \rightarrow \gamma$ (Conditionalisation)

Supraclassicality is the very natural condition that anything logically weaker than α should also be a plausible consequence of α . Under the assumption that $\alpha \sim \beta$, Cut can be seen as the converse of Cautious Monotonicity. It ensures that in the process of checking whether γ is a plausible consequence of α , it is sufficient to show that γ is a plausible consequence of α together with any plausible consequence β of α . Cumulativity is just Cut and Cautious Monotonicity thrown together, but is included here because it is an important nontrivial property of nonmonotonic reasoning systems. Together with Reflexivity, Left Logical Equivalence and Right Weakening, it provides a guarantee that adding to α any plausible consequences of α , will not in any way alter the plausible consequences obtained. It is thus markedly different from probabilistically motivated consequence relations in which the expression $\alpha \sim \beta$ is taken to mean that the conditional probability of β given α is above some threshold value. Reciprocity (referred to by Kraus et al. [1990] as Equivalence) shows that if α and β are “equivalent” under \sim , then the plausible consequences of α and β are exactly the same. Conditionalisation (referred to by Kraus et al. [1990] as rule S) is reminiscent of one part of the deduction theorem for classical propositional logic.

4.2.1 A semantics for preferential consequence relations

The method for constructing preferential consequence relations provided by Kraus et al. is semantic in nature and makes use of, what is called, *preferential models*. The idea is to place an ordering on a set of “states”, with the states lower down in the ordering being more “normal”, in some sense. A wff β is then taken to be a plausible consequence of α if β holds in the most normal states in which α holds. Intuitively, it has much in common with Shoham’s [1987a, 1987b] preferential models which, in turn, is a generalisation of the semantics for McCarthy’s circumscription [Lifschitz, 1987]. Technically, it generalises Shoham’s construction in two aspects. Firstly, it draws a clear distinction between the valuations of L and the set of “states”, and places an ordering on the states, not the valuations. A *labelling* function is used to associate every state with a particular valuation. States are thus more general than valuations, since different states may be associated with the same valuation. Secondly, it relaxes Shoham’s requirement that the ordering on interpretations be well-founded (i.e. that there are no infinite descending chains). This generality is needed in the representation theorem that links the preferential consequence relations to the preferential models.

Definition 4.2.2 Let S be any set. We refer to the elements of S as *states*.

1. A *labelling function* for S is a function from S to V , the set of valuations of L .
2. Given a labelling function l for S , the *l -models* of a wff $\alpha \in L$, denoted by $\hat{\alpha}$, is defined as $\hat{\alpha} = \{s \in S \mid l(s) \models \alpha\}$.
3. A *preferential model* P is an ordered triple (S, l, \prec) , where l is a labelling function for S , \prec is a strict partial order on S , and for every $\alpha \in L$, \prec is $\hat{\alpha}$ -smooth (see definition 3.2.5).

□

Given a preferential model P , the *P -induced* consequence relation \sim_P is defined in terms of P as follows:

(Def \sim_P from P) $\alpha \sim_P \beta$ iff for every $s \in S$ that is \prec -minimal in $\hat{\alpha}$, $s \in \hat{\beta}$

Kraus et al. then show that the binary relations on L defined in terms of the preferential models using (Def \sim_P from P) are precisely the preferential consequence relations.

Theorem 4.2.3 [Kraus et al., 1990] *Every binary relation on L defined in terms of a preferential model P using (Def \sim_P from P) is a preferential consequence relation. Conversely, every preferential consequence relation can be defined in terms of some preferential model P using (Def \sim_P from P).*

The insistence on the $\hat{\alpha}$ -smoothness, in the set of l -models, of the strict partial order \prec , for every wff α , is necessary for the satisfaction of Cautious Monotonicity, and it is a much weaker condition than Shoham's requirement that the ordering be well-founded. In fact, if \prec is required to be well-founded, the converse part of theorem 4.2.3 does not hold [Lehmann and Magidor, 1992]. The use of states instead of valuations is also necessary for the converse part of theorem 4.2.3 to hold, as the following example of Kraus et al. [1990] shows.

Example 4.2.4 Let L be the propositional language generated by the atoms p and q , and let (V, \models) be the valuation semantics for L with $V = \{11, 10, 00\}$. Let $P = (S, l, \prec)$ be a preferential model, with $S = \{s_1, s_2, s_3, s_4\}$, $\prec = \{(s_1, s_3), (s_2, s_4)\}$, and with l defined as follows:

$$l(s_1) = 00, l(s_2) = 10, l(s_3) = 11, l(s_4) = 11.$$

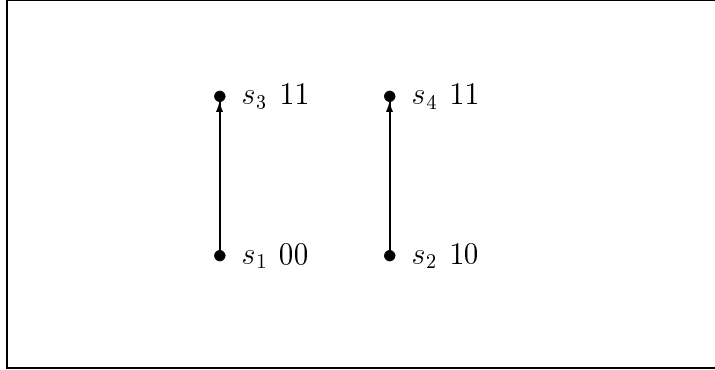


Figure 4.1: The preferential model $P = (S, l, \prec)$ used in example 4.2.4. The strict partial order \prec is determined by the arrows.

Figure 4.1 contains a graphical representation of the preferential model P . We show that the preferential consequence relation \sim_P defined in terms of P using (Def \sim_P from P) cannot be defined in terms of any preferential model whose labelling function is the identity function. Assume, to the contrary, that there is a preferential model $P' = (S', l', \prec')$ for which l' is the identity function, such that the preferential consequence relation $\sim_{P'}$ defined in terms of P' using (Def \sim_P from P) is identical to \sim_P . It is easily verified that $\neg p \wedge \neg q \not\sim_{P'} \perp$, $p \wedge \neg q \not\sim_{P'} \perp$, $p \wedge q \not\sim_{P'} \perp$, but $\neg p \wedge q \sim_{P'} \perp$, from which it follows that $S' = \{s'_1, s'_2, s'_3\}$, with $l'(s'_1) = 00$, $l'(s'_2) = 10$ and $l'(s'_3) = 11$. Furthermore, $(p \wedge q) \vee \neg q \sim_{P'} \neg q$, but $(p \wedge q) \vee \neg q \not\sim_{P'} \neg p \wedge \neg q$ and $(p \wedge q) \vee \neg q \not\sim_{P'} p \wedge \neg q$, which means that the \prec -minimal elements of $(p \wedge q) \vee \neg q$ are the states s'_1 and s'_2 . So either $s'_1 \prec s'_3$ or $s'_2 \prec s'_3$, or both. But from $p \leftrightarrow q \not\sim_{P'} \neg p \wedge \neg q$ and $p \not\sim_{P'} p \wedge \neg q$ it follows respectively that $s'_1 \not\prec s'_3$ and $s'_2 \not\prec s'_3$; a contradiction. \square

It is worth noting at this stage that Kraus et al. see preferential models only as technical tools to aid in the study of the preferential consequence relations, and do not regard the former as suitable representations of the part of an epistemic state pertaining to nonmonotonic reasoning [see Kraus et al., 1990, p. 170].

4.3 Rational consequence relations

There seems to be a fair amount of agreement that any reasonable nonmonotonic consequence relation should at least be a preferential consequence relation. A more

controversial question is whether one should cut down any further by picking out some strict subset of the preferential consequence relations, and if so, how to go about it. A particularly attractive proposal in this regard is the one advanced by Lehmann and Magidor [1992], in which they propose the addition of the following three postulates:

(NR) If $\alpha \sim \beta$ then either $\alpha \wedge \gamma \sim \beta$ or $\alpha \wedge \neg\gamma \sim \beta$ (Negation Rationality)

(DR) If $\alpha \vee \beta \sim \gamma$ then either $\alpha \sim \gamma$ or $\beta \sim \gamma$ (Disjunctive Rationality)

(RM) If $\alpha \sim \gamma$ then either $\alpha \wedge \beta \sim \gamma$ or $\alpha \sim \neg\beta$ (Rational Monotonicity)

Kraus et al. [1990] already considered these postulates, and described them as necessary properties for a rational reasoner. Negation Rationality stipulates that if we regard β as a plausible consequence of α , we must have some reason for doing so. Since exactly one of γ or $\neg\gamma$ holds, it has to be the case that β is a plausible consequence when adding either γ or $\neg\gamma$ to α . Disjunctive Rationality is a slightly generalised version of the same idea. If γ is a plausible consequence of $\alpha \vee \beta$ then, since one of α or β has to hold, γ should be a plausible consequence of either α or β . Rational Monotonicity requires of a reasoner to comply with Monotonicity unless there is a very good reason not to. If γ is a plausible consequence of α then γ should also be a plausible consequence when adding β to α , unless $\neg\beta$ is a plausible consequence of α . It is easily verified that, for each of these three postulates, there is a preferential consequence relation in which it does not hold. In fact, in the presence of the postulates for preferential consequence relations, Rational Monotonicity is strictly stronger than Disjunctive Rationality which, in turn, is strictly stronger than Negation Rationality [Lehmann and Magidor, 1992].

Definition 4.3.1 A *rational* consequence relation is a preferential consequence relation that also satisfies Rational Monotonicity. \square

To obtain a semantic characterisation of the rational consequence relations, we restrict ourselves to those preferential models in which the strict partial orders on states are also modular.

Definition 4.3.2 A strict partial order \prec on a set X is called *modular* iff for every $x, y, z \in X$, if $x \not\prec y$, $y \not\prec x$ and $z \prec x$ then $z \prec y$. \square

The modular strict partial orders are the strict versions of the total preorders on X .

Definition 4.3.3 A *ranked* model is a preferential model $R = (S, l, \prec)$ in which the strict partial order \prec is modular as well. \square

Every modular strict partial order \prec partitions the states into levels, with comparable states being on different levels, and incomparable states considered to be on the same level. We can thus think of \prec as being obtained from a ranking function that ranks states according to normality — the lower the rank of a state, the more normal it is.

Theorem 4.3.4 [Lehmann and Magidor, 1992] *Every binary relation on L defined in terms of a ranked model R using (Def \vdash_P from P) is a rational consequence relation. Conversely, every rational consequence relation can be defined in terms of some ranked model R using (Def \vdash_P from P).*

Lehmann and Magidor regard Rational Monotonicity as a natural condition that should be satisfied by all nonmonotonic consequence relations, and they thus tend to favour the rational consequence relations as *the* set of nonmonotonic consequence relations. This is not a view shared by everyone. For example, Makinson [1994] regards Rational Monotonicity as too strong a condition to insist upon. He argues as follows: If γ is a plausible consequence of α then, even if $\neg\beta$ does not follow plausibly from α , α may still suggest the possibility of $\neg\beta$ strongly enough to undermine the plausibility of γ given $\alpha \wedge \beta$. Makinson is in favour of removing some of the preferential consequence relations, though. He seems to be of the opinion that all nonmonotonic consequence relations should satisfy Disjunctive Rationality. In section 4.4.2 we present an argument supporting the viewpoint of Lehmann and Magidor.

We now come to properties that are *not* satisfied by all rational consequence relations. As expected, it is easily shown that some rational consequence relations do not satisfy Monotonicity. What is perhaps surprising is that some rational consequence relations *do* satisfy Monotonicity. For example, it is easily verified that the entailment relation \models obtained from any valuation semantics (V, \models) for L is a rational consequence relation. One simply needs to examine the ranked model (S, l, \prec) where $S = V$, $l(s) = s$ for every $s \in S$, and \prec is the empty relation. So the classical entailment relations of the logics we consider are all instances of the rational consequence relations! While it might seem strange to include consequence relations that satisfy Monotonicity in a family of nonmonotonic consequence relations, it can be justified as follows. As explained on page 59, the intuition that we are trying to formalise is one of jumping to conclusions in a systematic fashion. And it seems reasonable to include, as a sceptical

extreme, the case where an agent refuses to jump to any conclusions other than those sanctioned by classical logic. So perhaps it is not the inclusion of these monotonic consequence relations that should be called into question, but rather the choice of the name “nonmonotonic reasoning” for this field of study.

Kraus et al. also consider the following three properties.

(EHD) If $\alpha \sim \beta \rightarrow \gamma$ then $\alpha \wedge \beta \sim \gamma$

(Trans) If $\alpha \sim \beta$ and $\beta \sim \gamma$ then $\alpha \sim \gamma$ (Transitivity)

(Cont) If $\alpha \sim \beta$ then $\neg \beta \sim \neg \alpha$ (Contraposition)

EHD is reminiscent of one part of the deduction theorem for classical propositional logic; hence the acronym “EHD” which stands for the “Easy Half of the Deduction theorem”. Kraus et al. show that for the preferential consequence relations, EHD, Transitivity and Monotonicity are equivalent, and Contraposition is stronger than Monotonicity. It is thus clear that these are not suitable properties for nonmonotonic reasoning.

Two properties which are worth considering are given below.

(DP) If $\alpha \sim \gamma$ then either $\alpha \wedge \beta \sim \gamma$ or $\alpha \wedge \beta \sim \neg \gamma$ (Determinacy Preservation)

(CP) If $\alpha \sim \perp$ then $\models \neg \alpha$ (Consistency Preservation)

Like Rational Monotonicity, Determinacy Preservation (which was first suggested by Makinson [see 1994, p. 93]) requires Monotonicity to hold unless there is a very good reason not to. If γ is a plausible consequence of α , then it must also be a plausible consequence obtained when adding β to α , unless the addition of β to α has $\neg \gamma$ as a plausible consequence. It is easily verified that for the preferential consequence relations, Determinacy Preservation is strictly stronger than Rational Monotonicity, and its acceptance would thus amount to a restriction to a strict subset of the rational consequence relations. It is a desirable property in many instances since it promotes considerations of irrelevance. The notion of *irrelevance* involves the idea that irrelevant additional evidence should not influence our plausible consequences. So, given the default information that birds normally fly, we should conclude that Tweety can fly from the evidence that Tweety is a red bird, since being red is irrelevant to Tweety’s flying abilities. In general then, an irrelevance postulate will usually have the following

form: If γ is a plausible consequence of α , and β is irrelevant in terms of γ 's plausibility when presented with α as evidence, then γ is a plausible consequence of $\alpha \wedge \beta$. Of course, the question is how to formalise the intuition that β is irrelevant in terms of γ 's plausibility when α is given as evidence. In the case of Determinacy Preservation, the formalisation comes down to the requirement that $\neg\gamma$ should not be plausible when presented with $\alpha \wedge \beta$ as evidence, i.e., that $\alpha \wedge \beta \not\vdash \neg\gamma$.

In the context of our red bird example above, Determinacy Preservation can be explained as follows. Since we are willing to conclude that Tweety can fly on learning that it is a bird, and since the additional evidence that Tweety is red does not lead us to conclude that Tweety can't fly, we have to conclude that Tweety can fly when presented with the evidence that Tweety is a red bird. Although Determinacy Preservation is appropriate in this example, it is too strong a condition to impose in all situations. For example, from the evidence that Tweety is a bird, it is, as we have argued, reasonable to conclude that Tweety can fly. But when retracting this conclusion on learning that Tweety was spotted in Oudtshoorn — an area in South Africa where ostriches are not uncommon — we do not necessarily want to be forced into concluding that Tweety can't fly. After all, the information about Tweety's whereabouts might raise the possibility that Tweety is an ostrich without rendering it so plausible that one would be willing to act on such a claim.

Consistency Preservation stipulates that logically invalid wffs may only be plausible consequences of logically invalid wffs. It seems to be a reasonable condition, and it is therefore surprising that it is not satisfied by all rational consequence relations. As the next example shows, the failure of Consistency Preservation can be attributed to the fact that the labelling functions of ranked models need not be surjective.

Example 4.3.5 Let L be the propositional language generated by the two atoms p and q , and let (V, \models) be the \models -valuation semantics for L where $V = \{11, 10, 01, 00\}$. Let $R = (S, l, \prec)$ be the ranked model where $S = \{s\}$, $l(s) = 00$ and $\prec = \emptyset$, and let \sim_R be the consequence relation defined in terms of R using (Def \sim_P from P). It is easily verified that $p \sim_R \perp$, even though $\not\models \neg p$, and \sim_R therefore does not satisfy Consistency Preservation. \square

In the next section we shall have more to say about those rational consequence relations that satisfy Consistency Preservation.

4.4 Nonmonotonic reasoning as theory revision

From the results above it should be obvious that there are close similarities between the ranked models that characterise the rational consequence relations and the semantic characterisation of AGM theory revision. Both employ orderings to define some kind of minimal model semantics, although the elements on which the orderings are placed, are not quite the same. Based on these similarities, we can move from theory revision to nonmonotonic reasoning and back as follows [Makinson and Gärdenfors, 1991]: Given a belief set K and a theory revision operation, define a nonmonotonic consequence relation by letting the set of plausible consequences of α coincide with the new belief set obtained from an α -revision of K . Conversely, given a nonmonotonic consequence relation, fix a belief set K in some appropriate fashion and then define an α -revision of K by letting the resulting belief set be equal to the set of plausible consequences of α .

This translation scheme provides a nice way of comparing postulates for theory revision with postulates for nonmonotonic reasoning and vice versa, as is done by Makinson and Gärdenfors [1991] and Gärdenfors and Rott [1995]. One can also use the translation method as the basis for an investigation aimed at discovering the extent to which the two fields of research overlap. The results of Gärdenfors and Makinson [1994], which we describe below, are evidence of the success of such an approach.

4.4.1 Expectation based consequence relations

The fact that the semantic structures used in AGM theory revision and KLM nonmonotonic reasoning are similar, should not be too surprising. The idea of an ordering on worlds or states, with elements lower down (or higher up, as the case may be) in the ordering as somehow being “better”, can be traced back to work done in the 1960s and 1970s on conditional logic and counterfactual reasoning [Lewis, 1973, Adams, 1975, Burgess, 1981, Stalnaker et al., 1981, van Benthem, 1984, Ginsberg, 1986]. Makinson [1993] also provides a survey of research areas employing some kind of minimal model semantics. What *is* perhaps surprising is how easily the structures used in AGM theory revision and KLM nonmonotonic reasoning can be made identical. The two differences to be eliminated are that AGM theory revision uses total preorders, not modular strict partial orders, and places them on the interpretations of L , not on a set of states. From results by Gärdenfors and Makinson [1994] it follows indirectly that these differences

can easily be done away with. We show that those rational consequence relations satisfying Consistency Preservation can be defined in terms of the faithful total preorders as follows:²

(Def \sim from \preceq) $\alpha \sim \beta$ iff $Min_{\preceq}(\alpha) \subseteq M(\beta)$

Definition 4.4.1 An *expectation based* consequence relation is a rational consequence relation that also satisfies Consistency Preservation. \square

The underlying intuition provided by Gärdenfors and Makinson is that the reasoning of an agent is guided, not just by its firm beliefs, but also by its *expectations*. Every expectation based consequence relation \sim is *based* on a set of expectations E , playing a role that is analogous to that of a belief set K in theory change. In fact, every expectation set E is, technically speaking, a belief set, as we shall see below. Intuitively, E is the “current” set of expectations of the agent, and the plausible consequences of a wff α (i.e., every wff β for which $\alpha \sim \beta$ holds) are those wffs that follow logically from α together with “as many as possible” of the elements of E that are compatible with α . The set of expectations E is not explicitly mentioned in the definition of an expectation based consequence relation \sim , but a suitable E can be recovered from \sim as follows.

(Def E from \sim) $E = \{\alpha \mid \top \sim \alpha\}$

That is, E is taken as the set of plausible consequences of any logically valid wff. This process of recovery is justified by noting that the plausible consequences of a wff α can be seen as the “new” set of expectations that an agent would be willing to embrace whenever it is willing to accept α as a new piece of evidence. Now, offering a logically valid wff as a new piece of evidence is just a roundabout way of saying that we are in a situation in which no new evidence has been obtained. And when an agent is not presented with any new evidence, it is reasonable to require that its current set of expectations should not change. Hence the definition of E as in (Def E from \sim).

²In fact, the rational consequence relations *can* be defined in terms of total preorders on interpretations of L , but it includes total preorders on strict subsets of the interpretations of L as well. The inclusion of Consistency Preservation is thus necessary solely for the purpose of ensuring that we only consider the total preorders on all the interpretations of L .

Theorem 4.4.2 *Given any belief set K , every binary relation \sim on L defined in terms of a K -faithful total preorder using (Def \sim from \preceq) is an expectation based consequence relation. Conversely, every expectation based consequence relation \sim can be defined in terms of a K -faithful total preorder \preceq using (Def \sim from \preceq), where K is some satisfiable belief set.*

Proof Pick any belief set K and any K -faithful total preorder \preceq , and let \sim be the binary relation on L defined in terms of \preceq using (Def \sim from \preceq). To show that \sim satisfies Reflexivity, Left Logical Equivalence, Right Weakening, and And, is trivial. For Or, suppose that $\alpha \sim \gamma$ and $\beta \sim \gamma$. That is, $\text{Min}_{\preceq}(\alpha) \subseteq M(\gamma)$ and $\text{Min}_{\preceq}(\beta) \subseteq M(\gamma)$. In the case where at least one of α or β is logically invalid, it follows trivially that $\alpha \vee \beta \sim \gamma$. So we suppose that this is not the case. If the \preceq -minimal models of α are strictly below the \preceq -minimal models of β , then $\text{Min}_{\preceq}(\alpha \vee \beta) = \text{Min}_{\preceq}(\alpha)$, and so $\alpha \vee \beta \sim \gamma$. A similar argument holds if the \preceq -minimal models of β are strictly below the \preceq -minimal models of α . In the remaining case, it follows from the properties of a K -faithful total preorder that $\text{Min}_{\preceq}(\alpha \vee \beta) = \text{Min}_{\preceq}(\alpha) \cup \text{Min}_{\preceq}(\beta)$, from which it also follows that $\alpha \vee \beta \sim \gamma$. For Cautious Monotonicity, suppose that $\alpha \sim \beta$ and $\alpha \sim \gamma$. That is, $\text{Min}_{\preceq}(\alpha) \subseteq M(\beta)$ and $\text{Min}_{\preceq}(\alpha) \subseteq M(\gamma)$. If α is logically invalid then clearly $\alpha \wedge \beta \sim \gamma$. So we suppose that this is not the case. From the $M(\alpha)$ -smoothness of \preceq it then follows that $\text{Min}_{\preceq}(\alpha \wedge \beta) = \text{Min}_{\preceq}(\alpha)$, and so $\alpha \wedge \beta \sim \gamma$. For Rational Monotonicity, suppose that $\alpha \sim \gamma$ and that $\alpha \not\sim \neg\beta$. So $\text{Min}_{\preceq}(\alpha) \subseteq M(\gamma)$ and $\text{Min}_{\preceq}(\alpha) \cap M(\beta) \neq \emptyset$. From the properties of a K -faithful total preorder it then follows that $\text{Min}_{\preceq}(\alpha \wedge \beta) = \text{Min}_{\preceq}(\alpha) \cap M(\beta)$, and so $\alpha \wedge \beta \sim \gamma$. Finally, for Consistency Preservation, suppose that $\alpha \sim \perp$. That is, $\text{Min}_{\preceq}(\alpha) = \emptyset$. By the smoothness of \preceq it then has to be the case that $M(\alpha) = \emptyset$, and so $\models \neg\alpha$.

Conversely, let \sim be any expectation based consequence relation. Now consider the following definition of a binary relation \sim on L in terms of the selection functions of definition 2.2.2.

(Def \sim from s_K) $\alpha \sim \beta$ iff $\beta \in \bigcap \{K' + \alpha \mid K' \in s_K(K \perp \neg\alpha)\}$

A binary relation on L is called *transitively relational* iff it is defined in terms of a selection function s_K using (Def \sim from s_K), where s_K is defined in terms of a transitive relation \subseteq on $K \perp L$ using (Def s_K from \subseteq) (see page 23), and K is some *satisfiable* belief set. Gärdenfors and Makinson [1994] prove that the expectation based consequence relations are precisely the transitively relational binary relations on L . So

there is a satisfiable belief set K , and a transitive relation \subseteq on $K \perp L$ that defines a selection function s_K using (Def s_K from \subseteq), and \vdash can be defined in terms of s_K using (Def \vdash_K from s_K). But by theorem 2.2.6, the K -removal defined in terms of s_K using (Def \sim from s_K) is an AGM K -contraction. By theorem 2.1.6 it then follows that the K -revision $*$ defined as

$$K * \alpha = \bigcap \{K' + \alpha \mid K' \in s_K(K \perp \neg \alpha)\} \text{ for every } \alpha \in L$$

is an AGM K -revision. And by theorem 3.2.6 it follows that there is a K -faithful total preorder \preceq such that $*$ is defined in terms of \preceq using (Def $*$ from \preceq). Since \vdash can be defined in terms of s_K using (Def \vdash from s_K), \vdash can also be defined in terms of \preceq using (Def \vdash from \preceq). \square

With the aid of theorem 4.4.2 we can show that, technically at least, it makes sense to regard every expectation based consequence relation \vdash as being based, not just on the set of plausible consequences of any logically valid wff, but also on the unsatisfiable belief set. The idea is that whenever an expectation based consequence relation \vdash can be defined in terms of a K -faithful total preorder from using (Def \vdash from \preceq), then K is the expectation set on which \vdash is based. The question of whether it is appropriate to view an unsatisfiable belief set as an expectation set will be discussed in section 4.4.2.

Lemma 4.4.3 *Let \vdash be an expectation based consequence relation and let K be the set of wffs defined in terms of \vdash using (Def K from \vdash). Then K is a satisfiable belief set, and \vdash can be defined in terms of at least one K -faithful total preorder, and at least one $Cn(\perp)$ -faithful total preorder using (Def \vdash from \preceq). Furthermore, \vdash cannot be defined in terms of any K' -faithful total preorder, using (Def \vdash from \preceq), for any satisfiable belief set K' that is not equal to K .*

Proof From theorem 4.4.2 it follows that \vdash can be defined in terms of a K'' -faithful total preorder \preceq where K'' is a satisfiable belief set. So

$$K = \{\alpha \mid \top \vdash \alpha\} = Th(Min_{\preceq}(\top)),$$

and since it follows from K'' -faithfulness that $Min_{\preceq}(\top) = M(K'')$, we thus have that $K'' = K$. So we have shown that K is a satisfiable belief set and that \vdash can be defined in terms of at least one K -faithful total preorder using (Def \vdash from \preceq). By noting that

\preceq is also a $Cn(\perp)$ -faithful total preorder, we immediately get that \sim can be defined in terms of at least one $Cn(\perp)$ -faithful total preorder using (Def \sim from \preceq). And finally, pick any satisfiable belief set $K' \neq K$, any K' -faithful total preorder \preceq' , and let \sim' be the expectation based consequence relation defined in terms of \preceq' using (Def \sim from \preceq). So

$$\{\alpha \mid \top \sim' \alpha\} = Th(Min_{\preceq'}(\top)) = Th(M(K')) = K'.$$

For at least one wff β , either $\top \sim \beta$ but $\top \not\sim' \beta$, or $\top \not\sim \beta$ but $\top \sim' \beta$, and so \sim' is not identical to \sim . \square

Lemma 4.4.3 is the justification for the following definition. It allows us to associate expectation sets with expectation based consequence relations in the same way that belief sets are associated with theory change operations.

Definition 4.4.4 An expectation based consequence relation \sim is said to be *based on* E iff either $E = Cn(\perp)$, or E is defined in terms of \sim using (Def E from \sim). For brevity we shall refer to an expectation based consequence relation based on E as an *E -based consequence relation*. \square

From lemma 4.4.3 it follows that the expectation based consequence relations can be partitioned into equivalence classes according to the satisfiable belief sets on which they are based, and that all the expectation based consequence relations are based on the unsatisfiable belief set. This enables us to associate, for every belief set K , the K -based consequence relations with the AGM K -revisions, using the following two definitions, which can be seen as a formalisation of the procedure for translating between theory revision and nonmonotonic reasoning, and vice versa.

(Def \sim from $*$) $\alpha \sim \beta$ iff $\beta \in K * \alpha$

(Def $*$ from \sim) $K * \alpha = \{\beta \mid \alpha \sim \beta\}$

Corollary 4.4.5 Let K be any belief set and let \preceq be any K -faithful total preorder. The AGM K -revision $*$ defined in terms of \preceq using (Def $*$ from \preceq), and the K -based consequence relation \sim defined in terms of \preceq using (Def \sim from \preceq), can also be defined in terms of each other using (Def \sim from $*$) and (Def $*$ from \sim).

Proof The proofs follow easily from theorems 3.2.6 and 4.4.2, and lemma 4.4.3, and are omitted. \square

4.4.2 Expectations, beliefs and epistemic states

The semantic construction of the expectation based consequence relations suggests that the nonmonotonic reasoning abilities of agents can be modelled by ordered pairs of the form (E, \preceq) , where E is a belief set representing the expectations of the agent, and \preceq is an E -faithful total preorder on the interpretations of L . We shall refer to these structures as *expectation states*. An expectation state is thus a part of an epistemic state involved with nonmonotonic reasoning abilities. An information-theoretic view, with \preceq seen as a total preorder on infatoms instead of on interpretations, suggests the following interpretation of expectation states. Think of the expectations of an agent as being built up from infatoms. For a given expectation state (E, \preceq) , E is, of course, built up from $Cont(E)$, the content bits of E . Since \preceq is an E -faithful total preorder, the lowest infatoms in \preceq are precisely those that do *not* form part of $Cont(E)$. The total preorder \preceq should thus be seen as a representation of the extent to which infatoms form part of the current expectations (the content of E). Infatoms higher up in \preceq are less easily dislodged from $Cont(E)$, with the lowest infatoms in the ordering representing the limiting case of those that do not form part of $Cont(E)$ to begin with. The plausible consequences of a wff α are then taken to be all the wffs whose content are included in the set of infatoms, obtained by augmenting the content of α with “as many as possible” of the content bits of E . All that remains is to give a precise description of the phrase “as many as possible”. Now, the only set containing too many infatoms is the set of all infatoms, since it is the only set of infatoms corresponding to an unsatisfiable set of wffs. So, when adding content bits of E to the content of α , the main consideration is to avoid ending up with the set of all infatoms, something that can only occur if the content of E contains all the content bits of $\neg\alpha$. It thus boils down to the question of determining which content bits of $\neg\alpha$ should *not* be added to the content of α . With the help of the total preorder \preceq and the principle of Indifference, the decision is an easy one. The content bits of $\neg\alpha$ not to include, are the ones that are most easily dislodged from $Cont(E)$, i.e. the \preceq -minimal content bits of $\neg\alpha$.

The expectation states and the way they are used to define nonmonotonic reasoning thus coincide exactly with the modelling of theory revision as proposed in section 3.3. So the reasoning *process* employed in nonmonotonic reasoning and theory revision is identical. Does it then follow that nonmonotonic reasoning *is* theory revision (and vice

versa)? Gärdenfors and Makinson argue that this is not the case. Their argument is based on the fact that there is a difference between beliefs and expectations. Fuhrmann and Levi [1994] use this difference as an argument in favour of the claim that there is a difference in the processes of reasoning involved in theory revision and nonmonotonic reasoning. They do not question the appropriateness of AGM revision, but cast doubt on the desirability of Rational Monotonicity for nonmonotonic reasoning. Rabinowicz [1995] provides yet another perspective. He takes issue with the use of “mere expectations” (which he regards as being too weak) when interpreting nonmonotonic reasoning as belief change, and suggests the use of “assumptions”, which are taken to provide a basis for both reasoning and action. But in doing so, he rejects AGM revision as an appropriate framework for dealing with assumptions.

In our opinion, the crux of the matter is to determine what a particular reasoning process is intended to produce. For belief revision it is not an issue. The set of wffs obtained when an agent revises its current set of beliefs, is clearly intended to be the new set of beliefs of the agent. For expectation based nonmonotonic reasoning, though, matters are not so clear. How should we interpret the set of plausible consequences of a wff α ? It is our contention that it cannot be interpreted as anything other than the new set of expectations that an agent is willing to embrace when presented with the evidence α . In other words, expectation based nonmonotonic reasoning is the process of moving from one expectation set to another when confronted with new evidence. The main motivation for this claim centres around the identification of the current set of expectations with the plausible consequences of a logically valid wff, and can be explained as follows. Since the expression $\alpha \sim \beta$ is understood to mean that β is a plausible consequence of the new evidence α at my disposal, it seems reasonable to interpret the situation in which α is a logically valid wff as one in which no new evidence has become available. So my current set of expectations consists of the plausible consequences of the currently available evidence. And it therefore stands to reason that if I am willing to accept α as new evidence, my new set of expectations will be the plausible consequences of α . We shall have more to say about such a dynamic view of nonmonotonic reasoning in section 4.5.

The acceptance of this viewpoint has some interesting consequences for the relation between nonmonotonic reasoning and belief revision. Firstly, it requires the new belief set obtained when revising by a particular wff α , to be a subset of the plausible consequences of α , because the latter is precisely the set of wffs making up the expec-

tation set obtained when accepting α as new evidence. Secondly, it requires of us to associate a reasoning process with nonmonotonic reasoning that is at least as strong as the process of reasoning used in belief revision. A phrase such as “a stronger reasoning process” is, of course, highly ambiguous, but at least one sensible interpretation would be to insist that the postulates satisfied by belief revision should also be satisfied by nonmonotonic reasoning. The acceptance of AGM revision as an appropriate modelling for belief revision then forces us to accept Rational Monotonicity as a desirable postulate for nonmonotonic reasoning.

Our justification of the *dynamic* view of nonmonotonic reasoning presented above is, to a large extent, based on the premise that the current set of expectations of an agent can be identified with the plausible consequences of any logically wff. However, such an identification is slightly at odds with the idea, expressed in lemma 4.4.3, of obtaining the expectation set of an agent from a K -faithful total preorder, since this lemma shows that every expectation based consequence relation is not just based on some satisfiable belief set K , but also on the unsatisfiable belief set. Now, the latter certainly does not correspond to the set of plausible consequences of any logically valid wff (nor, for that matter, does it correspond to the plausible consequences of any wff other than one of the logically invalid ones). This presents us with the following dilemma. Should the unsatisfiable set be seen as an expectation set? A negative answer is not unlike the assumption frequently made in the theory change literature, where the current set of beliefs of an agent is assumed to be satisfiable. Indeed, in the representation results of Gärdenfors and Makinson [1994] that apply to this discussion, they restrict themselves to satisfiable expectation sets. But there are at least two reasons to consider unsatisfiable expectation sets as well. Firstly, in the context of the dynamic character that we attribute to nonmonotonic reasoning, the unsatisfiable belief set is a legitimate expectation set — the one obtained when accepting a logically invalid wff as evidence. And secondly, a broader view of the reasoning abilities of an agent might well include other forms of defeasible reasoning in which the acceptance of evidence, represented by wffs other than logically invalid ones, will give rise to an unsatisfiable belief set.³ On the other hand, if we accept the unsatisfiable belief set as

³A case in point is that of *base change* in which the beliefs of an agent are represented by a base, which is taken to be a set of wffs that is not closed under classical entailment. In such cases, the theory generated by a base may be unsatisfiable, but the base itself might contain enough structure to enable us to define appropriate change operations. See, for example, [Fuhrmann, 1991, p. 186],

a legitimate expectation set, how are we to explain the association of every expectation based consequence relation with two expectation sets, one of which is the unsatisfiable belief? The solution, we believe, is a simple one. It is just a matter of realising that the notion of an expectation state is more fundamental than that of a nonmonotonic consequence relation. An expectation set is thus obtained from an expectation state, not from the derived notion of an expectation based consequence relation. Viewed information-theoretically, an expectation state (E, \preceq) with $E = Cn(\perp)$ should simply be seen as an ordering on infatoms in which the lowest level, containing the infatoms that do not form part of $Cont(E)$, is empty.

In conclusion, there *is* a difference between theory revision and expectation based nonmonotonic reasoning, although the reasoning process involved in both might very well be identical. As a consequence, our representation of an epistemic state ought to be modified such that both forms of reasoning can be recovered from it. That is, an epistemic state should contain, not just the information necessary for performing theory change; it should also incorporate the information in an expectation state so as to be able to perform expectation based nonmonotonic reasoning. In section 4.7 we present one method for doing so.

4.5 A dynamic view of nonmonotonic reasoning

As discussed on page 59, the role of the set of nonmonotonic consequence relations is to provide a framework in which legitimate forms of nonmonotonic reasoning can be expressed. It is usually motivated in terms of examples such as the following. Consider a transparent propositional language containing the atoms $b(t)$, $f(t)$ and $o(t)$, respectively representing the assertion that Tweety is a bird, Tweety can fly, and Tweety is an ostrich. Given the fixed information that ostriches are birds and the default information that birds normally fly, but that ostriches usually don't, it is reasonable to conclude that Tweety can fly when learning that Tweety is a bird, but that Tweety can't fly when obtaining the additional evidence that Tweety is an ostrich. One should thus be able to find at least one nonmonotonic consequence relation \sim containing both $b(t) \sim f(t)$ and $b(t) \wedge o(t) \sim \neg f(t)$.

Examples such as the one above have a definite dynamic flavour to them. It involves the adjustment of the current set of plausible consequences when obtaining the initial

[Hansson, 1993b, p. 641], and chapter 8 of this dissertation.

evidence (that Tweety is a bird), only to be followed by a readjustment when presented with the additional evidence (that Tweety is an ostrich). Given the dynamic nature of such examples, the formalisation presented above offers a curiously *static* view of nonmonotonic reasoning. The idea of additional evidence (that Tweety is an ostrich) being added to the initial evidence (that Tweety is a bird) somehow gets lost in the formalisation of the example. So, although $b(t) \wedge o(t) \sim \neg f(t)$ is intended to signify that the addition of the new evidence $o(t)$ to the initial evidence $b(t)$ will result in $\neg f(t)$ as a plausible consequence, the given interpretation of \sim simply takes it to mean that $\neg f(t)$ is a plausible consequence of $b(t) \wedge o(t)$, and nothing more. There are only two ways to explain this seemingly anomalous behaviour. We can adopt a view of nonmonotonic reasoning as a kind of suppositional reasoning, in which evidence is put forward “for the sake of argument”, only to be discarded again when it has been determined what its plausible consequences would be. Such an interpretation seems to be in line with the aims of conditional logic [Adams, 1975, van Benthem, 1984], but it does not provide an accurate reflection of what nonmonotonic reasoning ought to be. Alternatively, we can attach both a static and a dynamic interpretation to expressions such as $b(t) \wedge o(t) \sim \neg f(t)$. In general then, we would take the expression $\alpha \wedge \beta \sim \gamma$ to mean that γ is a plausible consequence of $\alpha \wedge \beta$, as well as to convey the intuition that, when presented with α as initial evidence, followed by β as additional evidence, we will be able to draw the plausible conclusion γ . The dynamic interpretation describes a process in which evidence is being accumulated systematically, and can be seen as a kind of iterated version of nonmonotonic reasoning. In fact, it ties up nicely with the view of expectation based nonmonotonic reasoning, presented in section 4.4.2, as a process of moving from one expectation set to another when faced with new evidence. Let E be our current set of expectations, i.e. the wffs that we currently regard as plausible, and let \sim_E be the E -based consequence relation describing our current nonmonotonic reasoning process. When confronted with evidence in the form of a wff α , our new set of expectations E' consists of all the plausible consequences of α . But having accepted the evidence α (at least for the moment), there is every reason to believe that modifications will be made, not just to our expectation set E , but also to the very process of nonmonotonic reasoning that we employ. In other words, we don't just move to a new expectation set E' , but also to a new (E' -based) consequence relation $\sim_{E'}^\alpha$, and any additional evidence β will now be evaluated in terms of $\sim_{E'}^\alpha$. The decision to attach both a static and a dynamic interpretation to expressions of the form $\alpha \wedge \beta \sim \gamma$

can thus be formalised as the following property:

$$\text{If } \alpha \models \neg\beta \text{ then } \alpha \wedge \beta \sim_E \gamma \text{ iff } \beta \sim_{E'}^\alpha \gamma$$

where \sim_E is an E -based consequence relation, $E' = \{\delta \mid \alpha \sim_E \delta\}$, and where $\sim_{E'}^\alpha$ is the E' -based consequence relation adopted when presented with the evidence α . In section 7.5.1 we reconsider this property in the context of iterated belief change, and show that, when slightly modified, it has an interesting model-theoretic interpretation.

4.6 Representing default information

Although our description of the three types of information used by nonmonotonic reasoning systems (see page 60) has, for the most part, been of an informal nature thus far, it is clear that both the fixed information and the evidence can be represented adequately by sets of wffs of the language L . (In fact, the approach we have followed only makes provision for single wffs to represent evidence.) When it comes to the representation of default information, however, the situation is not so clear. One solution is to be satisfied with an implicit representation of default rules. For example, suppose L is a transparent propositional language containing the predicate symbols b and f , with b representing the property of being a bird and f the property of being able to fly. Then any expectation based consequence relation \sim containing all the elements of the form $b(x) \sim f(x)$, with x being replaced by all the terms in L , contains an implicit representation of the default rule that “birds normally fly”. This is the viewpoint advanced by Gärdenfors and Makinson [1994, p. 224], at least when it comes to expectation based nonmonotonic reasoning. Of course, such an approach still leaves unanswered the question of how an agent chooses a particular expectation based consequence relation, or equivalently, how it arrives at a particular expectation state.

In many instances though, of which the Tweety example is a case in point, it seems more natural to have an explicit way of representing default information. The question then becomes one of deciding on the most appropriate form of explicit representation. A first attempt might involve the expansion of the language L to introduce another object level connective \rightsquigarrow , which is used to encode default information. Thus, the default rule asserting that birds normally fly might be represented as the set of wffs of the form $b(x) \rightsquigarrow f(x)$, with x being replaced by all the terms in L . But this approach is bound to complicate matters enormously, since such default wffs can then also occur as

fixed information, as evidence and as plausible consequences of the evidence at hand. In fact, it is tantamount to the introduction of a conditional connective into L , and can easily lead to a variant of Gärdenfors' triviality result [Gärdenfors, 1988, pp. 156–166]. A more realistic initial approach, and one that avoids the complicated issue of an agent reasoning about its own reasoning, is to view \rightsquigarrow as a *meta-connective*. The language L thus remains unchanged, but in specifying default information we use expressions of the form $\alpha \rightsquigarrow \beta$, where α and β are wffs of L . This is the method of representing default information in quite a number of recently developed nonmonotonic reasoning systems, [Kraus et al., 1990, Pearl, 1990, Lehmann and Magidor, 1992, Geffner and Pearl, 1992, Goldszmidt and Pearl, 1993, 1996]. These systems are all based on notions independently developed by Lehmann and Magidor [1992] on the one hand, and Pearl [1990] on the other hand, which we briefly discuss below.

Lehmann and Magidor [1992] present three nonmonotonic reasoning systems, all of which involve the specification of a *conditional knowledge base*. In our terminology, a conditional knowledge base CK is a set of default rules of the form $\alpha \rightsquigarrow \beta$, with $\alpha, \beta \in L$. They refer to such default rules as *conditional assertions*. The idea is that one should be able to derive a set of conditional assertions from any conditional knowledge base. When viewed as a binary relation on L , such a derived set of conditional assertions can then be seen as a nonmonotonic consequence relation. So, for example, if we are able to derive the conditional assertion $\alpha \rightsquigarrow \beta$ from CK , we would take β to be a plausible consequence of α in the presence of CK . The question is then to determine which conditional assertions we should be able to derive from a particular conditional knowledge base.

Lehmann and Magidor's first proposal is based on the preferential consequence relations (see definition section 4.2.1), and is termed preferential entailment.

Definition 4.6.1 A conditional knowledge base CK *preferentially entails* a conditional assertion $\alpha \rightsquigarrow \beta$ iff for every preferential consequence relation \sim containing CK (in the sense that $\gamma \sim \delta$ for every $\gamma \rightsquigarrow \delta \in CK$), $\alpha \sim \beta$ holds. \square

So preferential entailment only permits us to draw those plausible conclusions that we will be able to draw from *every* preferential consequence relation respecting the default information contained in CK . This is one of the reasons that it has been advocated by Pearl [1989] as the *conservative core* that should be contained in any nonmonotonic reasoning system.

Clearly the set of conditional assertions preferentially entailed by a conditional knowledge base CK also contains CK , and preferential entailment can thus be seen as a closure operation of some kind. When viewed as a binary relation on L , it turns out that every set of conditional assertions preferentially entailed by some conditional knowledge base is itself a preferential consequence relation. Preferential entailment is thus seen as too weak, since it cannot be described as a set of rational consequence relations.

For Lehmann and Magidor's second proposal of what a conditional knowledge base should entail, they apply the construction used above to the rational consequence relations (see definition 4.3.1).

Definition 4.6.2 A conditional knowledge base CK *rationally entails* a conditional assertion $\alpha \rightsquigarrow \beta$ iff for every rational consequence relation \vdash containing CK (in the sense that $\gamma \vdash \delta$ for every $\gamma \rightsquigarrow \delta \in CK$), $\alpha \vdash \beta$ holds. \square

Remarkably, it turns out that rational entailment is equivalent to preferential entailment. To be more precise, Lehmann and Magidor show that a conditional knowledge base CK rationally entails a conditional assertion $\alpha \rightsquigarrow \beta$ iff CK preferentially entails $\alpha \rightsquigarrow \beta$. Even more remarkable, perhaps, is the fact that for finitely generated propositional languages, Pearl's ϵ -entailment [1988], which is a proposal to deal with default information on qualitative probabilistic grounds, is also equivalent to preferential entailment [Geffner and Pearl, 1992].

Since rational entailment is equivalent to preferential entailment, the former is thus also regarded as too weak. Lehmann and Magidor [Lehmann, 1989, Lehmann and Magidor, 1992] propose to rectify the situation as follows. Consider the set of conditional assertions rationally entailed by a conditional knowledge base CK . Viewed as a binary relation on L , this set is a preferential consequence relation. The idea is to find a sensible way to extend the preferential consequence relation to obtain a rational consequence relation. This rational consequence relation, termed the *rational closure* of CK , seems to be a genuine improvement on rational and preferential entailment, since it is able to handle accounts of irrelevance as well as specificity.⁴ The interested

⁴Specificity refers to the ability to give priority to more "specific" default information. For example, if we know that birds normally fly, but that ostriches normally don't fly, the latter rule should have priority over the former when dealing with a bird that also happens to be an ostrich. See section 4.3 for a description of irrelevance.

reader is invited to consult [Lehmann, 1989] and [Lehmann and Magidor, 1992] for more details. Interestingly enough, Goldszmidt and Pearl [1990] have shown that, for the finitely generated propositional languages, rational closure is equivalent to system-Z [Pearl, 1990], another of Pearl's nonmonotonic reasoning systems based on qualitative probabilities.

4.7 Unifying cautious and bold reasoning

We have seen (see section 4.4.2) that Gärdenfors and Makinson [1994] use the expectation based consequence relations as the basis for a unified treatment of theory revision and nonmonotonic reasoning, arguing that they can be seen as the same process, although used for two different purposes. In this section we show that a closer examination of the Gärdenfors-Makinson claim is the gateway to a theory of cautious and bold reasoning, encompassing both AGM theory revision and nonmonotonic reasoning (in the form of the expectation based consequence relations) as special cases. Such a theory thus provides a truly unified picture of the two areas.

Let us first consider the claim that theory revision and nonmonotonic reasoning can be seen as the same process. With AGM revision and the expectation based consequence relations in mind, the interpretation to attach to this assertion is straightforward. If the belief set K of an agent in a particular situation, and the expectation set E of the same agent in a (possibly) different situation are identical, then the reasoning process involved when revising K by a wff α should be the same as when trying to incorporate the evidence α into E . In other words, the permissible ways of revising K by α should be exactly the same as the permissible ways of obtaining the plausible consequences of α , given the expectation set E . What then, about the statement that this process is used for two different purposes? According to Gärdenfors and Makinson, it boils down to the difference between beliefs and expectations. For them, expectations include not only our beliefs as a limiting case, but also other wffs that are regarded as plausible enough to be used as a basis for inference. The set of expectations of an agent will thus always include its set of beliefs. If we take seriously this relationship between expectations and beliefs, we are one step closer to a unified view of theory revision and nonmonotonic reasoning. For, such a relationship does not just involve the current belief set K and the current set of expectations E . It also requires the belief set obtained when revising K by α to be a subset of the plausible consequences

of α , given the expectation set E . In other words, we should have both $K \subseteq E$ and $K * \alpha \subseteq \{\beta \mid \alpha \sim \beta\}$ for any AGM K -revision $*$ and any E -based consequence relation \sim . In fact, we might as well use AGM revisions to represent expectation based reasoning, as long as we keep in mind that when doing so, the belief sets used are to be interpreted as expectation sets. This is the route we shall take in the remainder of this section.

Gärdenfors and Makinson have chosen to differentiate between beliefs and expectations, but there is nothing preventing us from introducing even further distinctions between sets of conclusions. Rabinowicz [1995], for example, proposes the use of a set of *assumptions*, which is intended to be included in the set of expectations, and also to include the set of beliefs. Formally, there is, of course, no problem with drawing such distinctions. In fact, we might as well continue in this fashion, and make room for an arbitrary finite sequence of sets of wffs, each one including its predecessor and being included in its successor. But what would we be gaining epistemologically? One answer to this question concerns the actions to be taken by agents under various circumstances. For example, Rabinowicz's reason for introducing sets of assumptions is related to his dissatisfaction with the use of "mere expectations" when identifying the process of theory revision with nonmonotonic reasoning. He argues that expectations, as understood by Gärdenfors and Makinson, are too provisional to be used for purposes of deliberation and action, and suggests the use of assumptions instead. It is our view that the qualitative difference between beliefs, assumptions, expectations and the like, can perhaps best be expressed, not in terms of *whether* an agent is willing to act on them, as Rabinowicz contends, but rather in terms of *how* it is willing to act on them. For example, a detective investigating a murder case may be willing to draw tentative conclusions in order to get his investigation going. He may even be willing to act on such conclusions by, for example, following up certain leads. But he may not have sufficient faith in these conclusions to bring suspects in for questioning, or to obtain a warrant for searching the house of the main suspect. And even when the evidence at his disposal provides, in his opinion, sufficient grounds for assuming the main suspect to be guilty, he may not be willing to hand the case over for prosecution. In this example then, his *expectations* might determine his actions related to initial investigative work, his *assumptions* might determine when to take actions with possible negative ramifications, and his *beliefs* might determine when to close the investigation.

Intuitively, such a sequence of sets of wffs thus corresponds to various degrees of

beliefs, with a set of wffs earlier in the sequence being associated with the outcome of a more cautious form of reasoning, and one later in the sequence representing the outcome of a bolder form of reasoning. In the spirit of Gärdenfors and Makinson we take all these different forms of reasoning to be driven by the same reasoning process, and we represent them as a sequence of AGM revisions.

Definition 4.7.1 An *n-reasoning context* is a sequence of ordered pairs

$$\langle (K_1, *_1), \dots, (K_n, *_n) \rangle$$

where, for every i from 1 to n , $*_i$ is an AGM K_i -revision, and for every i from 1 to $n - 1$, and every $\alpha \in L$, $K_i \subseteq K_{i+1}$ and $K_i *_i \alpha \subseteq K_{i+1} *_i \alpha$. \square

We can then, for example, represent a setup involving beliefs, assumptions and expectations as a 3-reasoning context in which K_1 corresponds to the set of beliefs, K_2 to the set of assumptions, and K_3 to the expectation set.

It turns out that in the finitely generated propositional case, at least, the n -reasoning contexts can be constructed elegantly in terms of sequences of successively refined ordered pairs, each consisting of a belief set and a faithful total preorder.

Definition 4.7.2 For any $n > 0$, an *n-refined epistemic state* is an n -tuple of epistemic states $\langle \Phi_1, \dots, \Phi_n \rangle$ (with every Φ_i being an ordered pair (K_i, \preceq_i) , where K_i is a belief set, and \preceq_i is K_i -faithful total preorder) such that, for every i from 1 to $n - 1$:

1. $K_i \subseteq K_{i+1}$,
2. for every $x, y \in U$ (the set of interpretations of L), if $x \equiv_{\preceq_{i+1}} y$ then $x \equiv_{\preceq_i} y$, and
3. for every $x, y \in U$, if $x \prec_i y$ then $x \prec_{i+1} y$.

\square

Intuitively, finer grained total preorders represent more adventurous forms of reasoning. From an information-theoretic point of view, it ensures that an agent is better able to discriminate between infatoms, and will therefore remove fewer infatoms during a revision process.

Theorem 4.7.3 *Pick any positive integer n .*

1. Every n -refined epistemic state $\langle \Phi_1, \dots, \Phi_n \rangle$ defines an n -reasoning context

$$\langle (K_1, *_1), \dots, (K_n, *_{*n}) \rangle$$

by letting every $*_i$ be the AGM K_i -revision defined in terms of \preceq_i using (Def * from \preceq).

2. If L is a finitely generated propositional language, then every n -reasoning context $\langle (K_1, *_1), \dots, (K_n, *_{*n}) \rangle$ can be defined in terms of some n -refined epistemic state $\langle \Phi_1, \dots, \Phi_n \rangle$ for which every $*_i$ is defined in terms of \preceq_i using (Def * from \preceq).

Proof 1. Pick any i such that $1 \leq i < n$ and any $\alpha \in L$. It suffices to show that $\text{Min}_{\preceq_{i+1}}(\alpha) \subseteq \text{Min}_{\preceq_i}(\alpha)$. So pick any $y \in \text{Min}_{\preceq_{i+1}}(\alpha)$. For every $x \in M(\alpha)$, $x \not\prec_{i+1} y$, and so $x \not\prec_i y$. That is, $y \in \text{Min}_{\preceq_i}(\alpha)$.

2. From theorem 3.2.6 it follows that every $*_i$ can be defined in terms of some K_i -faithful total preorder \preceq_i using (Def * from \preceq). Pick any i such that $1 \leq i < n$, and any $x, y \in U$, and let α be a wff that axiomatises the set $\{x, y\}$. We need to show that $x \equiv_{\preceq_{i+1}} y$ implies $x \equiv_{\preceq_i} y$, and $x \prec_i y$ implies $x \prec_{i+1} y$. Suppose firstly that $x \equiv_{\preceq_{i+1}} y$ but that $x \not\equiv_{\preceq_i} y$. Without loss of generality we can assume that $x \prec_i y$. It then follows that $\text{Min}_{\preceq_i}(\alpha) = \{x\}$ and $\text{Min}_{\preceq_{i+1}}(\alpha) = \{x, y\}$. But then $M(K_i *_{*i} \alpha) \subset M(K_{i+1} *_{*i+1} \alpha)$, contradicting the supposition that $K_i *_{*i} \alpha \subseteq K_{i+1} *_{*i+1} \alpha$. Next, suppose that $x \prec_i y$ but that $x \not\prec_{i+1} y$, i.e. $y \preceq_{i+1} x$. Then $\text{Min}_{\preceq_i}(\alpha) = \{x\}$ and either $\text{Min}_{\preceq_{i+1}}(\alpha) = \{x, y\}$ or $\text{Min}_{\preceq_{i+1}}(\alpha) = \{y\}$. Either way, $\text{Min}_{\preceq_{i+1}}(\alpha) \not\subseteq \text{Min}_{\preceq_i}(\alpha)$, and so $M(K_{i+1} *_{*i+1} \alpha) \not\subseteq M(K_i *_{*i} \alpha)$, contradicting the supposition that $K_i *_{*i} \alpha \subseteq K_{i+1} *_{*i+1} \alpha$. □

So the n -refined epistemic states provide a suitable abstract framework for a unified view of cautious and bold reasoning, including both AGM theory revision and expectation based inference.

4.8 Conclusion

Although sometimes viewed as two distinct albeit related fields, theory revision and nonmonotonic reasoning seem to be two sides of the same coin. In recent years, the

research conducted in these two areas have become more and more entwined, with frequent attempts at attaining a kind of synergy on various levels [Makinson, 1989, 1993, Katsuno and Mendelzon, 1991, Katsuno and Satoh, 1991, Lindström, 1991, Gärdenfors and Makinson, 1994, Boutilier, 1994, Goldszmidt and Pearl, 1996].

One of the reasons for the view that theory revision and nonmonotonic reasoning are motivated by different ideas, is that theory revision is usually seen as a description of the dynamic process of an agent modifying its set of beliefs, while nonmonotonic reasoning is viewed as the study of the seemingly static notion of jumping to conclusions in the face of uncertainty.⁵ But, as we have seen in sections 4.4.2 and 4.5, a closer look at the intuition underlying nonmonotonic reasoning reveals it to be of a dynamic nature as well. In fact, the word “nonmonotonic” can be seen as a reference to the willingness of an agent to modify its current set of plausible conclusions in the face of *additional* conflicting evidence. The prevalence of the static view of nonmonotonic reasoning is perhaps attributable to the fact that many of the nonmonotonic reasoning formalisms are firmly rooted in work originally done in the area of conditional logic [Adams, 1975, Stalnaker et al., 1981, van Benthem, 1984].

As we have shown in section 4.5 the accommodation of the dynamic nature of nonmonotonic reasoning in these formalisms is made possible by making certain implicit assumptions. In section 7.5.1 we show how these assumptions can be translated into a property of iterated theory revision, thus providing another example of nonmonotonic reasoning as theory revision.

In conclusion, it is clear that research involving both nonmonotonic reasoning and theory change will benefit both areas. As a contribution along these lines, we have presented a general theory of bold and cautious reasoning, with AGM theory revision and expectation based reasoning as special cases. From our perspective, though, the important advantage resulting from the comparison of theory revision and nonmonotonic reasoning presented in this chapter, is that it provides more support for the use of faithful total preorders as an appropriate way to represent parts of the epistemic states of agents.

⁵See Veltman [1996] for a different view.

Chapter 5

Epistemic entrenchment

Good order is the foundation of all things.

Edmund Burke (1729-97), Irish-born British politician

As indicated in chapter 1, belief change is concerned with the ability of an agent to modify its current view of the world in a coherent fashion when confronted with new information. To be able to effect such modifications, it is necessary to find a way to represent the epistemic states of agents. In our view, an appropriate representation of an epistemic state, at least in the case of theory contraction and revision, is as an ordered pair of the form (K, \preceq) , where K is a belief set and \preceq is a faithful preorder. But this is not the only possibility. Other proposals include a representation as a set of “conditional assertions” (see section 4.6 and Darwiche and Pearl [1997, p. 2]), and as an ordering of entrenchment among the wffs of L , [Nayak, 1994b, Nayak et al., 1996]. Our focus in this chapter is on the latter proposal.

The best-known version of such entrenchment orderings is the EE-orderings of Gärdenfors and Makinson [Gärdenfors, 1988, Gärdenfors and Makinson, 1988], discussed in section 2.3 and again in section 3.3.1. In this chapter we consider them yet again. We show how to formalise the intuition underlying the definition of AGM contraction in terms of the EE-orderings. Then we focus on new results regarding the relationship between the EE-orderings and the faithful total preorders. This leads to a surprising connection between the EE-orderings and the orderings on wffs obtained from Spohn’s ordinal conditional functions [1988].

Section 5.5 of this chapter is an expanded version of the paper by Meyer et al. [1999b].

The EE-orderings provide an adequate formalisation of the intuitive notion of the entrenchment of beliefs in most respects, but they also have some undesirable properties. We take a look at other approaches to entrenchment and discuss the extent to which they circumvent the problems associated with the EE-orderings. This includes a presentation of our own proposal for entrenchment, a refined version of the EE-orderings that is, perhaps not surprisingly, motivated by semantic considerations.

This chapter contains references to virtually every variation on entrenchment that has been put forward in the belief change literature. Remarkably, each and every one of these can, in some way or another, be constructed semantically in terms of some ordering on interpretations or infatoms; a result that is, in part, summarised in figure 5.6 on page 136.

5.1 AGM contraction via the EE-orderings

The intuition ascribed to the EE-orderings is that wffs higher up in the ordering are more entrenched in the belief set K . When forced to choose, we should thus rather discard the less entrenched wffs. This is Gärdenfors' intuitive description of contraction via epistemic entrenchment [1988, p. 89]; an intuition that is not in exact accordance with (Def – from \sqsubseteq_{EE}), the formal definition of AGM contraction in terms of the EE-orderings. (We discuss this matter in more detail in chapter 6.) In this section we show that it is possible to formalise the intuition above, by specifying exactly what it means to say that we are “forced to choose”. We describe AGM contraction in terms of the EE-orderings in a way that differs from (Def – from \sqsubseteq_{EE}). In doing so, we make use of the following identities:

(Def sc_{\sqsubseteq}) $sc_{\sqsubseteq}(\alpha) = \{\beta \mid \alpha \sqsubset \beta\}$

(Def ∇_{\preceq}) $\nabla_{\preceq}(\alpha) = \{x \mid \exists y \in Min_{\preceq}(\alpha), \text{ such that } x \preceq y\}$

Definition 5.1.1 1. Given a preorder \sqsubseteq on L , and a wff α , we define $sc_{\sqsubseteq}(\alpha)$, the *strict cut* of α , in terms of \sqsubseteq using (Def sc_{\sqsubseteq}).

2. For a faithful preorder \preceq (which need not be total), we define $\nabla_{\preceq}(\alpha)$, the *downset* of a wff α in terms of \preceq using (Def ∇_{\preceq}).

□

A strict cut of a wff α contains all the wffs that are more entrenched than α in \sqsubseteq . Strict cuts can be seen as the strict versions of the fallbacks of Lindström and Rabinowicz [1991]. On the other hand, for the faithful total preorders, $\nabla_{\preceq}(\alpha)$ is the set of interpretations that are not strictly above the minimal models of α . We show that there is a close connection between strict cuts and downsets.

Proposition 5.1.2 *Let \preceq be a faithful total preorder, and let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq). If $\not\models \alpha$ then $Th(\nabla_{\preceq}(\neg\alpha)) = sc_{\sqsubseteq_{EE}}(\alpha)$.*

Proof Suppose that $\not\models \alpha$ and pick any $\beta \in Th(\nabla_{\preceq}(\neg\alpha))$. It suffices to show that $\beta \not\sqsubseteq_{EE} \alpha$. Because $Min_{\preceq}(\neg\alpha) \subseteq \nabla_{\preceq}(\neg\alpha)$, there is a $y \in Min_{\preceq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \preceq y$, and thus $\beta \not\sqsubseteq_{EE} \alpha$. Conversely, pick a $\beta \in sc_{\sqsubseteq_{EE}}(\alpha)$, and pick any $y \in Min_{\preceq}(\neg\alpha)$. Since $\beta \not\sqsubseteq_{EE} \alpha$, $x \in M(\beta)$ for every $x \preceq y$, from which it follows that $\nabla_{\preceq}(\neg\alpha) \subseteq M(\beta)$. \square

Proposition 5.1.2 enables us to show that the wffs that are strictly more entrenched than α form the *core* of the wffs to be retained during an α -contraction of K .

Proposition 5.1.3 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq). If $\not\models \alpha$ then $sc_{\sqsubseteq_{EE}}(\alpha) \subseteq K - \alpha$.*

Proof Follows easily from proposition 5.1.2. \square

The remaining question is thus to determine which of the wffs that are at most as entrenched as α will be retained, and which will be discarded during an α -contraction of K . The intuition dictates that we only remove those wffs that we are forced to remove. Given proposition 5.1.3, it is clear that a wff β in K will have to be discarded if α is entailed by β together with $sc_{\sqsubseteq_{EE}}(\alpha)$.

Proposition 5.1.4 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq). If $\not\models \alpha$ and $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta$ then $\beta \notin K - \alpha$.*

Proof Suppose that $\not\models \alpha$ and $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta$. By proposition 5.1.3, $sc_{\sqsubseteq_{EE}}(\alpha) \subseteq K - \alpha$ and so, if $\beta \in K - \alpha$, then $\alpha \in K - \alpha$, contradicting (K-4). \square

Furthermore, if the addition of any two wffs β and γ to $sc_{\sqsubseteq_{EE}}(\alpha)$ yields α , then both will have to be removed from K , even when adding either of them on their own to $sc_{\sqsubseteq_{EE}}(\alpha)$ does not entail α .

Proposition 5.1.5 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq). Now suppose that $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta$ and $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \gamma$, but $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge \gamma$. Then $\beta \notin K - \alpha$ and $\gamma \notin K - \alpha$.*

Proof Because $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \gamma$, we have that $\not\models \alpha$, and by proposition 5.1.2 there is an $x \in M(Th(\nabla_{\preceq}(\neg\alpha)))$ such that $x \in M(\neg\alpha) \cap M(\gamma)$. Furthermore, since $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge \gamma$, $x \in M(\neg\beta)$. Since $M(K) \subseteq \nabla_{\preceq}(\neg\alpha)$, it then follows from lemma 1.3.5 that $x \in M(Th(M(K) \cup Min_{\preceq}(\neg\alpha)))$ and so $\beta \notin K - \alpha$. The proof for $\gamma \notin K - \alpha$ is similar. \square

And finally, we get a result that places an upper bound on the wffs to be removed from K . The next proposition ensures that there is a good reason for discarding a wff $\beta \in K$ during an α -contraction of K : We'll always be able to find a wff that is at least as entrenched as β , and which, together with β and the core, entail α .

Proposition 5.1.6 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq). For every $\beta \in K \setminus K - \alpha$ there is a $\gamma \in K$ such that $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \gamma$ and $\beta \sqsubseteq_{EE} \gamma$, but $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge \gamma$.*

Proof Pick any $\beta \in K \setminus K - \alpha$. We show that $\beta \rightarrow \alpha$ has the desired properties. Since $\beta \in K \setminus K - \alpha$, it follows that $\alpha \in K$, and so $\beta \rightarrow \alpha \in K$. Furthermore, because $\beta \in K \setminus K - \alpha$, there is an $x \in Min_{\preceq}(\neg\alpha) \subseteq \nabla_{\preceq}(\neg\alpha)$ such that $x \in M(\neg\alpha) \cap M(\neg\beta)$. So $x \in M(\beta \rightarrow \alpha) \cap \nabla_{\preceq}(\neg\alpha) \subseteq M(sc_{\sqsubseteq_{EE}}(\alpha) + \beta \rightarrow \alpha)$ by proposition 5.1.2, and thus $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta \rightarrow \alpha$. To show that $\beta \sqsubseteq_{EE} \beta \rightarrow \alpha$ it is enough to point out that $x \preceq y$ for every $y \in M(\neg(\beta \rightarrow \alpha))$, and to recall that $x \in M(\neg\beta)$. Finally, $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge (\beta \rightarrow \alpha)$ because $\{\beta, \beta \rightarrow \alpha\} \models \alpha$. \square

Combining the results above, we obtain the following representation theorem.

Theorem 5.1.7 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq). For every $\alpha, \beta \in L$,*

$$\beta \notin K - \alpha \text{ iff } \begin{cases} \beta \notin K, \text{ or} \\ \not\models \alpha \text{ and } \alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta, \text{ or} \\ \not\models \alpha \text{ and } \exists \gamma \in K \text{ such that } \alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta, \beta \sqsubseteq_{EE} \gamma, \\ \text{and } \alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge \gamma. \end{cases}$$

Proof Suppose that $\beta \notin K - \alpha$, that $\beta \in K$ and that either $\models \alpha$ or $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta$. $\models \alpha$ contradicts (K-6) and the fact that $\beta \in K \setminus K - \alpha$, so we suppose that $\not\models \alpha$ and $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta$. Then the required result follows from proposition 5.1.6. Conversely, if $\beta \notin K$ then by (K-2), $\beta \notin K - \alpha$. If $\not\models \alpha$ and $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta$ then $\beta \notin K - \alpha$ by proposition 5.1.4. So suppose that $\beta \in K$, $\not\models \alpha$, $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \beta$ and that there is a $\gamma \in K$ such that $\alpha \notin sc_{\sqsubseteq_{EE}}(\alpha) + \gamma$, $\beta \sqsubseteq_{EE} \gamma$, and $\alpha \in sc_{\sqsubseteq_{EE}}(\alpha) + \beta \wedge \gamma$. Then $\beta \notin K - \alpha$ by proposition 5.1.5. \square

Theorem 5.1.7 shows that a wff $\beta \in K$ will be discarded during an α -contraction of K for precisely one of the following two reasons:

- If α is entailed by β together with $sc_{\sqsubseteq_{EE}}(\alpha)$.
- If α is entailed by β together with $sc_{\sqsubseteq_{EE}}(\alpha)$ and some wff γ that is at least as entrenched as β .

So, during an α -contraction of K , we say that we are “forced to choose” between two wffs β and γ iff the core of wffs to be retained (the set $sc_{\sqsubseteq_{EE}}(\alpha)$) entails α when both β and γ are added to it.

5.2 EE-orderings and minimality

Implicit in the semantic description of the EE-orderings in section 3.3.1, is the idea that the entrenchment of wffs is a derived notion, based on orderings of interpretations, or perhaps more aptly, orderings of infatoms. Of course, theorem 3.3.1 also guarantees the construction of faithful total preorders in terms of some kind of converse of (Def \sqsubseteq_E from \preceq), leaving the door open for a view of the EE-orderings as at least as basic, epistemologically, as the faithful total preorders. Nevertheless, a number of

other factors make it difficult to escape the conclusion that the latter is the more fundamental of the two. In the first place, an appeal to the principle of Reductionism becomes appropriate in this context if we adopt the view that the EE-orderings are built up from orderings on infatoms, in much the same way that classical entailment relations are built up from the interpretations (or infatoms) of L . In addition, there is also the fact that different minimal-equivalent faithful total preorders (see definition 3.3.6) may define the same EE-ordering using (Def \sqsubseteq_E from \preceq). This last realisation is, in fact, the key to some important results about the connection between the EE-orderings and the minimal models of the wffs of L . The verification of these results is based on the following useful observations concerning the relationship between power orders on L , in the sense of (Def \sqsubseteq_E from \preceq), and the preorders from which they are obtained. These technical results will again prove to be most useful in section 5.5, where we shall have occasion to make extensive use of them without explicitly referring to lemma 5.2.1.

Lemma 5.2.1 *Let \preceq be any preorder (not necessarily total), and let \sqsubseteq be the ordering on L defined in terms of (Def \sqsubseteq_E from \preceq).*

1. $\alpha \sqsubseteq \beta$ iff for every $y \in \text{Min}_{\preceq}(\neg\beta)$ there is an $x \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \preceq y$.
2. $\alpha \not\sqsubseteq \beta$ iff there is a $y \in \text{Min}_{\preceq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \preceq y$.

Proof 1. Suppose that $\alpha \sqsubseteq \beta$ and pick any $y \in \text{Min}_{\preceq}(\neg\beta)$. From (Def \sqsubseteq_E from \preceq) it follows that there is a $z \in M(\neg\alpha)$ such that $x \preceq y$. By the smoothness of \preceq , there is an $x \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \preceq z$, and the required result then follows from transitivity. Conversely, suppose that for every $y \in \text{Min}_{\preceq}(\neg\beta)$ there is an $x \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \preceq y$, and pick any $v \in M(\neg\beta)$. By the smoothness of \preceq there is a $v' \in \text{Min}_{\preceq}(\neg\beta)$ such that $v' \preceq v$, and by supposition there is a $u \in \text{Min}_{\preceq}(\neg\alpha)$ such that $u \preceq v'$. The required result then follows from transitivity.

2. Suppose that $\alpha \not\sqsubseteq \beta$. That is, there is a $v \in M(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \preceq v$. The required result then follows from smoothness and transitivity. The converse follows immediately from the supposition that there is a $y \in \text{Min}_{\preceq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \preceq y$.

□

For any faithful *total* preorder, the set of minimal models of any wff is particularly well-behaved in the sense that they are all on the same level. As a consequence, the application of lemma 5.2.1 to the faithful total preorders and the EE-orderings obtained from them, using (Def \sqsubseteq_E from \preceq), leads to some interesting results. It shows that the EE-orderings can be completely determined by the minimal models of the wffs of L . In the non-trivial case of two wffs α and β that are both not logically valid, β is at most as entrenched as α iff the minimal models of $\neg\beta$ are at least as high up as the minimal models of $\neg\alpha$, and α will be strictly more entrenched than β if and only if the minimal models of $\neg\alpha$ are strictly below the minimal models of $\neg\beta$. Moreover, any two wffs α and β are equally entrenched if and only if the minimal models of $\neg\alpha$ and $\neg\beta$ are on the same level. In the next section we show that these results provide an interesting connection between the GE-orderings of Grove (see section 2.3.1), the EE-orderings, and the orderings on wffs obtained from Spohn's ordinal conditional functions [1988].

Corollary 5.2.2 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and let \sqsubseteq_{GE} be the GE-ordering defined in terms of \preceq using (Def \sqsubseteq_G from \preceq).¹*

1. *If $\not\models \alpha$ and $\not\models \beta$ then $\alpha \sqsubseteq_{EE} \beta$ iff $Min_{\preceq}(\neg\alpha) \preceq Min_{\preceq}(\neg\beta)$.*
2. *If $\not\models \alpha$ and $\not\models \beta$ then $\alpha \sqsubset_{EE} \beta$ iff $Min_{\preceq}(\neg\alpha) \prec Min_{\preceq}(\neg\beta)$.*
3. *$\alpha \equiv_{\sqsubseteq_{EE}} \beta$ iff $Min_{\preceq}(\neg\alpha) \equiv_{\preceq} Min_{\preceq}(\neg\beta)$.²*
4. *If $\not\models \neg\alpha$ and $\not\models \neg\beta$ then $\alpha \sqsubseteq_{GE} \beta$ iff $Min_{\preceq}(\alpha) \preceq Min_{\preceq}(\beta)$.*
5. *If $\not\models \neg\alpha$ and $\not\models \neg\beta$ then $\alpha \sqsubset_{GE} \beta$ iff $Min_{\preceq}(\alpha) \prec Min_{\preceq}(\beta)$.*
6. *$\alpha \equiv_{\sqsubseteq_{GE}} \beta$ iff $Min_{\preceq}(\alpha) \equiv_{\preceq} Min_{\preceq}(\beta)$.*

Proof Follows easily from lemma 5.2.1 and theorem 2.3.5. □

From an information-theoretic point of view, the results concerning the EE-orderings are particularly illuminating. Recall that a faithful total preorder can be seen as an

¹See section 1.3 for an explanation of the convention of applying \preceq , \prec and \equiv_{\preceq} to sets of interpretations.

²This is a well-known result in the context of Grove's systems of spheres [see Gärdenfors, 1988, p. 95].

ordering in which infatoms lower down are regarded as less entrenched. The entrenchment of a wff is thus completely determined by its least entrenched content bits, a view that is reminiscent of the saying that a chain is only as strong as its weakest link. It can be seen as a generalisation of a result by Gärdenfors and Makinson [1988], that when a belief set K is finite modulo Cn , an EE-ordering with respect to K is completely determined by the co-atoms of K , where the co-atoms of K are the logically weakest elements of $K \setminus Cn(\top)$.

5.3 Ordinal conditional functions

Spohn [1988] presents a representation of epistemic states inspired by probability theory. Let us restrict ourselves to a valuation semantics (V, \models) for L . Spohn defines an *ordinal conditional function* (OCF) κ to be a function from V , the set of valuations of L , into the class of ordinals, such that $\kappa(v) = 0$ for at least one $v \in V$. Intuitively, valuations with a smaller ordinal assigned to them are considered to be more plausible. The valuations assigned the ordinal 0 are thus seen as the most plausible, and consequently the current belief set is defined as $K_\kappa = Th(\{v \mid \kappa(v) = 0\})$. Since κ has to assign the ordinal 0 to at least one element of V , K_κ will always be satisfiable.

Clearly any OCF κ induces a total preorder \preceq on V as follows:

(Def \preceq from κ) $v \preceq w$ iff $\kappa(v) \leq \kappa(w)$

In fact, since every subset of V has a smallest ordinal associated with it, \preceq is a well-order, which means it will also be smooth (see definition 3.2.5). Also, \preceq will be a K_κ -faithful total preorder, provided that $\{v \mid \kappa(v) = 0\} = M(K_\kappa)$.³ Some K_κ -faithful total preorders, however, are not well-orders, and they can thus not be defined in terms of any OCF κ using (Def \preceq from κ). In this sense ordinal conditional functions are less general than faithful total preorders. On the other hand, the reference to ordinals ensures that OCFs allow for a representation of degrees of belief that is more sophisticated than any such notion defined in terms of faithful total preorders.

Spohn extends the ordinal conditional functions to functions from $\wp V \setminus \{\emptyset\}$ into the class of ordinals by associating every non-empty subset W of V with the smallest ordinal assigned to any of the valuations in W . That is, for any OCF κ , he defines

³The requirement that $\{v \mid \kappa(v) = 0\} = M(K_\kappa)$ is a technical restriction that can be traced back to the non-axiomatisability of infinitely generated propositional languages.

$\kappa(W) = \min\{\kappa(w) \mid w \in W\}$, and the extended κ thus defines a total preorder on $\wp V \setminus \{\emptyset\}$. It is easily verified that for every OCF κ and every $W \in \wp V \setminus \{V\}$, $\kappa(w) = 0$ for every $w \in W$ iff $\kappa(V \setminus W) > 0$. As a consequence, K_κ can also be described as the set of all wffs α such that $\kappa(M(\neg\alpha)) > 0$.

Since every wff of L is associated with a particular set of valuations — its set of models — every OCF κ defines a total preorder on L as follows:⁴

$$(\mathbf{Def} \sqsubseteq_\kappa \mathbf{from} \kappa) \quad \alpha \sqsubseteq_\kappa \beta \text{ iff } \begin{cases} \kappa(M(\alpha)) \leq \kappa(M(\beta)) & \text{if } \not\models \neg\alpha \text{ and } \not\models \neg\beta, \\ \models \neg\alpha & \text{otherwise} \end{cases}$$

Remarkably, it turns out that these orderings on wffs are instances of the GE-orderings of Grove.

Proposition 5.3.1 *Let κ be an OCF, let \preceq be the total preorder on V defined in terms of κ using (Def \preceq from κ), and let \sqsubseteq_κ be the total preorder on L , defined in terms of κ using (Def \sqsubseteq_κ from κ).*

1. \sqsubseteq_κ can also be defined in terms of \preceq using (Def \sqsubseteq_G from \preceq), where \preceq is the total preorder on V obtained from κ .
2. \sqsubseteq_κ is a GE-ordering.

Proof 1. The non-trivial cases, i.e. for satisfiable wffs that are not logically valid, follow from the definition of the extended κ and part (4) of corollary 5.2.2.

2. If \preceq is a K_κ -faithful total preorder, the result follows from part (1) and theorem 3.3.1. The case where \preceq is not K_κ -faithful corresponds to a violation of the technical restriction that the lowest level of \preceq has to contain all the models of K_κ . It is easily verified that in such a case, \preceq also defines a GE-ordering. From part (1) we then get the required result.

□

We now come to Spohn's definition of the plausibility of wffs. He takes a wff α to be *less* plausible than a wff β iff $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$ or $\kappa(M(\beta)) < \kappa(M(\alpha))$. Since κ only assigns ordinals to non-empty sets of valuations, we let this definition apply only to satisfiable wffs that are not logically valid. In order to accommodate all the wffs of

⁴The OCF determines the relationship between all satisfiable wffs. For completeness, we include the logically invalid wffs by placing them strictly below the satisfiable ones.

L , we extend the definition by letting the logically valid wffs be more plausible than all the other wffs, and letting the logically invalid wffs be less plausible than all the others. The plausibility ordering \sqsubset_P is then defined in terms of an OCF κ as follows:

$$(\text{Def } \sqsubset_P \text{ from } \kappa) \quad \alpha \sqsubset_P \beta \text{ iff } \begin{cases} \kappa(M(\neg\alpha)) < \kappa(M(\neg\beta)) \text{ or } \kappa(M(\beta)) < \kappa(M(\alpha)) \\ \text{if } \not\models \alpha, \not\models \neg\alpha, \not\models \beta, \text{ and } \not\models \neg\beta, \\ \models \neg\alpha \text{ and } \not\models \neg\beta, \text{ or } \not\models \alpha \text{ and } \models \beta, \text{ otherwise} \end{cases}$$

Spohn justifies his definition of plausibility in terms of the firmness with which a wff is believed or disbelieved. The basic idea is that if α is K_κ -established (i.e. $\kappa(M(\neg\alpha)) > 0$) then α is believed with a firmness of $\kappa(M(\neg\alpha))$, if $\neg\alpha$ is K_κ -refuted (i.e. $\kappa(M(\alpha)) > 0$), then α is disbelieved with a firmness of $\kappa(M(\alpha))$, and if α is K_κ -undecided (i.e. $\kappa(M(\alpha)) = \kappa(M(\neg\alpha)) = 0$), then α and $\neg\alpha$ are both believed and disbelieved, with a firmness of 0. A wff α will thus be less plausible than a wff β for one of the following reasons (where both α and β are satisfiable but not logically valid):

1. α is K_κ -established and β is K_κ -established. That is, α is disbelieved and β is believed. Then $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$ and $\kappa(M(\beta)) < \kappa(M(\alpha))$.
2. α is K_κ -undecided and β is K_κ -established. That is, α is less firmly believed than β . Then $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$.
3. Both α and β are K_κ -established and α is less firmly believed than β . Then $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$.
4. α is K_κ -refuted and β is K_κ -undecided. So α is more firmly disbelieved than β . Then $\kappa(M(\beta)) < \kappa(M(\alpha))$.
5. Both α and β are K_κ -refuted and α is more firmly disbelieved than β . Then $\kappa(M(\beta)) < \kappa(M(\alpha))$.

There is another way to justify (Def \sqsubset_P from κ) as a suitable proposal for obtaining plausibility orderings as well; one that involves the connection between the ordinal conditional functions and the GE-orderings. Recall that one of the primary purposes of an EE-ordering (with respect to K) is to compare the wffs in K . It regards all wffs that are not in K as equally entrenched. Similarly (see section 2.3.1), a GE-ordering distinguishes between wffs that are K -refuted, but regards all the wffs in K , together with all the K -undecided wffs, as equally plausible. Now suppose that we want to

obtain an entrenchment ordering that combines the best of the EE-orderings and the GE-orderings.⁵ We need to consider three cases. Firstly, wffs not in K need to be placed strictly below wffs in K . Secondly, when comparing wffs in K we need to use an EE-ordering. And thirdly, when comparing wffs not in K , we need to use the *inverse* of a GE-ordering (since GE-orderings regard wffs lower down as more plausible). We therefore define a *refined* ordering, an R-ordering \sqsubseteq_R in terms of a faithful total preorder \preceq as follows:

$$(\text{Def } \sqsubseteq_R \text{ from } \preceq) \quad \alpha \sqsubseteq_R \beta \text{ iff } \begin{cases} \forall y \in M(\neg\beta), \exists x \in M(\neg\alpha) \text{ such that } x \preceq y \\ \text{if } \alpha, \beta \in K, \\ \forall y \in M(\alpha), \exists x \in M(\beta) \text{ such that } x \preceq y \\ \text{if } \alpha, \beta \notin K, \\ \alpha \notin K \text{ and } \beta \in K \text{ otherwise} \end{cases}$$

It is easily shown that (Def \sqsubseteq_R from \preceq) is a formalisation of the verbal description given above.

Proposition 5.3.2 *Let \preceq be a faithful total preorder, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), let \sqsubseteq_{GE} be the GE-ordering defined in terms of \preceq using (Def \sqsubseteq_G from \preceq), and let \sqsubseteq_R be the R-ordering defined in terms of \preceq using (Def \sqsubseteq_R from \preceq). Then*

$$\alpha \sqsubseteq_R \beta \text{ iff } \begin{cases} \alpha \sqsubseteq_{EE} \beta \text{ if } \alpha, \beta \in K, \\ \beta \sqsubseteq_{GE} \alpha \text{ if } \alpha, \beta \notin K, \\ \alpha \notin K \text{ and } \beta \in K \text{ otherwise.} \end{cases}$$

Proof Follows from theorem 2.3.5. □

It turns out that Spohn's plausibility orderings are instances of the strict versions of the R-orderings.

Theorem 5.3.3 *Let κ be an OCF, let \preceq be the K_κ -faithful total preorder obtained in terms of κ using (Def \preceq from κ), let \sqsubset_P be the plausibility ordering obtained in terms of κ using (Def \sqsubset_P from κ), and let \sqsubseteq_R be the R-ordering obtained in terms of \preceq using (Def \sqsubseteq_R from \preceq). Then $\alpha \sqsubset_P \beta$ iff $\alpha \sqsubseteq_R \beta$ for every $\alpha, \beta \in L$.*

⁵This is a suggestion due to Rabinowicz [1995].

Proof We only consider the case where $\not\models \alpha$, $\not\models \neg\alpha$, $\not\models \beta$ and $\not\models \neg\beta$. Firstly, note that it follows readily from proposition 5.3.2 that the strict version of \sqsubseteq_R can be described as follows:

$$\alpha \sqsubset_R \beta \text{ iff } \begin{cases} \alpha \sqsubseteq_{EE} \beta \text{ if } \alpha, \beta \in K_\kappa, \\ \beta \sqsubseteq_{GE} \alpha \text{ if } \alpha, \beta \notin K_\kappa, \\ \alpha \notin K_\kappa \text{ and } \beta \in K_\kappa \text{ otherwise.} \end{cases}$$

Now suppose that $\alpha \sqsubset_P \beta$. That is, $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$ or $\kappa(M(\beta)) < \kappa(M(\alpha))$. If $\alpha, \beta \in K_\kappa$, then $\kappa(M(\alpha)) = \kappa(M(\beta)) = 0$, and therefore $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$. But this means that $\text{Min}_\preceq(\neg\alpha) \prec \text{Min}_\preceq(\neg\beta)$. By proposition 5.3.2 it then follows that $\alpha \sqsubseteq_{EE} \beta$, and so $\alpha \sqsubset_R \beta$. If $\alpha, \beta \notin K_\kappa$, then $\kappa(M(\neg\alpha)) = \kappa(M(\neg\beta)) = 0$, and therefore $\kappa(M(\beta)) < \kappa(M(\alpha))$. But this means that $\text{Min}_\preceq(\beta) \prec \text{Min}_\preceq(\alpha)$. By proposition 5.3.2 it then follows that $\beta \sqsubseteq_{GE} \alpha$, and so $\alpha \sqsubset_R \beta$. Then the only remaining possibility is for α not to be in K_κ and for β to be in K_κ . For if $\alpha \in K_\kappa$ and $\beta \notin K_\kappa$ then $\kappa(M(\alpha)) = 0$ and $\kappa(M(\neg\beta)) = 0$, contradicting the supposition that $\alpha \sqsubset_P \beta$. So we again have that $\alpha \sqsubset_R \beta$.

Conversely, suppose that $\alpha \sqsubset_R \beta$. If $\alpha, \beta \in K_\kappa$ then $\alpha \sqsubseteq_{EE} \beta$, and so, by corollary 5.2.2, $\text{Min}_\preceq(\neg\alpha) \prec \text{Min}_\preceq(\neg\beta)$. But this means that $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$, and so $\alpha \sqsubset_P \beta$. If $\alpha, \beta \notin K_\kappa$ then $\beta \sqsubseteq_{GE} \alpha$, and by corollary 5.2.2, $\text{Min}_\preceq(\beta) \prec \text{Min}_\preceq(\alpha)$. But then $\kappa(M(\beta)) < \kappa(M(\alpha))$, and so $\alpha \sqsubset_P \beta$. So we are left with the case where $\alpha \notin K_\kappa$ and $\beta \in K_\kappa$, which means that $\kappa(M(\neg\alpha)) = 0$ and $\kappa(M(\neg\beta)) > 0$. So $\kappa(M(\neg\alpha)) < \kappa(M(\neg\beta))$ and therefore $\alpha \sqsubset_P \beta$. \square

We shall encounter the ordinal conditional functions again in section 7.1 in the context of iterated belief change.

5.4 Generalised epistemic entrenchment

The EE-orderings of Gärdenfors and Makinson provide a satisfactory formalisation of the intuition of the entrenchment of wffs in many ways, but they have drawn criticism from various quarters, mainly for being too restrictive in three aspects [Lindström and Rabinowicz, 1991, Rott, 1992c, Gärdenfors and Makinson, 1994, Rabinowicz, 1995]. The first, and most serious objection, is that every EE-ordering is a *total* preorder. This has the unfortunate consequence of ruling out any kind of formal representation of the idea that some wffs are not comparable in terms of entrenchment. A second

objection concerns the minimality condition, imposed on the EE-orderings in the guise of the postulate (EE4). It ensures that the EE-orderings do not distinguish between wffs that are not in K , and are thus unable to give a proper account of entrenchment among wffs that are not in the current belief set of an agent. That this is undesirable is highlighted by the realisation that an agent cannot regard a wff α as being more entrenched than its negation $\neg\alpha$ without accepting α into its current set of beliefs. And thirdly, there is resistance to the maximality condition, imposed on the EE-orderings in terms of the postulate (EE5), which requires the most entrenched wffs to be nothing other than the logically valid wffs. In this section we consider proposals intended to rectify these shortcomings by providing entrenchment orderings that generalise the EE-orderings in one way or another.

5.4.1 LR-entrenchment

Lindström and Rabinowicz [1991] propose a generalised version of the EE-orderings aimed at rectifying the first objection mentioned above, subject to the following set of postulates:

(LR1) \sqsubseteq_{LR} is transitive.

(LR2) If $\alpha \models \beta$ then $\alpha \sqsubseteq_{LR} \beta$

(LR3) If $\alpha \sqsubseteq_{LR} \beta$ and $\alpha \sqsubseteq_{LR} \gamma$ then $\alpha \sqsubseteq_{LR} \beta \wedge \gamma$

(LR4) If $K \neq Cn(\perp)$ then $\alpha \notin K$ iff $\alpha \sqsubseteq_{LR} \beta$ for all β

(LR5) If $\top \sqsubseteq_{LR} \alpha$ then $\models \alpha$

Definition 5.4.1 A binary relation \sqsubseteq_{LR} on L is an *LR-ordering* (with respect to a belief set K) iff it satisfies (LR1) to (LR5). \square

With the exception of (LR3), which replaces the postulate (EE3), and (LR5), which is equivalent to (EE5) in the presence of (LR1) and (LR2), the LR-postulates are identical to the postulates for the EE-orderings. (LR3) is a weakened version of (EE3), and its adoption in the place of (EE3) ensures the possibility that wffs in K need not all be comparable. In fact, it is easy to see that if we only consider those LR-orderings in which all wff are comparable, we end up with precisely the EE-orderings. To see why,

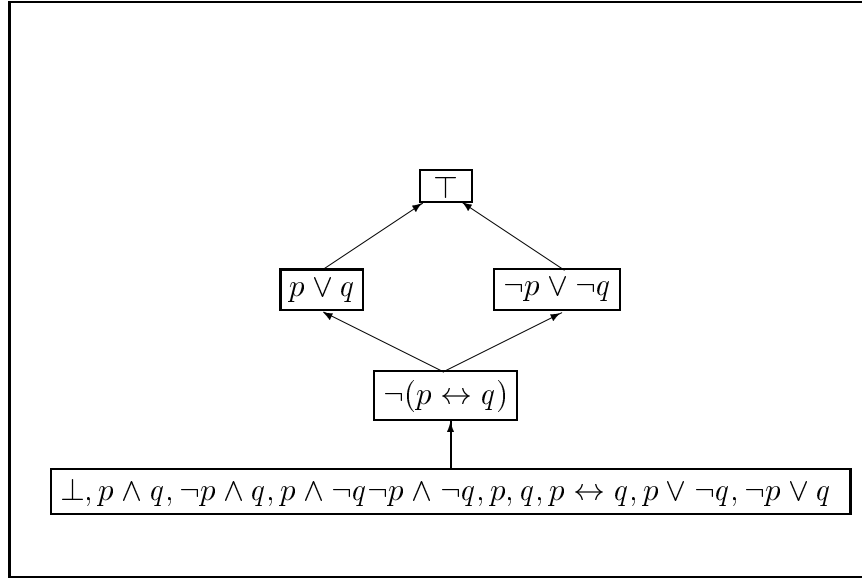


Figure 5.1: A graphical representation of the LR-ordering used in example 5.4.2. The ordering is obtained from the reflexive transitive closure of the relation determined by the arrows. Every wff in the figure is a canonical representative of the set of wffs that are logically equivalent to it.

note that it follows from (LR3) that if $\alpha \sqsubseteq_{LR} \beta$ then $\alpha \sqsubseteq_{LR} \alpha \wedge \beta$. Now, if all wffs are comparable, then we have either $\alpha \sqsubseteq_{LR} \beta$ or $\beta \sqsubseteq_{LR} \alpha$, from which we immediately get that $\alpha \sqsubseteq_{LR} \alpha \wedge \beta$ or $\beta \sqsubseteq_{LR} \alpha \wedge \beta$.

LR-entrenchment is thus a generalisation of the EE-orderings, but is it a proper generalisation? That is, are there any LR-orderings for which some wffs are indeed not comparable? The answer to this question is provided by the following simple example.

Example 5.4.2 Let L be the propositional language generated by the two atoms p and q , and let (V, \models) be the valuation semantics for L where $V = \{11, 10, 01, 00\}$. Now let $K = Cn(\neg(p \leftrightarrow q))$, and define the LR-ordering \sqsubseteq_{LR} as follows: $\alpha \sqsubseteq_{LR} \beta$ iff $\alpha \models \beta$ or $\alpha \notin K$. It is easily verified that \sqsubseteq_{LR} is indeed an LR-ordering. Figure 5.1 contains a graphical representation of \sqsubseteq_{LR} , from which it is easily seen that $\neg p \vee q$ and $p \vee q$ are incomparable. \square

Lindström and Rabinowicz provide a method for constructing the LR-orderings in terms of *fallback families* and prove an appropriate representation theorem. More interesting, from our perspective, is their second construction method. They show that the LR-orderings can also be obtained as the intersections of families of EE-orderings.

Theorem 5.4.3 [*Lindström and Rabinowicz, 1991*]

1. *The intersection of every family of EE-orderings is an LR-ordering.*
2. *For every LR-ordering \sqsubseteq_{LR} there is a family \mathcal{E} of EE-orderings such that $\sqsubseteq_{LR} = \cap \mathcal{E}$.*

In this view of the LR-orderings, the epistemic state of an agent is taken to be a class of EE-orderings. An appeal to the principle of Indifference then results in the construction of an entrenchment ordering in which a wff α is seen as at most as entrenched as a wff β iff every EE-ordering in \mathcal{E} regards α as at most as entrenched as β .

5.4.2 GEE-entrenchment

Rott [1992c] takes the view that it is more natural to consider *strict* relations on wffs and argues that the EE-orderings should be seen as converse complements of such strict relations (or equivalently, that these strict relations be obtained as the converse complements of the EE-orderings).⁶ He defines a set of generalised epistemic entrenchment orderings in terms of the following set of postulates:

(GEE1) $\top \not\sqsubseteq \top$

(GEE2 \uparrow) If $\alpha \sqsubseteq \beta$ and $\beta \models \gamma$ then $\alpha \sqsubseteq \gamma$

(GEE2 \downarrow) If $\alpha \sqsubseteq \beta$ and $\gamma \models \alpha$ then $\gamma \sqsubseteq \beta$

(GEE3 \uparrow) If $\alpha \sqsubseteq \beta$ and $\alpha \sqsubseteq \gamma$ then $\alpha \sqsubseteq \beta \wedge \gamma$

(GEE3 \downarrow) If $\alpha \wedge \beta \sqsubseteq \beta$ then $\alpha \sqsubseteq \beta$

Definition 5.4.4 A binary relation \sqsubseteq_{GEE} on L is a *GEE-ordering* (with respect to a belief set K) iff it satisfies (GEE1) to (GEE3 \downarrow). \square

⁶A relation S is the converse complement of a binary relation R on a set X iff for every $x, y \in X$, $(x, y) \in S$ iff $(y, x) \notin R$.

Rott is of the opinion that the EE-orderings should be seen as converse complements of the GEE-orderings. Since the EE-orderings are total preorders, taking the strict version of an EE-ordering is the same as taking its converse complement. He shows that the strict versions of the EE-orderings form a strict subset of the set of all GEE-orderings. The GEE-orderings are not subject to analogues of the minimality and maximality conditions imposed on the EE-orderings. The following four supplementary postulates for generalised epistemic entrenchment are intended to serve as such analogues.

(GEE4) If $K \neq L$ then $\perp \sqsubset \alpha$ iff $\alpha \in K$

(GEE4') If $\alpha \notin K$ and $\beta \in K$ then $\alpha \sqsubset \beta$

(GEE5) If $\not\models \alpha$ then $\alpha \sqsubset \top$

(GEE5') If $\alpha \sqsubset \top$ and $\beta \not\sqsubset \top$ then $\alpha \sqsubset \beta$

It is easily verified that the strict versions of the EE-orderings satisfy these four postulates as well.

It turns out that the GEE-orderings can be defined in terms of families of strict versions of the EE-orderings. With a small modification, the following results are obtained from [Rott, 1992c].

Theorem 5.4.5 *1. The intersection of every family of strict EE-orderings is a GEE-ordering that satisfies the four supplementary postulates as well.*

2. For every GEE-ordering \sqsubset_{GEE} that satisfies the four supplementary postulates as well, there is a family \mathcal{E} of strict EE-orderings such that $\sqsubset_{GEE} = \cap \mathcal{E}$.

Theorem 5.4.5 is remarkably similar to theorem 5.4.3, the representation theorem for the LR-orderings in terms of families of EE-orderings, and might lead one to suspect that the GEE-orderings (satisfying the four supplementary postulates) are precisely the strict versions of the LR-orderings. But as the next example shows, this is not the case.

Example 5.4.6 Consider the propositional language L generated by the two atoms p and q with the valuation semantics (V, \models) , where $V = \{11, 10, 01, 00\}$. Let $K =$

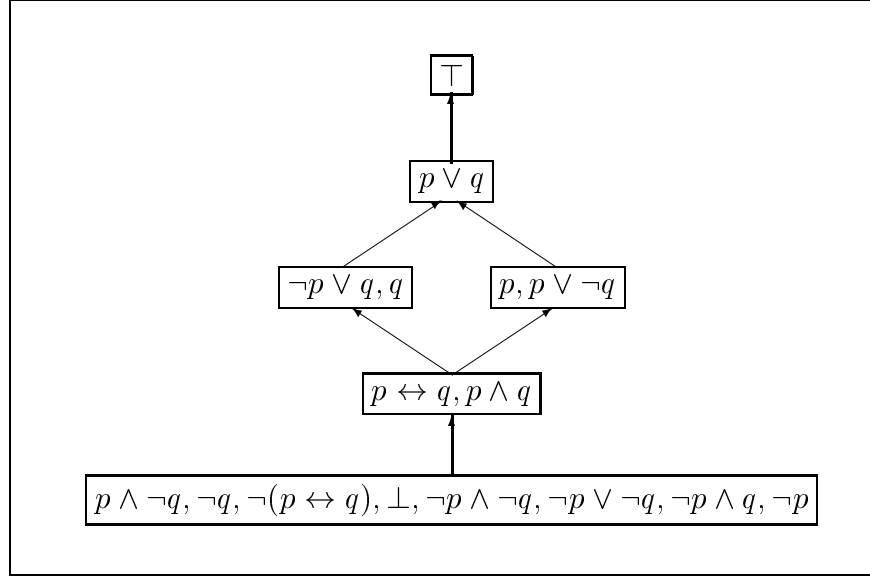


Figure 5.2: A graphical representation of the LR-ordering used in example 5.4.6. The ordering is obtained from the reflexive transitive closure of the relation determined by the arrows. Every wff in the figure is a canonical representative of the set of wffs that are logically equivalent to it.

$Cn(p \wedge q)$, and consider the LR-ordering \sqsubseteq_{LR} defined as follows:

$$\alpha \sqsubseteq_{LR} \beta \text{ iff } \begin{cases} \beta \in L \text{ if } \alpha \notin K, \\ p \wedge q \models \beta \text{ if } \alpha \equiv p \wedge q \text{ or } \alpha \equiv p \leftrightarrow q, \\ q \models \beta \text{ if } \alpha \equiv q \text{ or } \alpha \equiv \neg p \vee q, \\ p \models \beta \text{ if } \alpha \equiv p \text{ or } \alpha \equiv p \vee \neg q, \\ p \vee q \models \beta \text{ if } \alpha \equiv p \vee q, \\ \alpha \in L \text{ if } \models \beta. \end{cases}$$

Figure 5.2 contains a graphical representation of the LR-ordering \sqsubseteq_{LR} . An inspection of figure 5.2 shows that \sqsubseteq_{LR} is indeed an LR-ordering, but that the strict version \sqsubset_{LR} of \sqsubseteq_{LR} violates (GEE3 \uparrow), by taking α as $p \leftrightarrow q$, β as p , and γ as q , and violates (GEE3 \downarrow) by taking α as p and β as q . \square

5.5 Refined entrenchment

Although the faithful total preorders are sufficient for a complete characterisation of AGM theory change, it is possible to achieve the same effect with other preorders as well. We have a particular interest in a set of faithful preorders that are very closely related to the faithful total preorders.

Definition 5.5.1 A weak partial order \leq on a set X is called *modular* iff for every $x, y, z \in X$, if $x \parallel_{\leq} y$ and $z \leq x$, then $z \leq y$. \square

The modular weak partial orders are the reflexive versions of the modular partial orders of Ginsberg [1986] and Lehmann and Magidor [1992], which in turn, can also be described as the relations on a set X satisfying transitivity and virtual connectivity (see definition 2.4.4). Intuitively, a modular weak partial order ensures that the elements of X are arranged in levels, with incomparable elements being regarded as on the same level. Using this intuition, it is clear that the following two identities provide a natural connection between the total preorders and the modular weak partial orders.

(Def \leq from \preceq) $\leq = \preceq \setminus \{(x, y) \in X \times X \mid x \neq y \text{ and } x \equiv_{\preceq} y\}$

(Def \preceq from \leq) $\preceq = \leq \cup \{(x, y) \in X \times X \mid x \parallel_{\leq} y\}$

Definition 5.5.2 A faithful modular weak partial order and a faithful total preorder are *semantically related* iff they can be defined in terms of each other using (Def \leq from \preceq) and (Def \preceq from \leq) respectively. \square

It is easily seen that a faithful total preorder and its semantically related modular weak partial order are minimal-equivalent (see definition 3.3.6), and as a result, the set of faithful modular weak partial orders can also be used to characterise AGM theory change.⁷ So if we are only interested in minimality, as in the case of AGM theory change, a move from the faithful total preorders to the faithful modular weak partial orders is an inessential technical modification. But as we shall see below, other constructions involving orderings on interpretations are more sensitive to such a shift. (See also chapter 6.) From an information-theoretic point of view, there is also an important difference. In a faithful total preorder, infatoms on the same level are regarded as equally important or entrenched, while the semantically related faithful modular weak

⁷These results are special cases of proposition 5.7.3 and corollary 5.7.4.

partial order will regard them as incomparable. As we shall see in chapter 6, this can have important effects on basic principles such as the principle of Indifference and the principle of Preference.

In this section we use the faithful modular weak partial orders as the basis for the presentation and investigation of sets of refined versions of the EE-orderings that allow for the possibility of wffs being incomparable. These orderings are obtained by applying (Def \sqsubseteq_E from \preceq), not to the faithful total preorders, but to the faithful modular weak partial orders.

Definition 5.5.3 An *RE-ordering* \sqsubseteq_{RE} (refined entrenchment ordering) is a binary relation on L defined in terms of a faithful modular weak partial order using (Def \sqsubseteq_E from \preceq). We say that an EE-ordering and an RE-ordering, defined respectively in terms of a faithful total preorder and its semantically related faithful modular weak partial order, using (Def \sqsubseteq_E from \preceq), are *semantically related*. \square

The next proposition provides a preliminary list of properties of the RE-orderings.

Proposition 5.5.4 Let \sqsubseteq_{RE} be the RE-ordering defined in terms of the faithful modular weak partial order \leq using (Def \sqsubseteq_E from \preceq). Then \sqsubseteq_{RE} satisfies the following properties.

1. \sqsubseteq_{RE} is a preorder (that need not be total).
2. Suppose that the EE-ordering \sqsubseteq_{EE} is semantically related to \sqsubseteq_{RE} . If $\alpha \sqsubseteq_{RE} \beta$ then $\alpha \sqsubseteq_{EE} \beta$.
3. If $\alpha \models \beta$ then $\alpha \sqsubseteq_{RE} \beta$.
4. $\alpha \sqsubseteq_{RE} \beta$ for all α , iff $\models \beta$.
5. If $\alpha \equiv \beta$ then $\alpha \sqsubseteq_{RE} \gamma$ iff $\beta \sqsubseteq_{RE} \gamma$, and $\gamma \sqsubseteq_{RE} \alpha$ iff $\gamma \sqsubseteq_{RE} \beta$.
6. If K is satisfiable then $\{\alpha \mid \neg\alpha \in K\} = [\perp]_{\sqsubseteq_{RE}}$.
7. If $\alpha \notin K$ and $\beta \in K$ then $\alpha \sqsubset_{RE} \beta$.
8. If $\neg\beta \in K$ and $\neg\gamma \notin K$ then $\beta \sqsubset_{RE} \gamma$.
9. If $\alpha \notin K$ then $K \cup \{\alpha\} \models \beta$ iff $\alpha \sqsubseteq_{RE} \beta$.

10. If $\alpha \equiv_{\sqsubseteq_{RE}} \beta$ then $\alpha \wedge \beta, \alpha \vee \beta \in [\alpha]_{\sqsubseteq_{RE}} = [\beta]_{\sqsubseteq_{RE}}$.

11. $\alpha \sqsubseteq_{RE} \alpha \wedge \beta$, or $\beta \sqsubseteq_{RE} \alpha \wedge \beta$, or both $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ and $\beta \rightarrow \alpha \not\sqsubseteq_{RE} \beta$.

Proof Many of these results follow from proposition 5.7.6. We only prove the remaining parts. For part (6), we need to show that if K is satisfiable, then $[\perp]_{\sqsubseteq_{RE}} \subseteq \{\alpha \mid \neg\alpha \in K\}$. So suppose that K is satisfiable, pick any $\alpha \in [\perp]_{\sqsubseteq_{RE}}$ and assume that $\neg\alpha \notin K$. Then there is at least one model x of K that satisfies α , and thus $\alpha \not\sqsubseteq_{RE} \perp$, contradicting the supposition that $\alpha \in [\perp]_{\sqsubseteq_{RE}}$. For part (8), we need to show that if $\neg\beta \in K$ and $\neg\gamma \notin K$ then $\gamma \not\sqsubseteq_{RE} \beta$. So suppose that $\neg\beta \in K$ and $\neg\gamma \notin K$. Since $M(K) \cap M(\gamma) \neq \emptyset$, it follows from faithfulness that there is a $y \in M(\gamma) \cap M(K) \subseteq M(\neg\beta)$, such that $x \in M(\gamma)$ for every $x \leq y$, i.e. $\gamma \not\sqsubseteq_{RE} \beta$. For part (9), we need to show that if $\alpha \notin K$ and $\alpha \sqsubseteq_{RE} \beta$ then $K \cup \{\alpha\} \models \beta$. So let $\alpha \notin K$ and suppose that $K \cup \{\alpha\} \not\models \beta$. So there is a $y \in M(K) \cap M(\alpha)$ such that $y \in M(\neg\beta)$. That is, $y \in M(\neg\beta)$ and for every $x \leq y$, $x \in M(\alpha)$, which means that $\alpha \not\sqsubseteq_{RE} \beta$. For part (10), we need to show that if $\alpha \equiv_{\sqsubseteq_{RE}} \beta$ then $\alpha \vee \beta \in [\alpha]_{\sqsubseteq_{RE}} = [\beta]_{\sqsubseteq_{RE}}$. From part (3) of this proposition it follows that $\alpha \sqsubseteq_{RE} \alpha \vee \beta$. To show that $\alpha \vee \beta \sqsubseteq_{RE} \alpha$, assume that it is not the case. Then there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\alpha \vee \beta)$ for every $x \leq y$. Therefore $y \in M(\neg\alpha) \cap M(\beta)$. But since $\beta \sqsubseteq_{RE} \alpha$, there is a $z \in \text{Min}_{\leq}(\neg\beta)$ such that $z < y$ which, together with the minimality of y in $M(\neg\alpha)$, contradicts the fact that $\alpha \sqsubseteq_{RE} \beta$. \square

An inspection of the properties set out in proposition 5.5.4 reveals something of the structure of the RE-orderings. They are refined versions of the EE-orderings that need not be total. Furthermore, every RE-ordering partitions the set of wffs into four disjoint sets. The logically valid wffs are all equally entrenched and strictly more entrenched than all other wffs. Next comes the remaining wffs in K . While strictly more entrenched than the wffs not in K , they need not all be comparable. The third partition consists of the K -undecided wffs, which are all strictly less entrenched than the wffs in K and more entrenched than the K -refuted wffs. (If K is unsatisfiable, there are not any K -undecided wffs or K -refuted wffs.) So the RE-orderings are able to distinguish between wffs not in K . In fact, the part of an RE-ordering restricted to the wffs that are not in K , corresponds to classical entailment relative to K . This certainly has more intuitive appeal than regarding all the wffs that are not in K as equally entrenched, such as the EE-orderings do. For example, it makes much more

sense to regard a wff that is K -refuted as less entrenched than a wff that is merely K -undecided, than to regard them both as equally entrenched.

The last two parts of proposition 5.5.4 are worth singling out. Note that part (10) does not hold for the EE-orderings. An interesting example is the case of a wff α and its negation. In an EE-ordering \sqsubseteq_{EE} , it is perfectly acceptable to have $\neg\alpha \equiv_{\sqsubseteq_{EE}} \alpha$ as long as α is not logically valid or logically invalid. However, if this were the case in an RE-ordering \sqsubseteq_{RE} , part (10) of proposition 5.5.4 would require that $\alpha \vee \neg\alpha \in [\alpha]_{\sqsubseteq_{RE}}$, thus contradicting part (4) of the same proposition. Part (11) bears a vague resemblance to the postulate (EE3), and will be used in our characterisation of the RE-orderings in terms of postulates. In fact, so will the properties contained in the lemma below.

Lemma 5.5.5 *Let \sqsubseteq_{RE} be an RE-ordering.*

1. *If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\alpha \rightarrow \beta \sqsubseteq_{RE} \alpha$ or $\beta \rightarrow \gamma \sqsubseteq_{RE} \beta$.*
2. *If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\alpha \not\sqsubseteq_{RE} \beta$ or $\beta \rightarrow \gamma \sqsubseteq_{RE} \beta$.*
3. *If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\beta \not\sqsubseteq_{RE} \gamma$ or $\alpha \rightarrow \beta \sqsubseteq_{RE} \alpha$.*

Proof Let \leq be a faithful modular weak partial order from which \sqsubseteq_{RE} is defined using (Def \sqsubseteq_E from \preceq).

1. Suppose that $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ and $\beta \rightarrow \gamma \not\sqsubseteq_{RE} \beta$. By $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such $x \in M(\alpha \rightarrow \beta)$ for every $x \leq y$. And by the minimality of y in $M(\neg\alpha)$, $x \in M(\alpha) \cap M(\beta)$ for every $x < y$. Similarly, $\beta \rightarrow \gamma \not\sqsubseteq_{RE} \beta$ implies that there is a $v \in \text{Min}_{\leq}(\neg\beta)$ such that $u \in M(\beta) \cap M(\gamma)$ for every $u < v$. Since \leq is a modular weak partial order, it has to be the case that $v \not\leq y$. And this means that $z \in M(\alpha) \cap M(\gamma)$ for every $z < y$. So $y \in M(\neg\alpha)$ and $x \in M(\alpha \rightarrow \gamma)$ for every $x \leq y$. That is, $\alpha \rightarrow \gamma \not\sqsubseteq_{RE} \alpha$.
2. Suppose that $\alpha \sqsubseteq_{RE} \beta$ and $\beta \rightarrow \gamma \not\sqsubseteq_{RE} \beta$. As in part (1), $\beta \rightarrow \gamma \not\sqsubseteq_{RE} \beta$ means there is a $v \in \text{Min}_{\leq}(\neg\beta)$ such that $u \in M(\beta) \cap M(\gamma)$ for every $u < v$. So by $\alpha \sqsubseteq_{RE} \beta$ there is a $w \leq v$ such that $w \in \text{Min}_{\leq}(\neg\alpha)$. And since $w \leq v$, it follows that $u \in M(\alpha) \cap M(\gamma)$ for every $u < w$. So $u \in M(\alpha \rightarrow \gamma)$ for every $u \leq w$. That is, $\alpha \rightarrow \gamma \not\sqsubseteq_{RE} \alpha$.
3. Suppose that $\beta \sqsubseteq_{RE} \gamma$ and $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$. As in part (1), $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ means there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\alpha) \cap M(\beta)$ for every $x < y$. So by

$\beta \sqsubseteq_{RE} \gamma$, $z \not\leq y$ for every $z \in M(\neg\gamma)$. And therefore $x \in M(\gamma)$ for every $x < y$. So $y \in M(\neg\alpha)$ and $x \in M(\alpha \rightarrow \gamma)$ for every $x \leq y$. That is, $\alpha \rightarrow \gamma \not\sqsubseteq_{RE} \alpha$.

□

5.5.1 Refined entrenchment and the EE-orderings

The application of lemma 5.2.1 to the faithful modular weak partial orders and the RE-orderings obtained in terms of them, using (Def \sqsubseteq_E from \preceq), yields a useful result. It shows that two wffs are equally entrenched if and only if their negations have the same minimal models, and for any two wffs α and β , both of whom are not logically valid, α is strictly more entrenched than β if and only if the minimal models of $\neg\beta$ are either strictly above the minimal models of $\neg\alpha$, or form a strict subset of the minimal models of $\neg\alpha$. As a consequence, two wffs α and β are incomparable iff the minimal models of the negations of the two wffs are on the same level, the minimal models of $\neg\alpha$ include a model of β , and the minimal models of $\neg\beta$ include a model of α .

Corollary 5.5.6 *Let \leq be a faithful modular weak partial order, and let \sqsubseteq_{RE} be the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq).*

1. $\alpha \equiv_{\sqsubseteq_{RE}} \beta$ iff $Min_{\leq}(\neg\beta) = Min_{\leq}(\neg\alpha)$.
2. If $\not\equiv \alpha$ and $\not\equiv \beta$ then $\alpha \sqsubset_{RE} \beta$ iff $Min_{\leq}(\neg\beta) \subset Min_{\leq}(\neg\alpha)$ or $Min_{\leq}(\neg\alpha) < Min_{\leq}(\neg\beta)$.
3. $\alpha \parallel_{\sqsubseteq_{RE}} \beta$ iff $Min_{\leq}(\neg\alpha) \not\subset M(\neg\beta)$, $Min_{\leq}(\neg\beta) \not\subset M(\neg\alpha)$, and $x \parallel_{\leq} y$ or $x = y$ for every $y \in Min_{\leq}(\neg\beta)$ and every $x \in Min_{\leq}(\neg\alpha)$.

Proof 1. Follows easily from lemma 5.2.1.

2. Suppose that $\not\equiv \alpha$, $\not\equiv \beta$, and $\alpha \sqsubset_{RE} \beta$, and suppose there is a $y \in Min_{\leq}(\neg\beta)$ and an $x \in Min_{\leq}(\neg\alpha)$, such that $x \not\leq y$. From $\beta \not\sqsubseteq_{RE} \alpha$ there is a $v \in Min_{\leq}(\neg\alpha)$ such that $u \in M(\beta)$ for every $u \leq v$. So, for every $s \in Min_{\leq}(\neg\beta)$ and every $t \in Min_{\leq}(\neg\alpha)$, $s \not\leq t$. Therefore the minimal models of $\neg\alpha$ and $\neg\beta$ lie on the same level. Now pick any $u \in Min_{\leq}(\neg\beta)$. By $\alpha \sqsubseteq_{RE} \beta$, $u \in Min_{\leq}(\neg\alpha)$. Furthermore v is a minimal model of $\neg\alpha$ that is not a minimal model of $\neg\beta$ and so $Min_{\leq}(\neg\beta) \subset Min_{\leq}(\neg\alpha)$. The converse follows easily, and is omitted.

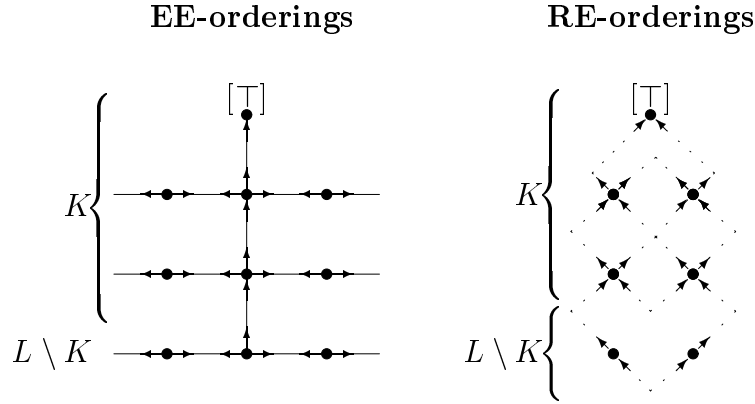


Figure 5.3: A graphical comparison of the EE-orderings and the RE-orderings

3. Suppose $\alpha \parallel_{\sqsubseteq_{RE}} \beta$. So there is a $v \in \text{Min}_{\leq}(\neg\alpha)$ such that $u \in M(\beta)$ for every $u \leq v$, and there is a $y \in \text{Min}_{\leq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \leq y$. The required result then follows from the fact that $v \parallel_{\leq} y$ and that \leq is a modular weak partial order. The converse follows easily. \square

A consequence of corollaries 5.2.2 and 5.5.6 is that the RE-ordering which is semantically related to an EE-ordering \sqsubseteq_{EE} maintains the ordering between the equivalence classes of wffs modulo \sqsubseteq_{EE} but offers an exploded view of each of these equivalence classes. Figure 5.3 gives a graphical representation of this situation. From an information-theoretic point of view, the results of corollary 5.5.6 are quite interesting. They show that the RE-orderings have more of the underlying entailment relation \models built into them than their semantically related EE-orderings. Thus, two wffs are equally entrenched when they have *exactly* the same set of least entrenched contents bits, not when their least entrenched content bits are merely on the same level, as is the case for the EE-orderings. Similarly, a wff β will be more entrenched than a wff α , not only when the least entrenched content bits of β are more entrenched than the least entrenched contents bits of α , but also when the least entrenched content bits of α strictly includes the least entrenched content bits of β . And continuing in the same vein, the incomparability of two wffs α and β , in terms of refined entrenchment, then occurs when their least entrenched content bits are on the same level, but neither set is included in the other.

The strong link with the entailment relation \models is evident when we consider the faithful modular weak partial order \leq in which the countermodels of K are all on the same level. In such a case, the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq) corresponds exactly to entailment relative to K . In fact, an even stronger link exists in the limiting case where K just contains the logically valid wffs. In this case, there is only one RE-ordering and one EE-ordering with respect to K . And whereas the RE-ordering is exactly the entailment relation \models , the semantically related EE-ordering regards all wffs, except for the logically valid ones, as equally entrenched. An elegant explanation for this difference can be found by looking at \leq and the semantically related faithful total preorder \preceq . It is easily verified that \leq is the identity relation on U , while \preceq is the Cartesian product $U \times U$. So \leq represents the epistemic state of an agent for whom all infatoms are incomparable. In the absence of any preference for certain bits of information, it has no choice but to revert back to the logical content of wffs as a measure of the entrenchment. Hence the use of the classical entailment relation \models as the associated entrenchment ordering. On the other hand, the faithful total preorder \preceq represents the epistemic state of an agent who regards all infatoms as *equally* entrenched. Hence all wffs, except the logically valid wffs, are seen as equally entrenched.

In light of the similarity between the methods of constructing the RE-orderings and the EE-orderings, it is natural to wonder whether they can be defined in terms of one another. The next theorem shows that this can be accomplished by the following two identities:

(Def \sqsubseteq_{RE} from \sqsubseteq_{EE}) $\alpha \sqsubseteq_{RE} \beta$ iff $\models \beta$ or $\alpha \sqsubseteq_{EE} \beta$ or $\beta \sqsubseteq_{EE} \alpha \rightarrow \beta$

(Def \sqsubseteq_{EE} from \sqsubseteq_{RE}) $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \sqsubseteq_{RE} \beta$ or $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$

Theorem 5.5.7 *Let the RE-ordering \sqsubseteq_{RE} and the EE-ordering \sqsubseteq_{EE} be semantically related.*

1. \sqsubseteq_{RE} can also be defined in terms of \sqsubseteq_{EE} using (Def \sqsubseteq_{RE} from \sqsubseteq_{EE})

2. \sqsubseteq_{EE} can also be defined in terms of \sqsubseteq_{RE} using (Def \sqsubseteq_{EE} from \sqsubseteq_{RE})

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} is defined using (Def \sqsubseteq_E from \preceq), and let \preceq be the semantically related faithful total preorder in terms of which \sqsubseteq_{EE} is defined using (Def \sqsubseteq_E from \preceq).

1. Suppose that $\alpha \sqsubseteq_{RE} \beta$. By part (2) of proposition 5.5.4, $\alpha \sqsubseteq_{EE} \beta$. Suppose further that $\alpha \not\sqsubseteq_{EE} \beta$, i.e. $\beta \sqsubseteq_{EE} \alpha$, and that $\not\models \beta$. Then $\not\models \alpha$ (by part (4) of proposition 5.5.4), and by corollary 5.2.2 it follows that for every $y \in \text{Min}_{\leq}(\neg\beta)$ and every $x \in \text{Min}_{\leq}(\neg\alpha)$, $x \equiv_{\leq} y$. Because $\not\models \beta$ there is thus a $v \in \text{Min}_{\leq}(\neg\beta)$ such that for every $u \prec v$, $u \in M(\alpha) \cap M(\beta)$. Combined with $\alpha \sqsubseteq_{RE} \beta$ this means that for every $w \equiv_{\leq} v$, $w \in M(\neg\beta) \cap M(\neg\alpha)$ or $w \in M(\beta)$. Therefore $z \in M(\alpha \rightarrow \beta)$ for every $z \preceq v$, and so $\alpha \rightarrow \beta \not\sqsubseteq_{EE} \beta$, i.e. $\beta \sqsubseteq_{EE} \alpha \rightarrow \beta$. Conversely, if $\models \beta$ then $\alpha \sqsubseteq_{RE} \beta$ follows vacuously. If $\alpha \sqsubseteq_{EE} \beta$, i.e. $\beta \not\sqsubseteq_{EE} \alpha$, there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \preceq y$. So $y < u$ for every $u \in \text{Min}_{\leq}(\neg\beta)$ and therefore $\alpha \sqsubseteq_{RE} \beta$. Finally, suppose that $\beta \sqsubseteq_{EE} \alpha \rightarrow \beta$, i.e. $\alpha \rightarrow \beta \not\sqsubseteq_{EE} \beta$. Then there is a $y \in \text{Min}_{\leq}(\neg\beta)$ such that $x \in M(\alpha \rightarrow \beta)$ for every $x \preceq y$, and so $\text{Min}_{\leq}(\neg\beta) = \text{Min}_{\leq}(\neg\beta) \subseteq M(\neg\alpha)$. So for every $v \in M(\neg\beta)$, there is a $u \in M(\neg\alpha)$ such that $u \leq v$, i.e. $\alpha \sqsubseteq_{RE} \beta$.
2. Suppose that $\alpha \sqsubseteq_{EE} \beta$ and that $\alpha \not\sqsubseteq_{RE} \beta$. By $\alpha \not\sqsubseteq_{RE} \beta$ there is a $y \in \text{Min}_{\leq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \leq y$. So $z \in M(\alpha) \cap M(\beta)$ for every $z < y$, and from $\alpha \sqsubseteq_{EE} \beta$ it thus follows that there is a $v \equiv_{\leq} y$ such that $v \in M(\neg\alpha)$. So $u \in M(\alpha \rightarrow \beta)$ for every $u \leq v$ and thus $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$. Conversely, if $\alpha \sqsubseteq_{RE} \beta$ then $\alpha \sqsubseteq_{EE} \beta$ by part (2) of proposition 5.5.4. And if $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ then there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\alpha \rightarrow \beta)$ for every $x \leq y$. Therefore $z \in M(\alpha) \cap M(\beta)$ for every $z < y$. So $u \not\leq y$ for every $u \in M(\neg\beta)$, from which $\alpha \sqsubseteq_{EE} \beta$ follows easily.

□

Theorem 5.5.7 also shows that the identities (Def \sqsubseteq_{EE} from \sqsubseteq_{RE}) and (Def \sqsubseteq_{RE} from \sqsubseteq_{EE}) are interchangeable. That is, if we start with either an RE-ordering or an EE-ordering, and then apply (Def \sqsubseteq_{EE} from \sqsubseteq_{RE}) and (Def \sqsubseteq_{RE} from \sqsubseteq_{EE}) in the appropriate order, we end up with the same ordering that we started with. It is thus appropriate to think of refined entrenchment as an alternative to the EE-orderings. Indeed, in view of theorem 5.5.7, there is a one-to-one correspondence between the RE-orderings and the EE-orderings, obtained by applying the two identities (Def \sqsubseteq_{EE} from \sqsubseteq_{RE}) and (Def \sqsubseteq_{RE} from \sqsubseteq_{EE}).

The close relationship between the RE-orderings and the EE-orderings raises the question of whether the two notions ever coincide. One part of the answer to this

question is easy. Whenever a faithful preorder \preceq is both a total preorder and a modular weak partial order, the EE-ordering and the RE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq) are, by definition, identical. Now, it is easy to see that this is the case only when \preceq , restricted to the countermodels of K , is a linear order.

Definition 5.5.8 For a belief set K , a K -linear order \preceq is a faithful total preorder such that $\preceq \cap ((U \setminus M(K)) \times (U \setminus M(K)))$ is a linear order. \square

Proposition 5.5.9 Let \leq be any K -linear order. The binary relation defined in terms of \leq using (Def \sqsubseteq_E from \preceq) is an EE-ordering and an RE-ordering.

Proof Follows immediately from the fact that \leq is both a faithful total preorder and a faithful modular weak partial order. \square

In general, there may be instances of faithful total preorders, or faithful modular weak partial orders, as the case may be, that are not K -linear orders, but that nevertheless define the same EE-orderings (or RE-orderings) as some K -linear order. More interesting, no doubt, is that, at least in the finitely generated propositional case, if an EE-ordering cannot be defined in terms of a K -linear order using (Def \sqsubseteq_E from \preceq), then it is not an RE-ordering, and vice versa.

Proposition 5.5.10 Let L be a finitely generated propositional language with a valuation semantics (V, \models) .

1. Let \sqsubseteq_{RE} be an RE-ordering that cannot be defined in terms of a K -linear order using (Def \sqsubseteq_E from \preceq). Then \sqsubseteq_{RE} is not an EE-ordering.
2. Let \sqsubseteq_{EE} be an EE-ordering that cannot be defined in terms of a K -linear order using (Def \sqsubseteq_E from \preceq). Then \sqsubseteq_{EE} is not an RE-ordering.

Proof 1. By definition, \sqsubseteq_{RE} can be defined in terms of a faithful modular weak partial order \leq , that is not a K -linear order, using (Def \sqsubseteq_E from \preceq). That means there are at least two distinct countermodels x and y of K such that $x \parallel_{\leq} y$. Let α_x be a wff that axiomatises x and let α_y be a wff that axiomatises y . (By our choice of L , there are such wffs.) So $\text{Min}_{\leq}(\alpha_x) = \{x\}$ and $\text{Min}_{\leq}(\alpha_y) = \{y\}$, and thus $\text{Min}_{\leq}(\alpha_x) \not\subseteq M(\alpha_y)$, $\text{Min}_{\leq}(\alpha_y) \not\subseteq M(\alpha_x)$, and $\text{Min}_{\leq}(\alpha_x) \parallel_{\leq} \text{Min}_{\leq}(\alpha_y)$. By part (iii) of corollary 5.5.6 it then follows that $\neg\alpha_x \parallel_{\sqsubseteq_{RE}} \neg\alpha_y$, and so \sqsubseteq_{RE} cannot be an EE-ordering.

2. Assume that \sqsubseteq_{EE} is an RE-ordering. Then it follows, as in part (i), that \sqsubseteq_{EE} cannot be an EE-ordering; a contradiction.

□

5.5.2 Postulates for refined entrenchment

In this section we present a description of the RE-orderings in terms of postulates, and give a representation theorem to prove that the postulates do indeed provide a characterisation of the RE-orderings. The postulates are given below.

(RE1) \sqsubseteq_{RE} is transitive

(RE2) If $\alpha \models \beta$ then $\alpha \sqsubseteq_{RE} \beta$

(RE3a) If $\alpha, \beta \in K$ then $\alpha \sqsubseteq_{RE} \alpha \wedge \beta$, or $\beta \sqsubseteq_{RE} \alpha \wedge \beta$,

or both $\alpha \rightarrow \beta \not\sqsubseteq_{RE} \alpha$ and $\beta \rightarrow \alpha \not\sqsubseteq_{RE} \beta$

(RE3b) If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\alpha \rightarrow \beta \sqsubseteq_{RE} \alpha$ or $\beta \rightarrow \gamma \sqsubseteq_{RE} \beta$

(RE3c) If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\alpha \not\sqsubseteq_{RE} \beta$ or $\beta \rightarrow \gamma \sqsubseteq_{RE} \beta$

(RE3d) If $\alpha \rightarrow \gamma \sqsubseteq_{RE} \alpha$ then $\beta \not\sqsubseteq_{RE} \gamma$ or $\alpha \rightarrow \beta \sqsubseteq_{RE} \alpha$

(RE4a) If $\alpha \notin K$ and $\beta \in K$, then $\alpha \sqsubset_{RE} \beta$

(RE4b) If $\alpha, \beta \notin K$, then $\alpha \sqsubseteq_{RE} \beta$ iff $K \cup \{\alpha\} \models \beta$

(RE5) If $\alpha \sqsubseteq_{RE} \beta$ for all α , then $\models \beta$

To a certain extent, the postulates for refined entrenchment follow the same pattern as the postulates for the EE-orderings, and this is reflected in the labelling scheme we use. (RE1), (RE2) and (RE5) are identical to (EE1), (EE2), and (EE5) respectively. And while (RE3a) bears a vague resemblance to (EE3), it is a bit more difficult to describe the intuition associated with (RE3b), (RE3c) and (RE3d). Technically though, they seem to be necessary for a complete description of the relationship between the wffs in K . (EE4) gives a complete description of how an EE-ordering treats the wffs that are not in K , while the handling of such wffs by the RE-orderings are described by the two independent postulates, (RE4a) and (RE4b). (RE4a) describes the relationship between wffs in K and wffs not in K , while (RE4b) is a prescription for the treatment of

any two wffs, neither of which are in K . To obtain the desired representation theorem, we need the following two lemmas.

Lemma 5.5.11 *If \sqsubseteq is a relation on L that satisfies (RE1) to (RE5) then the relation \sqsubseteq_{EE} defined as: $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \sqsubseteq \beta$ or $\alpha \rightarrow \beta \not\sqsubseteq \alpha$, satisfies (EE1) to (EE5).*

Proof For (EE1), suppose that $\alpha \sqsubseteq_{EE} \beta$ and $\beta \sqsubseteq_{EE} \gamma$. That is, $\alpha \sqsubseteq \beta$ or $\alpha \rightarrow \beta \not\sqsubseteq \alpha$, and $\beta \sqsubseteq \gamma$ or $\beta \rightarrow \gamma \not\sqsubseteq \beta$. This can be divided into four cases: (i) $\alpha \sqsubseteq \beta$ and $\beta \sqsubseteq \gamma$, (ii) $\alpha \sqsubseteq \beta$ and $\beta \rightarrow \gamma \not\sqsubseteq \beta$, (iii) $\beta \sqsubseteq \gamma$ and $\alpha \rightarrow \beta \not\sqsubseteq \alpha$, and (iv) $\alpha \rightarrow \beta \not\sqsubseteq \alpha$ and $\beta \rightarrow \gamma \not\sqsubseteq \beta$. For (i), $\alpha \sqsubseteq \gamma$ follows from (RE1). For (ii), (iii), and (iv), $\alpha \rightarrow \gamma \not\sqsubseteq \alpha$ follows from (RE3c), (RE3d), and (RE3b) respectively. So in all four cases, either $\alpha \sqsubseteq \gamma$ or $\alpha \rightarrow \gamma \not\sqsubseteq \alpha$. That is, $\alpha \sqsubseteq_{EE} \gamma$. (EE2) follows from (RE2) and (EE3) follows from (RE3a). For (EE4), suppose that $K \neq L$, and let $\alpha \notin K$. Assume there is a β such that $\alpha \not\sqsubseteq_{EE} \beta$. That is, $\alpha \not\sqsubseteq \beta$ and $\alpha \rightarrow \beta \sqsubseteq \alpha$. By (RE4a) $\beta \notin K$, and so, by (RE4b), $K \cup \{\alpha \rightarrow \beta\} \models \alpha$. But this means $\alpha \in K$; a contradiction. Conversely, suppose that $\alpha \in K$. So $\neg\alpha \notin K$, and $\neg\alpha \sqsubset \alpha$ by (RE4a). And since $\alpha \rightarrow \neg\alpha \equiv \neg\alpha$, we have that $\alpha \not\sqsubseteq \neg\alpha$ and $\alpha \rightarrow \neg\alpha \sqsubseteq \alpha$. That is, $\alpha \not\sqsubseteq_{EE} \neg\alpha$. For (EE5), suppose that $\not\models \beta$. By (RE5), $\top \not\sqsubseteq \beta$ and by (RE2), $\top \rightarrow \beta \sqsubseteq \top$. That is, $\top \not\sqsubseteq_{EE} \beta$. \square

Lemma 5.5.12 *Let \sqsubseteq be a relation on L that satisfies (RE1) to (RE5). If \sqsubseteq_{EE} is defined as: $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \sqsubseteq \beta$ or $\alpha \rightarrow \beta \not\sqsubseteq \alpha$, and \sqsubseteq_{RE} is defined as: $\alpha \sqsubseteq_{RE} \beta$ iff $\models \beta$ or $\alpha \sqsubset_{EE} \beta$ or $\beta \sqsubset_{EE} \alpha \rightarrow \beta$, then $\sqsubseteq = \sqsubseteq_{RE}$.*

Proof By lemma 5.5.11, \sqsubseteq_{EE} is an EE-ordering, and thus a total preorder. By keeping in mind that $\alpha \sqsubset_{EE} \beta$ iff $\beta \not\sqsubseteq_{EE} \alpha$, noting that $(\alpha \rightarrow \beta) \rightarrow \beta \equiv \alpha \vee \beta$, and combining the definitions of \sqsubseteq_{EE} and \sqsubseteq_{RE} , it suffices to show that

$$\alpha \sqsubseteq \beta \text{ iff } \begin{cases} \models \beta, \text{ or} \\ \beta \not\sqsubseteq \alpha \text{ and } \beta \rightarrow \alpha \sqsubseteq \beta, \text{ or} \\ \alpha \rightarrow \beta \not\sqsubseteq \beta \text{ and } \alpha \vee \beta \sqsubseteq \alpha \rightarrow \beta. \end{cases}$$

So suppose that $\alpha \sqsubseteq \beta$, $\not\models \beta$, and either $\beta \sqsubseteq \alpha$ or $\beta \rightarrow \alpha \not\sqsubseteq \beta$. We have to show that $\alpha \rightarrow \beta \not\sqsubseteq \beta$ and $\alpha \vee \beta \sqsubseteq \alpha \rightarrow \beta$. Assume that $\alpha \rightarrow \beta \sqsubseteq \beta$. There are two cases. Either $\beta \sqsubseteq \alpha$ or $\beta \rightarrow \alpha \not\sqsubseteq \beta$. In the former case, $\alpha \rightarrow \beta \sqsubseteq \beta \sqsubseteq \alpha$. By (RE3c) it thus follows that $\alpha \not\sqsubseteq \beta$ or $\beta \rightarrow \beta \sqsubseteq \beta$, contradicting $\alpha \sqsubseteq \beta$ and $\not\models \beta$ combined with (RE5). In the latter case, note that $\alpha \sqsubseteq \beta \sqsubseteq \alpha \rightarrow \beta$ by (RE2), and since $(\alpha \rightarrow \beta) \rightarrow \alpha \equiv \alpha$, $(\alpha \rightarrow \beta) \rightarrow \alpha \sqsubseteq \alpha \rightarrow \beta$. By (RE3c) we then have that $\alpha \rightarrow \beta \not\sqsubseteq \beta$, or $\beta \rightarrow \alpha \sqsubseteq \beta$; a

contradiction. So we have shown that $\alpha \rightarrow \beta \not\sqsubseteq \beta$. Now assume that $\alpha \vee \beta \not\sqsubseteq \alpha \rightarrow \beta$. By (RE2), $\alpha \sqsubseteq \beta \sqsubseteq \alpha \rightarrow \beta$. And since $\alpha \equiv (\alpha \rightarrow \beta) \rightarrow \alpha$, $(\alpha \rightarrow \beta) \rightarrow \alpha \sqsubseteq \alpha \rightarrow \beta$. By (RE3b) it then follows that $(\alpha \rightarrow \beta) \rightarrow \beta \sqsubseteq \alpha \rightarrow \beta$, or $\beta \rightarrow \alpha \sqsubseteq \beta$. And since $(\alpha \rightarrow \beta) \rightarrow \beta \equiv \alpha \vee \beta$, it has to be the case that $\beta \rightarrow \alpha \sqsubseteq \beta$. But since we have, by supposition, that $\beta \sqsubseteq \alpha$ or $\beta \rightarrow \alpha \not\sqsubseteq \beta$, this means that $\beta \sqsubseteq \alpha$. From $\beta \rightarrow \alpha \sqsubseteq \beta$ it also follows by (RE3c) that $\beta \not\sqsubseteq \alpha$ or $\alpha \rightarrow \alpha \sqsubseteq \alpha$. So $\alpha \rightarrow \alpha \sqsubseteq \alpha$, and by (RE5), $\models \alpha$. But this contradicts $\not\models \beta$, $\alpha \sqsubseteq \beta$, and (RE5). \square

We are now in a position to prove that the postulates given above provide a characterisation of the RE-orderings.

Theorem 5.5.13 *Every binary relation on L defined in terms of a faithful modular weak partial order using (Def \sqsubseteq_E from \preceq) satisfies (RE1) to (RE5). Conversely, every binary relation on L that satisfies (RE1) to (RE5) can be defined in terms of a faithful modular weak partial order using (Def \sqsubseteq_E from \preceq).*

Proof Let \sqsubseteq_{RE} be a binary relation on L defined in terms of a faithful modular weak partial order using (Def \sqsubseteq_E from \preceq). The required result follows from proposition 5.5.4 and lemma 5.5.5. For the converse, let \sqsubseteq be a relation on L that satisfies (RE1) to (RE5). Now define a relation \sqsubseteq_{EE} on L as follows: $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \sqsubseteq \beta$ or $\alpha \rightarrow \beta \not\sqsubseteq \alpha$. By lemma 5.5.11, \sqsubseteq_{EE} is an EE-ordering, and by theorem 3.3.1, there is thus a faithful total preorder \preceq from which \sqsubseteq_{EE} can be obtained using (Def \sqsubseteq_E from \preceq). By theorem 5.5.7, the faithful modular weak partial order semantically related to \preceq defines the RE-ordering \sqsubseteq_{RE} using (Def \sqsubseteq_{RE} from \sqsubseteq_{EE}). And by lemma 5.5.12, \sqsubseteq and \sqsubseteq_{RE} are identical. \square

5.5.3 Refined entrenchment and AGM contraction

Just as in the case of the EE-orderings and AGM contraction, the RE-orderings and AGM contraction are interdefinable; in this case using the following two identities:

(Def \sqsubseteq_{RE} from \sim) $\alpha \sqsubseteq_{RE} \beta$ iff $\alpha \rightarrow \beta \in K \sim \alpha \wedge \beta$

(Def – from \sqsubseteq_{RE}) $\beta \in K - \alpha$ iff $\beta \in K$ and $\begin{cases} \alpha \not\sqsubseteq_{RE} \beta \rightarrow \alpha, \text{ or} \\ \alpha \notin K \end{cases}$

Definition 5.5.14 An RE-ordering and an AGM contraction are *semantically related* iff they can be defined in terms of the same faithful modular weak partial order using (Def \sqsubseteq_E from \preceq) and (Def \sim from \preceq). \square

Theorem 5.5.15 Suppose that the RE-ordering \sqsubseteq_{RE} and the AGM contraction $-$ are semantically related.

1. \sqsubseteq_{RE} can also be defined in terms of $-$ using (Def \sqsubseteq_{RE} from \sim).
2. $-$ can also be defined in terms of \sqsubseteq_{RE} using (Def $-$ from \sqsubseteq_{RE}).

Proof 1. Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and $-$ are defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from \preceq). Now suppose that $\alpha \rightarrow \beta \notin K - \alpha \wedge \beta$. So there is a $y \in M(K) \cup \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ such that $y \in M(\alpha \wedge \neg\beta)$. If $y \in M(K)$ then $x \in M(\alpha)$ for every $x \leq y$ and so $\alpha \not\sqsubseteq_R \beta$. And similarly, if $y \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ then $x \in M(\alpha)$ for every $x \leq y$ and so $\alpha \not\sqsubseteq_{RE} \beta$. Conversely, suppose that $\alpha \not\sqsubseteq_{RE} \beta$. Then there is a $y \in \text{Min}_{\leq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \leq y$. So $y \in M(\alpha)$ and $y \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$, and thus $\alpha \rightarrow \beta \notin K - \alpha \wedge \beta$.

2. Follows from theorems 3.3.4 and 5.5.7. \square

And as in similar cases discussed before, the identities (Def \sqsubseteq_{RE} from \sim) and (Def $-$ from \sqsubseteq_{RE}) are interchangeable. That is, if we start with either an AGM contraction or an RE-ordering, and apply (Def \sqsubseteq_{RE} from \sim) and (Def $-$ from \sqsubseteq_{RE}) in the appropriate order, we end up with the same AGM contraction or RE-ordering. In fact, we can extend the interchangeability of identities further by noting that the identity (Def $-$ from \sqsubseteq_{EE}), when applied to the EE-ordering \sqsubseteq_{EE} , and the identity (Def $-$ from \sqsubseteq_{RE}), when applied to the RE-ordering \sqsubseteq_{RE} which is semantically related to \sqsubseteq_{EE} , both yield exactly the same AGM contraction $-$.

Corollary 5.5.16 Let the RE-ordering \sqsubseteq_{RE} , the EE-ordering \sqsubseteq_{EE} , and the AGM contraction $-$ be semantically related. Then $-$ can also be defined in terms of \sqsubseteq_{RE} using (Def $-$ from \sqsubseteq_{RE}), as well as in terms of \sqsubseteq_{EE} using (Def $-$ from \sqsubseteq_{EE}).

Proof Follows easily from proposition 3.3.4, and theorems 5.5.15 and 5.5.7. \square

In the case of a finitely generated propositional language, the definition of AGM contraction in terms of the RE-orderings can be simplified considerably. Consider a faithful modular weak partial order \leq on the interpretations of such a finite L , and let \sqsubseteq_{RE} be the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq). From part (10) of proposition 5.5.4, it follows that for every α , there is a $\beta \in [\alpha]_{\sqsubseteq_{RE}}$ such that $\gamma \models \beta$ for every $\gamma \in [\alpha]_{\sqsubseteq_{RE}}$. That is, every equivalence class $[\alpha]_{\sqsubseteq_{RE}}$ contains a logically weakest wff. We use $\ulcorner \alpha \urcorner_{RE}$ to denote a canonical representative of the logically weakest wffs in $[\alpha]_{\sqsubseteq_{RE}}$, and show below that if $\alpha, \beta \in K$, then $\beta \in K - \alpha$ iff $\beta \rightarrow \alpha \models \ulcorner \alpha \urcorner_{RE}$. That is, if α is in K , then checking whether a wff $\beta \in K$ is retained in $K - \alpha$ is a matter of checking whether $\beta \rightarrow \alpha$ entails a logically weakest wff in $[\alpha]_{\sqsubseteq_{RE}}$.

Proposition 5.5.17 *Let L be a finitely generated propositional language, \leq a faithful modular weak partial order, $-$ the AGM contraction defined in terms of \leq using (Def \sim from \preceq), and \sqsubseteq_{RE} the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq). If $\alpha, \beta \in K$ then $\beta \in K - \alpha$ iff $\beta \rightarrow \alpha \models \ulcorner \alpha \urcorner_{RE}$ (where $\ulcorner \alpha \urcorner_{RE}$ is a canonical representative of the logically weakest wffs in $[\alpha]_{\sqsubseteq_{RE}}$).*

Proof By theorem 5.5.15, if $\alpha, \beta \in K$, then $\beta \in K - \alpha$ iff $\alpha \not\sqsubseteq_{RE} \beta \rightarrow \alpha$. Since $\alpha \models \beta \rightarrow \alpha$, it follows from part (3) of proposition 5.5.4 that $\alpha \sqsubseteq_{RE} \beta \rightarrow \alpha$, and this result can thus be rewritten as follows: If $\alpha, \beta \in K$, then $\beta \in K - \alpha$ iff $\beta \rightarrow \alpha \in [\alpha]_{\sqsubseteq_{RE}}$. Now suppose that $\alpha, \beta \in K$. If $\beta \in K - \alpha$, then $\beta \rightarrow \alpha \in [\alpha]_{\sqsubseteq_{RE}}$, and since $\ulcorner \alpha \urcorner_{RE}$ is logically weaker than every wff in $[\alpha]_{\sqsubseteq_{RE}}$, $\beta \rightarrow \alpha \models \ulcorner \alpha \urcorner_{RE}$. Conversely, if $\beta \rightarrow \alpha \models \ulcorner \alpha \urcorner_{RE}$ then, by part (3) of proposition 5.5.4, $\beta \rightarrow \alpha \sqsubseteq_{RE} \ulcorner \alpha \urcorner_{RE}$. Furthermore, since $\ulcorner \alpha \urcorner_{RE} \in [\alpha]_{\sqsubseteq_{RE}}$, we get that $\ulcorner \alpha \urcorner_{RE} \sqsubseteq_{RE} \alpha$, and so, by the transitivity of \sqsubseteq_{RE} , $\beta \rightarrow \alpha \sqsubseteq_{RE} \alpha$. And because $\alpha \models \beta \rightarrow \alpha$, it follows from part (3) of proposition 5.5.4 that $\alpha \sqsubseteq_{RE} \beta \rightarrow \alpha$. Thus $\beta \rightarrow \alpha \in [\alpha]_{\sqsubseteq_{RE}}$, and it follows from the result above that $\beta \in K - \alpha$. \square

5.5.4 A comparison with generalised entrenchment

Since the EE-orderings are all instances of the LR-orderings of Lindström and Rabinowicz, the one-to-one correspondence between the EE-orderings and the RE-orderings provide an indirect relationship between RE-entrenchment and certain instances of LR-entrenchment. But we can also obtain a different connection by noting that the RE-orderings all satisfy the postulate (LR3).

Proposition 5.5.18 *Every RE-ordering \sqsubseteq_{RE} satisfies (LR3).*

Proof Suppose $\alpha \sqsubseteq_{RE} \beta$ and $\alpha \sqsubseteq_{RE} \gamma$, let \leq be a faithful modular weak partial order from which \sqsubseteq_{RE} is defined, using (Def \sqsubseteq_E from \preceq), and pick a $y \in M(\neg(\beta \wedge \gamma))$. So $y \in M(\neg\beta)$ or $y \in M(\neg\gamma)$. In the former case it follows from $\alpha \sqsubseteq_{RE} \beta$ that there is an $x \in M(\neg\alpha)$ such that $x \leq y$. And in the latter case it follows from $\alpha \sqsubseteq_{RE} \gamma$ that there is an $x \in M(\neg\alpha)$ such that $x \leq y$. So $\alpha \sqsubseteq_{RE} \beta \wedge \gamma$. \square

Since the LR-orderings require that all the wffs not in K be equally entrenched, the RE-orderings do not qualify as instances of the LR-orderings. However, the RE-orderings conform to the conditions imposed by the LR-orderings on the wffs in K . In this sense, there is an LR-ordering corresponding to every RE-ordering. On the other hand, the following example shows that some LR-orderings do not correspond to any RE-ordering, even when we restrict ourselves to just the wffs in K .

Example 5.5.19 Consider the propositional language L generated by the two atoms p and q with the valuation semantics (V, \models) , where $V = \{11, 10, 01, 00\}$. Now, let $K = Cn(p \wedge q)$, and consider the LR-ordering \sqsubseteq_{LR} defined as follows:

$$\alpha \sqsubseteq_{LR} \beta \text{ iff } \left\{ \begin{array}{l} \beta \in L \text{ if } \alpha \notin K, \\ p \wedge q \models \beta \text{ if } \alpha \equiv p \wedge q \text{ or } \alpha \equiv p \leftrightarrow q \text{ or } \alpha \equiv p \text{ or} \\ \alpha \equiv q \text{ or } \alpha \equiv \neg p \vee q, \\ p \vee q \models \beta \text{ if } \alpha \equiv p \vee q, \\ p \vee \neg q \models \beta \text{ if } \alpha \equiv p \vee \neg q, \\ \alpha \in L \text{ if } \models \beta. \end{array} \right.$$

Figure 5.4 contains a graphical representation of \sqsubseteq_{LR} . An inspection of figure 5.4 reveals that \sqsubseteq_{LR} is indeed an LR-ordering, but it can be verified that the part of \sqsubseteq_{LR} restricted to the elements of K does not coincide with the restriction of any RE-ordering \sqsubseteq_{RE} to K . \square

We now turn to a comparison of Rott's GEE-orderings (see section 5.4.2) and the RE-orderings. Observe firstly that, since the EE-orderings are total preorders, taking the converse complement of an EE-ordering \sqsubseteq_{EE} is the same as taking its strict version \sqsubset_{EE} . In the case of the RE-orderings, this is not the case, though. One way to obtain a comparison of the RE-orderings with the GEE-orderings is to check whether the RE-orderings satisfy the following translations of the GEE postulates into assertions about the converse complements of the GEE-orderings:

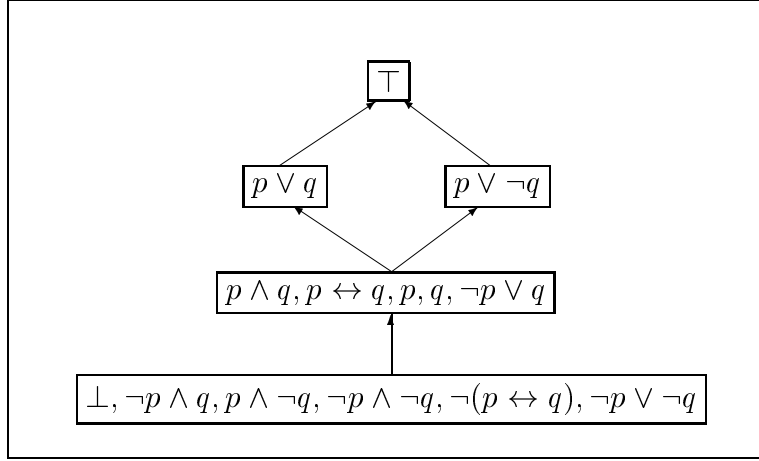


Figure 5.4: A graphical representation of the LR-ordering used in example 5.5.19. The ordering is obtained from the reflexive transitive closure of the relation determined by the arrows. Every wff in the figure is a canonical representative of the set of wffs that are logically equivalent to it.

(CGEE1) $\top \sqsubseteq \top$

(CGEE2 \uparrow) If $\gamma \sqsubseteq \alpha$ and $\beta \models \gamma$ then $\beta \sqsubseteq \alpha$

(CGEE2 \downarrow) If $\beta \sqsubseteq \gamma$ and $\gamma \models \alpha$ then $\beta \sqsubseteq \alpha$

(CGEE3 \uparrow) If $\beta \wedge \gamma \sqsubseteq \alpha$ then $\beta \sqsubseteq \alpha$ or $\gamma \sqsubseteq \alpha$

(CGEE3 \downarrow) If $\beta \sqsubseteq \alpha$ then $\beta \sqsubseteq \alpha \wedge \beta$

(CGEE4) If $K \neq Cn(\perp)$ then $\alpha \sqsubseteq \perp$ iff $\alpha \notin K$

(CGEE4') If $\beta \in K$ and $\beta \sqsubseteq \alpha$ then $\alpha \in K$

(CGEE5) If $\top \sqsubseteq \alpha$ then $\models \alpha$

(CGEE5') If $\top \sqsubseteq \beta$ and $\beta \sqsubseteq \alpha$ then $\top \sqsubseteq \alpha$

It is easily verified that the RE-orderings satisfy (CGEE1), (CGEE2 \uparrow) and (CGEE2 \downarrow), the three postulates regarded by Rott as minimal conditions of rationality for every relation designed to formalise the concept of epistemic entrenchment. Furthermore,

they also satisfy (CGEE3 \downarrow), (CGEE4'), (CGEE5) and (CGEE5'), but do not satisfy (CGEE3 \uparrow) or (CGEE4). They do satisfy the left-to-right direction of (CGEE4), though.

Proposition 5.5.20 *The RE-orderings satisfy the postulates (CGEE1), (CGEE2 \uparrow), (CGEE2 \downarrow), (CGEE3 \downarrow), (CGEE4'), (CGEE5) and (CGEE5'). Furthermore, they do not necessarily satisfy (CGEE3 \uparrow) or (CGEE4), but they do satisfy the left-to-right direction of (CGEE4).*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} can be defined using (Def \sqsubseteq_E from \preceq). (CGEE1) follows from (RE2), and (CGEE2 \uparrow) and (CGEE2 \downarrow) both from (RE1) and (RE2). For (CGEE3 \downarrow), suppose that $\beta \sqsubseteq_{RE} \alpha$ and pick any $y \in M(\neg(\alpha \wedge \beta))$. So $y \in M(\neg\beta)$ or $y \in M(\neg\alpha)$. We have to show that there is an $x \leq y$ such that $x \in M(\neg\beta)$. If $y \in M(\neg\beta)$, this follows from the reflexivity of \leq , and if $y \in M(\neg\alpha)$, it follows from the fact that $\beta \sqsubseteq_{RE} \alpha$. (CGEE4') follows from (RE4a), (CGEE5) from (RE2) and (RE5), and (CGEE5') follows from (RE1).

To show that the RE-orderings do not always satisfy (CGEE3 \uparrow), let $K = Cn(\{p \leftrightarrow q\})$ and consider the RE-ordering \sqsubseteq_{RE} , with respect to K , which is defined as follows:

$$\alpha \sqsubseteq_{RE} \beta \text{ iff } \begin{cases} \beta \in L \text{ if } \alpha \notin K, \\ \beta \in K \text{ if } \alpha \equiv p \leftrightarrow q, \\ p \rightarrow q \models \beta \text{ if } \alpha \equiv p \rightarrow q, \\ q \rightarrow p \models \beta \text{ if } \alpha \equiv q \rightarrow p, \\ \models \beta \text{ if } \models \alpha. \end{cases}$$

It is readily verified that \sqsubseteq_{RE} is indeed an RE-ordering. By letting $\alpha = p \leftrightarrow q$, $\beta = q \leftarrow q$, $\gamma = p \leftarrow q$, and observing that $\beta \wedge \gamma \equiv \alpha$, we see that \sqsubseteq_{RE} violates (CGEE3 \uparrow).

To show that the RE-orderings do not always satisfy (CGEE4), it is sufficient to observe that the entailment relation \models is an RE-ordering with respect to the belief set $Cn(\top)$. Finally, that every RE-ordering satisfies the left-to-right direction of (CGEE4) follows from (RE4a). \square

As observed above, the converse complement of an EE-ordering is the same as its strict version. It might therefore be instructive to determine whether or not the strict versions of the RE-orderings are instances of the GEE-orderings. It turns out that the

strict RE-orderings satisfy (GEE1), (GEE2 \uparrow), and (GEE2 \downarrow), but do not always satisfy (GEE3 \uparrow) and (GEE3 \downarrow).

Proposition 5.5.21 *Let \sqsubseteq_{RE} be the strict version of an RE-ordering. Then \sqsubseteq_{RE} satisfies (GEE1), (GEE2 \uparrow) and (GEE2 \downarrow), but does not, in general, satisfy either (GEE3 \uparrow) or (GEE3 \downarrow).*

Proof (GEE1) is trivial. (GEE2 \uparrow) and (GEE2 \downarrow) both follow from (RE1) and (RE2). To show that (GEE3 \uparrow) and (GEE3 \downarrow) do not always hold, consider the LR-ordering \sqsubseteq_{LR} in example 5.5.19. It is easily verified that \sqsubseteq_{LR} is also an RE-ordering, defined in terms of the faithful modular weak partial order \leq using (Def \sqsubseteq_E from \preceq), where \leq is defined as follows:

$$\leq = \{(x, x) \mid x \in U\} \cup \{(11, y) \mid y \in U\} \cup \{(10, 00), (01, 00)\}.$$

Figure 5.5 contains a graphical representation of \leq and the RE-ordering \sqsubseteq_{RE} defined in terms of \leq using (Def \sqsubseteq_E from \preceq). Note that the LR-ordering in example 5.5.19 is identical to \sqsubseteq_{RE} . As noted in example 5.5.19, \sqsubseteq_{LR} violates both (GEE3 \uparrow) and (GEE3 \downarrow). \square

With regard to the supplementary postulates, the strict RE-orderings satisfy all but the left-to-right direction of (GEE4).

Proposition 5.5.22 *The strict version \sqsubseteq_{RE} of an RE-ordering satisfies the right-to-left direction of (GEE4) (but not the left-to-right direction), as well as (GEE4'), (GEE5) and (GEE5').*

Proof (RE4b) ensures that the left-to-right direction of (GEE4) does not always hold, and (RE4a) ensures that the right-to-left direction holds. (GEE4') follows from (RE4a), and both (GEE5) and (GEE5') follow from (RE2) and (RE5). \square

The results above seem to suggest that the GEE-orderings and the RE-orderings have quite different intuitions associated with them. Whereas the GEE-orderings constitute a proper generalisation of the EE-orderings, the RE-orderings should be seen as alternatives to the EE-orderings. This is highlighted when the link with contraction is investigated. Rott applies (Def – from \sqsubseteq_{EE}) to the GEE-orderings to obtain a set of contractions that is a strict superset of AGM contraction. In contrast, (Def – from \sqsubseteq_{RE}) applied to the RE-orderings yields precisely the set of AGM contractions.

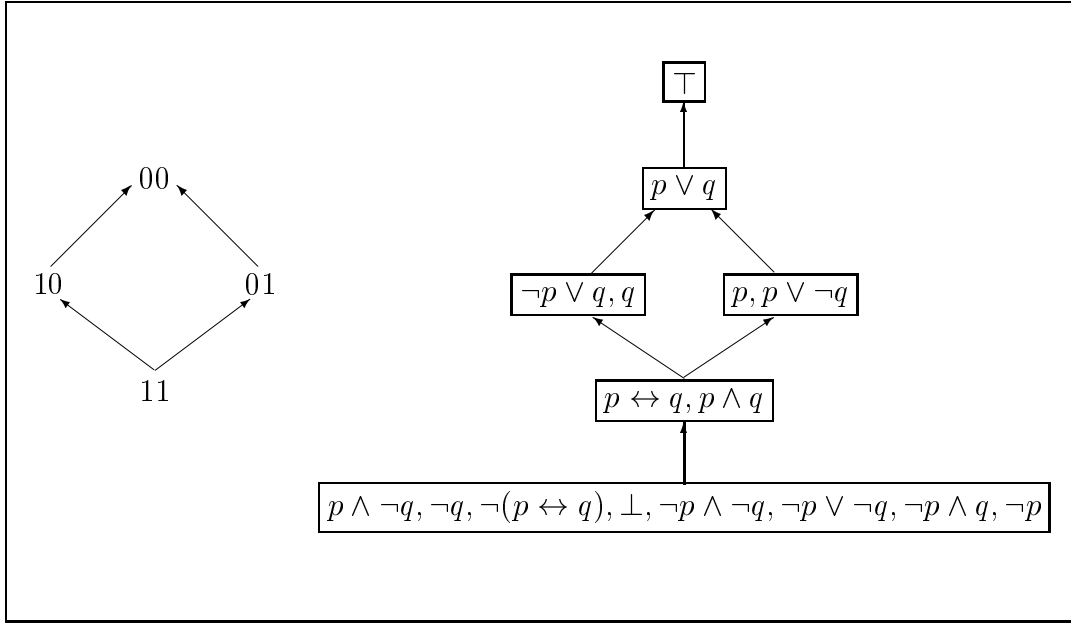


Figure 5.5: A graphical representation of the faithful modular weak partial order \leq used in proposition 5.5.21, as well as the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq). Both orderings are obtained from the reflexive transitive closure of the respective relations determined by the arrows. Every wff in the representation of the RE-ordering is a canonical representative of the set of wffs that are logically equivalent to it.

5.5.5 Refined G-plausibility

We have seen in theorem 3.3.1 that the duals of the EE-orderings (the GE-orderings of Grove) can be defined in terms of the faithful total preorders using (Def \sqsubseteq_G from \preceq). In a similar manner, by applying (Def \sqsubseteq_G from \preceq) to the faithful modular weak partial orders, we can obtain a set of orderings that are dual to the RE-orderings. We shall refer to them as the RG-orderings.

Definition 5.5.23 An *RG-ordering* is a binary relation on L obtained in terms of a faithful modular weak partial order using (Def \sqsubseteq_G from \preceq). We say that a GE-ordering and an RG-ordering, defined respectively in terms of a faithful total preorder and its semantically related modular weak partial order, using (Def \sqsubseteq_G from \preceq), are *semantically related*. \square

From the definitions of the RE-orderings and the RG-orderings it is clear that they can be defined in terms of one another using (Def \sqsubseteq_E from \sqsubseteq_G) and (Def \sqsubseteq_G from \sqsubseteq_E). By virtue of (Def \sqsubseteq_G from \sqsubseteq_E), the results about the RE-orderings can thus be translated into results about the RG-orderings. While such an exercise would serve little purpose in most cases, it is our intention to concentrate on two aspects pertaining to the use of the RG-orderings. The first is a comparison of the suitability of the RG-orderings and the GE-orderings as orderings of plausibility. The second aspect involves the definition of AGM revision in terms of the RG-orderings. The results provided in the proposition below will be used in the discussion of these two aspects.

Proposition 5.5.24 *Let \sqsubseteq_{RG} be the RG-ordering defined in terms of the faithful modular weak partial order \leq using (Def \sqsubseteq_G from \preceq). Then \sqsubseteq_{RG} satisfies the following properties.*

1. \sqsubseteq_{RG} is a preorder (that need not be total).
2. Suppose that the GE-ordering \sqsubseteq_{GE} and the \sqsubseteq_{RG} are semantically related. If $\alpha \sqsubseteq_{RG} \beta$ then $\alpha \sqsubseteq_{GE} \beta$.
3. If $\alpha \models \beta$ then $\beta \sqsubseteq_{RG} \alpha$.
4. $\models \neg\alpha$ iff for all $\beta \in L$, $\beta \sqsubseteq_{RG} \alpha$.
5. If K is satisfiable then $K = [\top]_{\sqsubseteq_{RG}}$.
6. If $K \models \alpha$ and $K \not\models \beta$ then $\alpha \sqsubset_{RG} \beta$.
7. If $K \not\models \neg\beta$ and $K \models \neg\gamma$ then $\beta \sqsubset_{RG} \gamma$.
8. If $K \cup \{\alpha\} \not\models \perp$ then $K \cup \{\alpha\} \models \beta$ iff $\beta \sqsubseteq_{RG} \alpha$.

Proof The proofs are similar to that of proposition 5.5.4 and are omitted. \square

These properties reveal that the RG-orderings are finer grained versions of the GE-orderings. They are preorders like the GE-orderings, but they need not be total. For every satisfiable belief set K , they partition the set of wffs into four disjoint sets, and not three, as the GE-orderings do. The logically invalid wffs are all equivalent and strictly above all other wffs, followed by the rest of the K -refuted wffs, just as with the GE-orderings. However, whereas the GE-orderings lump the K -established wffs

together with the wffs that are K -undecided, the RG-orderings distinguish between these two sets. In particular, the wffs in K are all equivalent and strictly lower than all other wffs, while the K -undecided wffs are strictly below all K -refuted wffs. Finally, the K -undecided wffs are not all comparable. Instead, any RG-ordering restricted to the K -undecided wffs (or, in fact, restricted to all wffs except those that are K -refuted) is exactly the inverse of entailment relative to K . So the only part of any RG-ordering that is not completely specified by K itself, is the ordering restricted to the wffs that are not logically invalid, but are nevertheless K -refuted.

In view of these results, the RG-orderings seem to be more suitable as plausibility orderings than the GE-orderings. This is due mainly to the fact that they are finer grained versions of the GE-orderings. Unlike the GE-orderings, an RG-ordering (with respect to K) does not regard the wffs in K and the K -undecided wffs as equally plausible, or equally close to the belief set K . Instead, all the wffs in K are seen as more plausible than the K -undecided wffs, a result that surely is more in line with the intuition of plausibility. More important, perhaps, is that the added structure of the RG-orderings also extends to the K -refuted wffs, enabling us to give a definition of revision that is both simpler and more intuitively appealing than (Def * from \sqsubseteq_{GE}).

The intuition underlying the use of the RG-orderings to define AGM revision can be described as follows. When revising a belief set K with a wff α , $K * \alpha$ should consist of a set of wffs that entails α , while still being satisfiable. Now, the set consisting of the wffs that are precisely as plausible as α , certainly entails α (since it includes α). So if this set is satisfiable, all the wffs in it should be included in $K * \alpha$. As a result, we should choose $K * \alpha$ to consist of all the wffs entailed by the set of wffs that is precisely as plausible as α . That is, AGM revision can be defined in terms of the RG-orderings as follows:

(Def * from \sqsubseteq_{RG}) $K * \alpha = Cn([\alpha]_{\sqsubseteq_{RG}})$

Theorem 5.5.25 *Let \leq be a faithful modular weak partial order and let \sqsubseteq_{RG} be the RG-ordering defined in terms of \leq using (Def \sqsubseteq_G from \leq). The revision defined in terms of \leq using (Def * from \leq) can also be defined in terms of \sqsubseteq_{RG} using (Def * from \sqsubseteq_{RG}).*

Proof It suffices to show that for all $\alpha, \beta \in L$, $[\alpha]_{\sqsubseteq_{RG}} \models \beta$ iff $\beta \in Th(Min_{\leq}(\alpha))$. So suppose that $[\alpha]_{\sqsubseteq_{RG}} \models \beta$. That is, $\beta \in Th(\bigcap \{M(\gamma) \mid \gamma \in [\alpha]_{\sqsubseteq_{RG}}\})$. By corollary

5.5.6 and theorem 2.3.5, $Min_{\leq}(\alpha) = Min_{\leq}(\gamma)$ for every $\gamma \in [\alpha]_{\sqsubseteq_{RG}}$. So $Min_{\leq}(\alpha) \subseteq \bigcap \{M(\gamma) \mid \gamma \in [\alpha]_{\sqsubseteq_{RG}}\}$, and thus $\beta \in Th(Min_{\leq}(\alpha))$. Conversely, suppose that $\beta \in Th(Min_{\leq}(\alpha))$. Clearly, we also have that $\alpha \wedge \beta \in Th(Min_{\leq}(\alpha))$. So, for every $y \in M(\alpha)$ there is an $x \in M(\alpha \wedge \beta)$ such that $x \leq y$, which means that $\alpha \wedge \beta \sqsubseteq_{RG} \alpha$. Furthermore, because $\alpha \wedge \beta \models \alpha$, it follows from part (3) of proposition 5.5.24 that $\alpha \sqsubseteq_{RG} \alpha \wedge \beta$. So $\alpha \wedge \beta \equiv_{\sqsubseteq_{RG}} \alpha$ and thus $[\alpha]_{\sqsubseteq_{RG}} \models \beta$. \square

For the sake of completeness, we include two identities that can be used to define the GE-orderings and the RG-orderings in terms of each other. That these two identities can indeed be used for this purpose follows easily by applying (Def \sqsubseteq_G from \sqsubseteq_E), (Def \sqsubseteq_{EE} from \sqsubseteq_{RE}), (Def \sqsubseteq_{RE} from \sqsubseteq_{RG}), (Def \sqsubseteq_{RE} from \sqsubseteq_{EE}) and (Def \sqsubseteq_E from \sqsubseteq_G).

(Def \sqsubseteq_{GE} from \sqsubseteq_{RG}) $\alpha \sqsubseteq_{GE} \beta$ iff $\alpha \sqsubseteq_{RG} \beta$ or $\neg\alpha \wedge \beta \not\sqsubseteq_{RG} \alpha$

(Def \sqsubseteq_{RG} from \sqsubseteq_{GE}) $\alpha \sqsubseteq_{RG} \beta$ iff $\models \neg\beta$ or $\alpha \sqsubset_{GE} \beta$ or $\beta \sqsubseteq_{GE} \neg\alpha \wedge \beta$

5.6 Other alternatives

In this section we take a brief look at ways to remove the minimality and maximality conditions on the EE-orderings (see section 2.3). Two proposals in which both these conditions do not feature are the GEE-orderings of Rott, considered in section 5.4.2, and the expectation orderings of Gärdenfors and Makinson [1994]. The expectation orderings are required to satisfy the postulates (EE1), (EE2) and (EE3), but not (EE4) and (EE5), and can thus be seen as the EE-orderings without the minimality and maximality conditions imposed on them. They are used to define the expectation based consequence relations, discussed in section 4.4.1, as follows:

(Def \sim from \sqsubseteq) $\alpha \sim \beta$ iff $\beta \in Cn(\{\alpha\} \cup \{\gamma \mid \neg\alpha \sqsubset \gamma\})$

Interestingly enough, Gärdenfors and Makinson point out that the expectation based consequence relations can also be defined in terms of the EE-orderings using (Def \sim from \sqsubseteq). So if the interest in the expectation orderings is motivated solely on their relationship with the expectation based consequence relations, we might as well stick to the EE-orderings.

Let us now take a closer look at these two conditions individually. We first consider the maximality condition — the requirement that the most entrenched wffs are nothing

but the logically valid wffs. The most obvious way to remove this requirement is to remove the corresponding postulate (or postulates). Thus, for the EE-orderings it is a matter of removing (EE4), for the LR-orderings the removal of (LR5), and for the GEE-orderings the removal of (GEE5) and (GEE5'). But there is also an elegant semantic way in which to consider this issue. Intuitively, the objection to this maximality condition is that some of the beliefs of an agent might be so entrenched as to be on the same level as the logically valid wffs. It is reasonable to regard these wffs as being entrenched to such an extent that they cannot be dislodged from the belief set of the agent. As such, they should rather be seen as part of the *knowledge* of the agent. We can achieve the desired effect by moving to a new semantics for L in which these wffs are logically valid. This new semantics is obtained from the current one by taking the new set of interpretations as the set of models of these wffs. That is, if B is the set of beliefs that should be regarded as knowledge, we replace the current semantics (U, \Vdash) by the new semantics $(M(B), \Vdash_B)$, where \Vdash_B is just the satisfaction relation \Vdash with the first coordinates restricted to $M(B)$.

We now turn to the minimality condition — that all wffs not in K should be equally entrenched and strictly less entrenched than the wffs in K . The objection to this requirement is, of course, concerned with the insistence that all wffs not in K be equally entrenched, and not with the decision to place the wffs that are not in K strictly below the wffs in K . In fact, it seems reasonable to require that all versions of entrenchment should satisfy (RE4a) on page 115.⁸ This is the condition termed *stability* by Rabinowicz [1995], and is clearly satisfied by all the versions of entrenchment that we have considered so far.

With regard to the issue of the comparability of the wffs not in K , we can distinguish between three approaches. The first is to apply the same conditions that are being placed on the comparability of wffs in K . Thus, for the EE-orderings it is a matter of applying (EE3) to the wffs not in K as well, and replacing (EE4) by (RE4a), while such a suggestion applied to the LR-orderings merely involves the replacement of (LR4) by (RE4a).

⁸If there is no explicit mention of a belief set, the extraction of a suitable one from the entrenchment ordering should be performed in such a way as to ensure the satisfaction of (RE4a). For example, Rott's basic GEE-orderings do not refer to a belief set, but the belief set obtained from a GEE-ordering \sqsubset_{GEE} is taken as the set $\{\alpha \mid \perp \sqsubset_{GEE} \alpha\}$.

A different suggestion due to Rabinowicz [1995], and one that relates specifically to the EE-orderings, is to use their semantically related GE-orderings (obtained using (Def \sqsubseteq_G from \sqsubseteq_E)) to distinguish between the wffs not in K . As we have seen in section 5.3, this corresponds to Spohn's R-orderings. For a satisfiable belief set K , an R-ordering \sqsubseteq_R partitions the wffs of L into three classes: Those that are K -believed (the wffs in K), those that are K -disbelieved (the K -refuted wffs), and those that are K -neutral (the K -undecided wffs). The K -neutral wffs are all equally entrenched, strictly less entrenched than those wffs that are K -believed, but strictly more entrenched than the K -disbelieved wffs. Note, however, that the relative ordering of the K -believed wffs is a mirror image of the relative ordering of the K -disbelieved wffs, and vice versa. That is, if α and β are both K -believed, or if both are K -disbelieved, then $\alpha \sqsubseteq_R \beta$ iff $\neg\beta \sqsubseteq_R \neg\alpha$. The R-orderings thus involve the imposition of a kind of symmetry between the ordering of the belief set and the *disbelief set* of an agent that seems difficult to justify.

As a result, we propose to generalise this idea as follows. Instead of combining an EE-ordering \sqsubseteq_{EE} and the particular GE-ordering obtained in terms of \sqsubseteq_{EE} using (Def \sqsubseteq_G from \sqsubseteq_E), we combine \sqsubseteq_{EE} and *any* GE-ordering (with respect to the same belief set K). To be more specific, given any EE-ordering \sqsubseteq_{EE} and any GE-ordering \sqsubseteq_{GE} , both with respect to the same belief set K , we define a *combined entrenchment* ordering \sqsubseteq_C in terms of \sqsubseteq_{EE} and \sqsubseteq_{GE} as follows:

$$(\text{Def } \sqsubseteq_C \text{ from } \sqsubseteq_{EE} \text{ and } \sqsubseteq_{GE}) \quad \alpha \sqsubseteq_C \beta \text{ iff } \begin{cases} \alpha \sqsubseteq_{EE} \beta \text{ if } \alpha, \beta \in K, \\ \beta \sqsubseteq_{GE} \alpha \text{ if } \alpha, \beta \notin K, \\ \alpha \notin K \text{ and } \beta \in K \text{ otherwise} \end{cases}$$

The combined entrenchment orderings retain the partitioning of the R-orderings, as well as the property that all K -neutral wffs are equally entrenched, but have the advantage of not being subject to the requirements of symmetry between the belief set and the disbelief set of an agent.

We conclude this section with some thoughts on the way refined entrenchment handles the comparability of wffs not in K . Although the RE-orderings are able to distinguish between the entrenchment of wffs not in K , this ability is little more than a reflection of the underlying entailment relation \models and does not seem to express a genuine difference in the entrenchment of such wffs. For a more satisfactory description of the relative entrenchment of such wffs, we have a choice between the two proposals related to the minimality condition, applied to the RE-orderings. The application of the first

proposal involves doing away with (RE4b), and applying (RE3a) to all wffs, and not just those in K . The second proposal involves the RE-orderings and the RG-orderings. Given any RE-ordering \sqsubseteq_{RE} and any RG-ordering \sqsubseteq_{RG} , both with respect to the same belief set K , we define a CR-ordering \sqsubseteq_{CR} in terms of \sqsubseteq_{RE} and \sqsubseteq_{RG} using (Def \sqsubseteq_C from \sqsubseteq_{EE} and \sqsubseteq_{GE}).

Definition 5.6.1 A binary relation on L is a *CR-ordering*, with respect to a belief set K , iff it is defined in terms of an RE-ordering with respect to K , and an RG-ordering with respect to K , using (Def \sqsubseteq_C from \sqsubseteq_{EE} and \sqsubseteq_{GE}). \square

From the properties of the RE-orderings and the RG-orderings, it follows that for a satisfiable belief set K , every CR-ordering \sqsubseteq_{CR} partitions the wffs of L into five classes:

1. The logically valid wffs are all equally entrenched, and more entrenched than all other wffs.
2. The wffs that are K -believed, but not logically valid, are strictly less entrenched than the logically valid wffs, and more entrenched than all other wffs.
3. The K -neutral wffs are less entrenched than the K -believed wffs and more entrenched than the K -disbelieved wffs. Moreover, the CR-ordering restricted to the K -neutral wffs corresponds to entailment relative to K .
4. The K -disbelieved wffs that are not logically invalid are less entrenched than the K -believed and the K -neutral wffs, but more entrenched than the logically invalid wffs.
5. And finally, the logically invalid wffs are all equally entrenched, and less entrenched than all other wffs.

5.7 Unifying epistemic and refined entrenchment

From the discussion on refined entrenchment it is clear that the RE-orderings are intended as alternatives to the EE-orderings. This view is supported by the results about the connection between the RE-orderings, the EE-orderings and AGM contraction. The main difference between the RE-orderings and the EE-orderings is that the RE-orderings are not all total preorders. And while this renders the RE-orderings

more appropriate in certain respects, it has its downside as well. For in embracing the RE-orderings at the expense of the EE-orderings, we also discard the property of being able to compare all wffs in all but some limiting cases. The question that arises is thus whether it is possible to obtain a unified view of entrenchment, encompassing both refined entrenchment and epistemic entrenchment. From a semantic viewpoint, there is a positive answer to this question. It involves the use of a set of faithful preorders which strictly includes the faithful total preorders and the faithful modular weak partial orders.

Definition 5.7.1 A preorder \preceq on a set V is called *layered* iff for every $x, y, z \in V$, if $z \prec x$ and either $x \equiv_{\preceq} y$ or $x \parallel_{\preceq} y$, then $z \prec y$. \square

Layered preorders appeal to the same intuition that underlies the total preorders, the modular weak partial orders and the modular (strict) partial orders. The idea is that the elements of V are arranged in levels, with elements in different layers being comparable. The difference between all these types of orderings concerns the way in which elements in the same layer are treated. So, while the total preorders regard all elements in the same layer as *comparable*, and the modular weak partial orders take all distinct elements in the same layer as *incomparable*, the layered preorders provide a compromise between these two extremes: they allow for both the comparability and the incomparability of elements in the same layer. Using this intuition, it is clear that every layered preorder is uniquely associated with a modular weak partial order and a total preorder. (And in fact, every total preorder and every modular weak partial is a layered preorder.)

Definition 5.7.2 A modular weak partial order \leq on a set X , a total preorder \preceq on X , and a layered preorder \preccurlyeq on X are *semantically related* iff \leq can be defined in terms of \preccurlyeq using (Def \leq from \preceq) and \preceq can be defined in terms of \preccurlyeq using (Def \preceq from \leq). \square

It is easily verified that a faithful layered preorder and its semantically related faithful total preorder and faithful modular weak partial order are minimal-equivalent (see definition 3.3.6).

Proposition 5.7.3 A removal and a revision defined in terms of a faithful layered preorder \preccurlyeq using (Def \sim from \preceq) and (Def $*$ from \preceq), can also be defined in terms

of its semantically related faithful total preorder \preceq , and its semantically related faithful modular weak partial order \leq , using (Def \sim from \preceq) and (Def $*$ from \preceq).

Proof Follows from the fact that $Min_{\preceq}(\alpha) = Min_{\leq}(\alpha) = Min_{\preceq}(\alpha)$ for every $\alpha \in L$. \square

As a result we can use either the faithful layered preorders, or the faithful modular weak partial orders, or the faithful total preorders to characterise AGM theory change.

Corollary 5.7.4 *Let \preceq be a faithful layered preorder, let \leq be the faithful modular weak partial order obtained in terms of \preceq using (Def \leq from \preceq), and let \preceq be the faithful total preorder obtained in terms of \preceq using (Def \preceq from \preceq).*

1. *The AGM revisions defined in terms of \leq , \preceq , and \preceq using (Def $*$ from \preceq) are identical.*
2. *The AGM contractions defined in terms of \leq , \preceq , and \preceq using (Def \sim from \preceq) are identical.*

Proof Follows from theorem 3.2.6 and proposition 5.7.3. \square

From an information-theoretic point of view, the faithful layered preorders provide us with a degree of freedom that is lacking in both the faithful total preorders and the faithful modular weak partial orders. It allows us to regard some infatoms as being incomparable with respect to entrenchment, and others to be equally entrenched. As a result, the faithful layered preorders can be used to define a class of entrenchment orderings that generalises both the RE-orderings and the EE-orderings.

Definition 5.7.5 A binary relation \sqsubseteq_{GRE} is a *GRE-ordering* iff it is defined in terms of a faithful layered preorder \preceq using (Def \sqsubseteq_E from \preceq). We say that a GRE-ordering, an RE-ordering, and an EE-ordering defined respectively in terms of a faithful layered preorder, its semantically related total preorder, and its semantically related modular weak partial order, using (Def \sqsubseteq_E from \preceq), are *semantically related*. \square

From theorem 3.3.1, definitions 5.5.3 and 5.7.5, and the fact that the faithful total preorders and the faithful modular weak partial orders are instances of the faithful layered preorders, it immediately follows that the EE-orderings and the RE-orderings are all instances of the GRE-orderings. We conclude with a list of properties of the

GRE-orderings. The obvious question, whether there is a set of postulates that gives a precise description of the GRE-orderings, seems to be a non-trivial one. We leave a proper investigation of this issue, and the quest for an appropriate representation theorem, for future research.

Proposition 5.7.6 *Let \preceq be a faithful layered preorder, and let \sqsubseteq_{GRE} be the GRE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq). Then \sqsubseteq_{GRE} satisfies the following properties.*

1. \sqsubseteq_{GRE} is a preorder (that need not be total).
2. Suppose that \sqsubseteq_{GRE} and an EE-ordering \sqsubseteq_{EE} are semantically related. If $\alpha \sqsubseteq_{GRE} \beta$ then $\alpha \sqsubseteq_{EE} \beta$.
3. Suppose that \sqsubseteq_{GRE} and an RE-ordering \sqsubseteq_{RE} are semantically related. If $\alpha \sqsubseteq_{RE} \beta$ then $\alpha \sqsubseteq_{GRE} \beta$.
4. If $\alpha \models \beta$ then $\alpha \sqsubseteq_{GRE} \beta$.
5. $\alpha \sqsubseteq_{GRE} \beta$ for all α iff $\models \beta$.
6. If $\alpha \equiv \beta$ then $\alpha \sqsubseteq_{GRE} \gamma$ iff $\beta \sqsubseteq_{GRE} \gamma$, and $\gamma \sqsubseteq_{GRE} \alpha$ iff $\gamma \sqsubseteq_{GRE} \beta$.
7. If K is satisfiable then $\{\alpha \mid \neg\alpha \in K\} \subseteq [\perp]_{\sqsubseteq_{RE}}$.
8. If $\alpha \notin K$ and $\beta \in K$ then $\alpha \sqsubset_{GRE} \beta$.
9. If $\neg\beta \in K$ and $\neg\gamma \notin K$ then $\beta \sqsubseteq_{GRE} \gamma$.
10. If $\alpha \notin K$ and $K \cup \{\alpha\} \models \beta$ then $\alpha \sqsubseteq_{GRE} \beta$.
11. If $\alpha \equiv_{\sqsubseteq_{GRE}} \beta$ then $\alpha \wedge \beta \in [\alpha]_{\sqsubseteq_{GRE}} = [\beta]_{\sqsubseteq_{GRE}}$.
12. $\alpha \sqsubseteq_{GRE} \alpha \wedge \beta$, or $\beta \sqsubseteq_{GRE} \alpha \wedge \beta$, or both $\alpha \rightarrow \beta \not\sqsubseteq_{GRE} \alpha$ and $\beta \rightarrow \alpha \not\sqsubseteq_{GRE} \beta$.

Proof The reflexivity and transitivity of \sqsubseteq_{GRE} are trivial. To show that \sqsubseteq_{GRE} need not be a total preorder, consider the example of a propositional language generated by two atoms, p and q . Now let $K = Cn(p)$ and consider the faithful layered preorder

$$\{(x, x) \mid x \in U\} \cup \{(x, y) \mid x \in M(K) \text{ and } y \notin M(K)\}.$$

It is easily verified that $q \not\sqsubseteq_{GRE} \neg q$ and $\neg q \not\sqsubseteq_{GRE} q$. For (2) and (3), let \preccurlyeq be a faithful layered preorder in terms of which \sqsubseteq_{GRE} can be defined using (Def \sqsubseteq_E from \preceq), let \leq be the faithful modular weak partial order that is semantically related to \preccurlyeq , and let \preceq be the faithful total preorder that is semantically related to \preccurlyeq . Then (2) follows from the fact that if $x \preccurlyeq y$ then $x \preceq y$, and (3) from the fact that if $x \leq y$ then $x \preccurlyeq y$. (4) is trivial. For (5), suppose that $\alpha \sqsubseteq_{GRE} \beta$ for all α . So in particular $\top \sqsubseteq_{GRE} \beta$, which can only be if $M(\neg\beta) = \emptyset$. Therefore $\models \beta$. Conversely, if $\models \beta$ then $M(\neg\beta) = \emptyset$, from which it follows vacuously that $\alpha \sqsubseteq_{GRE} \beta$ for all α . (6) is trivial. For (7), suppose that K is satisfiable and pick any α such that $\neg\alpha \in K$. $\perp \sqsubseteq_{GRE} \alpha$ follows from $\perp \models \alpha$ and part (4), and $\alpha \sqsubseteq_{GRE} \perp$ from the faithfulness of \preccurlyeq . For (8), suppose that $\alpha \notin K$ and $\beta \in K$. So $M(K) \cap M(\neg\beta) = \emptyset$, and since K has a model that satisfies $\neg\alpha$, it follows from faithfulness that for every $y \in M(\neg\beta)$ there is an $x \in M(\neg\alpha)$ such that $x \preccurlyeq y$. That is, $\alpha \sqsubseteq_{GRE} \beta$. On the other hand, since K has a model y that satisfies $\neg\alpha$, and since all models of K satisfy β , it follows from faithfulness that $x \in M(\beta)$ for every $x \preccurlyeq y$, and thus $\beta \not\sqsubseteq_{GRE} \alpha$. For (9), suppose that $\neg\beta \in K$ and $\neg\gamma \notin K$. $\beta \sqsubseteq_{GRE} \gamma$ follows from faithfulness. For the proof of (10), let $\alpha \notin K$ and suppose that $K \cup \{\alpha\} \models \beta$. Now pick any $y \in M(\neg\beta)$. If $y \notin M(K)$ then, because $M(K) \cap M(\neg\alpha) \neq \emptyset$, there is an $x \in M(K) \cap M(\neg\alpha)$ such that $x \preccurlyeq y$. And if $y \in M(K)$ then, because $M(K) \cap M(\alpha) \subseteq M(\beta)$, $y \notin M(\alpha)$, and there is thus an $x \in M(\neg\alpha)$ such that $x \preceq y$. So $\alpha \sqsubseteq_{GRE} \beta$. For the proof of (11), suppose that $\alpha \equiv_{GRE} \beta$. By part (4), $\alpha \wedge \beta \sqsubseteq_{GRE} \alpha$. To show that $\alpha \sqsubseteq_{GRE} \alpha \wedge \beta$, pick a $y \in M(\neg\alpha \vee \neg\beta)$. If $y \in M(\neg\beta)$ then $\alpha \sqsubseteq_{GRE} \beta$ guarantees that there is an $x \in M(\neg\alpha)$ such that $x \preccurlyeq y$, and the case where $y \in M(\neg\alpha)$ is trivial. For the proof of (12), suppose that $\alpha \not\sqsubseteq_{GRE} \alpha \wedge \beta$ and $\beta \not\sqsubseteq_{GRE} \alpha \wedge \beta$. Then there is a $y \in Min_{\preccurlyeq}(\neg\alpha \vee \neg\beta)$ such that $x \in M(\alpha)$ for every $x \preccurlyeq y$, and there is a $v \in Min_{\preccurlyeq}(\neg\alpha \vee \neg\beta)$ such that $u \in M(\beta)$ for every $u \preccurlyeq v$. So $y \in M(\alpha) \cap M(\neg\beta)$ and $x \in M(\alpha) \cap M(\beta)$ for every $x \prec y$. Similarly, $v \in M(\neg\alpha) \cap M(\beta)$ and $u \in M(\alpha) \cap M(\beta)$ for every $u \prec v$. Since \preccurlyeq is a layered preorder, it therefore has to be the case that $y \parallel_{\preccurlyeq} v$. So $y \in M(\neg\beta)$ and $x \in M(\beta \rightarrow \alpha)$ for every $x \preccurlyeq y$. That is, $\beta \rightarrow \alpha \not\sqsubseteq_{GRE} \beta$. And similarly for v , $\alpha \rightarrow \beta \not\sqsubseteq_{GRE} \alpha$. \square

5.8 Summary

Entrenchment orderings play an important role in belief change. They are regarded as more fundamental than theory change operations such as revision and contraction

[Gärdenfors, 1988, p. 88], and are seen as suitable representations of the epistemic states of an agent [Nayak, 1994a,b, Nayak et al., 1996], at least for the part pertaining to belief change. While we are in agreement with the idea of entrenchment being more basic than theory change operations, it should come as no surprise that our view concerning the representation of epistemic states is rather different. We regard the faithful layered preorders as more fundamental than entrenchment, a view that is supported by the results in this chapter. In particular, we saw that the different kinds of entrenchment orderings discussed all turn out to be have a semantic basis, and more specifically, are rooted in (some subset of) the faithful layered preorders. This prompts the following generalisation of definitions 3.3.8, 5.5.3 and 5.5.14.

Definition 5.8.1 An AGM contraction $-$, an AGM revision $*$, an EE-ordering \sqsubseteq_{EE} , a GE-ordering \sqsubseteq_{GE} , an RE-ordering \sqsubseteq_{RE} , and an RG-ordering \sqsubseteq_{RG} are *semantically related* iff there is a faithful total preorder \preceq and a semantically related faithful modular weak partial order \leq such that

1. $-$ can be defined in terms of \preceq (and \leq) using (Def \sim from \preceq),
2. $*$ can be defined in terms of \preceq (and \leq) using (Def $*$ from \preceq),
3. \sqsubseteq_{EE} can be defined in terms of \preceq using (Def \sqsubseteq_E from \preceq),
4. \sqsubseteq_{GE} can be defined in terms of \preceq using (Def \sqsubseteq_G from \preceq),
5. \sqsubseteq_{RE} can be defined in terms of \leq using (Def \sqsubseteq_E from \preceq), and
6. \sqsubseteq_{RG} can be defined in terms of \leq using (Def \sqsubseteq_G from \preceq).

□

Figure 5.6 contains a summary of some the results related to faithful layered preorders, and extends the results in figure 3.2 on page 58. Gärdenfors and Makinson [1994, p. 244] advance the view that entrenchment orderings such as their expectation orderings, are more fundamental than structures such as the faithful layered preorders. Their argument is that placing an ordering on sets of states (or worlds or infatoms) is epistemologically more advanced than placing an ordering on beliefs in the form of wffs of L . Accordingly, they see the former as being derived from the latter, and leave the question of how an agent obtains such an ordering on wffs to the field of

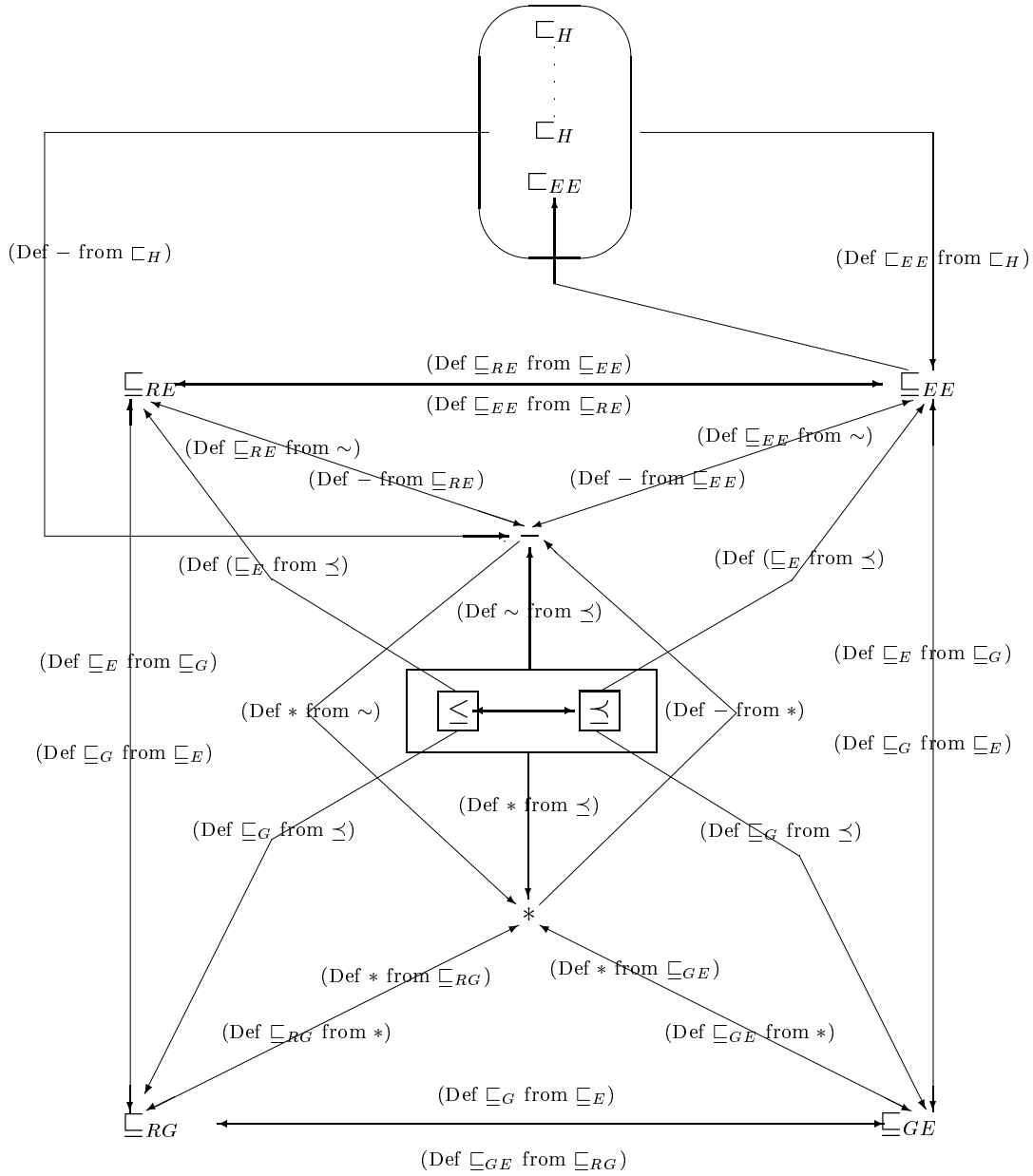


Figure 5.6: The relationship between minimal-equivalent faithful layered preorders, AGM contraction, AGM revision, the EE-orderings, the RE-orderings, the GE-orderings, the RG-orderings, and safe contraction in terms of regular virtually connected hierarchies.

cognitive science. While we agree that some kinds of orderings on wffs can be regarded as more fundamental than orderings on worlds, it is difficult to see that such a view can be applied to orderings on wffs that are as highly structured as the entrenchment orderings encountered in this chapter. In particular, it is difficult to escape the conclusion that the faithful layered preorders are used to *derive* orderings on wffs (in the form of the GRE-orderings), especially when adopting an information-theoretic point of view. Of course, this still leaves open the question of *how* to obtain such orderings on infatoms. One way to achieve this might indeed be in terms of priority orderings on wffs, in the spirit of Nebel's *epistemic relevance* orderings [1990, 1991, 1992]. But such orderings have a completely different character than orderings of entrenchment, since they disregard the logical relationship between wffs.

Finally, in this chapter we have concentrated on suitable properties for entrenchment orderings, but we have paid little attention to how these entrenchment orderings ought to be used. In the next chapter, our attention will be shifted to the latter question. More specifically, we show how the EE-orderings and the RE-orderings can be used to define withdrawals which differ from AGM contraction.

Chapter 6

Withdrawal

Believe it or not.

R.L. Ripley

Title of newspaper column

Although AGM theory contraction occupies a central position in the literature on belief change, there is one aspect about it that has created a fair amount of controversy. It involves the inclusion of (K-6), the postulate known as Recovery. The Recovery postulate is part of the AGM trio's formal expression of the principle of Informational Economy, the idea that an agent should try to keep the loss of information to a minimum.

In this chapter we undertake a detailed investigation of *withdrawals*, the removals obtained when Recovery is dropped from the basic AGM contraction postulates (see section 2.1). We commence with a motivation for the move from contraction to withdrawal by reviewing the main objections levelled at recovery, and then proceed with a description of the withdrawal operations found in [Levi, 1991, 1998, Hansson and Olsson, 1995, Rott and Pagnucco, 1999, Fermé, 1998, Fermé and Rodriguez, 1998]. Along the way, we also present a new addition to the family of withdrawal operations; *systematic withdrawal*. We define systematic withdrawal semantically, in terms of the faithful modular weak partial orders (see definition 5.5.1), and show that it can be characterised by a set of postulates.

In a comparison of withdrawal operations we show that AGM contraction, systematic withdrawal and the *severe withdrawal* of Rott and Pagnucco [1999] are intimately

connected by virtue of their definition in terms of sets of layered faithful preorders. These semantic constructions, together with similar semantic definitions of the EE-orderings (see theorem 3.3.1) and the RE-orderings (see definition 5.5.3), are then used to show that AGM contraction, systematic withdrawal, severe withdrawal, the EE-orderings, and the RE-orderings are all interdefinable; indeed interchangeable. The close connection between these constructions can be traced back to a shared feature. They are all defined in terms of faithful layered preorders; a result that is summarised in figure 6.7 on page 199.

6.1 To recover or not to recover

At a first glance, the Recovery postulate seems to be a reasonable requirement to impose on theory removal. It requires the changes to a belief set K resulting from an α -contraction to be small enough so that an α -expansion will be sufficient to recover all the discarded information. In other words, information is a valuable commodity, and it makes good sense to effect as little change as possible when circumstances dictate that our set of beliefs should be modified. Viewed as such, recovery is a formalisation of the principle of Informational Economy. And while this is clearly a useful principle, it can have undesirable consequences if it is allowed to become the overriding concern. This is the background against which the objections levelled at recovery should be seen.

The Recovery postulate has been criticised by various authors, and for several different reasons.¹ One of the reasons most frequently cited stems from the extension of theory change to *base change*. In base change, the set of wffs on which contractions and revisions are performed, termed the base, need not be a belief set. A base is taken to contain the “basic” beliefs of an agent, with the wffs logically entailed by the base being seen as “derived” beliefs. Under the assumption that only wffs in the current base are allowed to be retained after a (base) contraction — an assumption which underlies most approaches to base change — it is easy to find counterexamples to Recovery.

Example 6.1.1 Let L be the propositional language generated by the two atoms p and q with the valuation semantics (V, \models) , where $V = \{00, 01, 10, 11\}$. Contracting the base $\{p, q\}$ by $p \vee q$ clearly has to result in the empty base. Expanding with $p \vee q$ now

¹Those objections to the Recovery postulate raised by Tennant [1994, 1997] which are valid, are restatements of those in the references cited below.

yields the new base $\{p \vee q\}$, and it is thus not the case that $\{p, q\} \subseteq Cn(\emptyset) + p \vee q$, as (K-6'), applied to bases, would have it. (Recall from section 2.1, that (K-6') is an alternative formulation of the Recovery postulate.) \square

The retention of Recovery on the knowledge level (see page 3) is thus regarded as an obstacle to the acceptance of base change operations.² This argument, found in [Makinson, 1987, Fuhrmann, 1991, Hansson, 1992a, 1993c, 1996, Niederee, 1991], is certainly compelling if one accepts the requirement that a base contraction operation may only result in a new base that is a subset of the original one. But a number of researchers have defined base contraction operations that aren't bound by this restriction, and as a result, the theory contraction operations associated with these base contraction operations *do* satisfy Recovery ([Nebel, 1989, 1990, 1991, 1992, Nayak, 1994a, Meyer et al., 1999a], and chapter 8).³ The rejection of Recovery on these grounds thus boils down to a question of the kind of base contraction one is willing to accept.

A different argument against Recovery, one that operates purely on the theory change level, can be found in [Hansson, 1991, 1992a, 1996, Lindström and Rabinowicz, 1991], and to a certain extent, in [Niederee, 1991] as well. A general formulation of this argument is presented by Lindström and Rabinowicz [1991]. They point out that the following is a derived property of any removal that satisfies the six basic AGM postulates:

$$\text{If } \alpha \in K \text{ and } \alpha \models \beta \text{ then } \alpha \in (K - \beta) + \beta.$$

That is, it is impossible to get rid of a wff α in K by first contracting and then expanding with a wff that is logically weaker than α . This argument is made concrete by the following two convincing counterexamples to Recovery, due to Hansson [1991, 1992a], and also occurring in [Hansson, 1996, 1999].

Example 6.1.2 I read a book about Cleopatra, in which the claim is made that she had a son and a daughter. I subsequently discover that the book is fictional, which leads me to remove my belief that Cleopatra had a child. However, on consulting a history book I discover that Cleopatra indeed had a child, and I thus expand my belief set with this assertion.

²Indeed, in [Alchourrón et al., 1985], where the AGM postulates are phrased so as not to deal exclusively with belief sets, the Recovery postulate, in the form of (K-6'), is taken to hold only for belief sets.

³A theory contraction operation $-$ is *associated* with a base contraction operation \sim iff $Cn(B - \alpha) = Cn(B) - \alpha$.

Let L be a propositional language generated by the two atoms p and q . Let p denote the assertion that Cleopatra had a son, and q the assertion that she had a daughter. Then $K = Cn(p, q)$. The removal of the belief that she had a child is formalised as $K - (p \vee q)$. Since $p, q \in K$, Recovery requires that $K - (p \vee q) + (p \vee q) = K$. So expanding my belief set with the assertion that Cleopatra did, after all, have a child, will ensure that I again entertain the belief that she had a son and the belief that she had a daughter; a conclusion which seems unreasonable in this context. \square

Example 6.1.3 I have reason to believe that George is a mass murderer, and therefore a criminal. Then I receive information which leads me to give up my belief that George is a criminal. Since all mass murderers are criminals, I also have to give up my belief that George is a mass murderer. Then I receive new information which leads me to accept the belief that George is a shoplifter.

To formalise this example, let L be a propositional language generated by the three atoms p , q and r . Let p denote the assertion that George is a mass murderer, q the assertion that George is a criminal, and r the assertion that George is a shoplifter. Clearly, it is appropriate to use a semantics for L in which $p \models q$ and $r \models q$. Letting K denote my initial set of beliefs, we have that $q \in K$. Now, giving up my belief that George is a criminal results in the new set of beliefs $K - q$. By Recovery we then have that $(K - q) + q = K$, and since $r \models q$, $K = (K - q) + q \subseteq (K - q) + r$.

So, since I previously believed George to be a mass murderer, I can't regard him as a shoplifter without again believing that he is a mass murderer as well. \square

These counterexamples strongly suggest that concerns other than the retention of information should also play a role during the removal of beliefs. Such considerations also form the gist of Levi's criticism of Recovery [1991]. He argues that anything other than the use of maxichoice contraction (see section 2.2) already constitutes a radical departure from the requirement that as much information as possible be retained, and takes issue with AGM's restriction of the permissible withdrawals to those that can be defined in terms of the intersection of maxichoice contractions using (Def – from M) (see section 2.2). We discuss these matters in more detail in section 6.3.6.

Niederee [1991] considers a third reason for rejecting the Recovery postulate. This involves an extension to *multiple withdrawal*, in which a withdrawal from a belief set by a *set* of wffs, instead of just a single wff, is performed. He provides some plausible postulates for multiple withdrawal, and shows that multiple withdrawal operations

satisfying these postulates cannot be regarded as “extensions” of (single wff) withdrawal operations that satisfy Recovery (where “extension” is given a precisely defined meaning).

Despite the objections against Recovery, its inclusion yields some desirable properties as well. Firstly, the principal argument against recovery is that removal operations satisfying it, sometimes retain too much information. Yet, as Makinson points out [1987], full meet contraction (see section 2.2), which is a particularly cautious form of withdrawal, satisfies Recovery. In contrast, if Recovery is simply discarded, it permits removals that clearly remove too much information. Witness, for example, the specific withdrawal in which withdrawing any wff $\alpha \in K$, except a logically valid α , yields the set of logically valid wffs. And observe also that it is the inclusion of Recovery which ensures that the attempted removal of a logically valid wff from K results in retaining all of K . Furthermore, although Hansson’s counterexamples show that there are circumstances under which Recovery ought not to hold, it does not address the question of whether there is any situation in which Recovery *should* hold. But such examples do exist, as shown by Nayak [1994a]. Finally, Makinson [1997] points out that counterexamples to Recovery are presented with an implicit assumption of a particular pattern of justification among the beliefs held. He argues that such counterexamples show that Recovery is indeed inappropriate for belief sets augmented with additional structure of some kind, but that Recovery seems to be free of intuitive counterexamples in the idealised situation where a belief set is taken as a “naked” theory, without any extra-logical structure.

In summary then, it seems excessive to insist that *every* withdrawal should satisfy Recovery in order for it to be regarded as rational. Moreover, the advantages of Recovery discussed above are not so much arguments *for* its retention as they are arguments *against* its complete dismissal. It thus seems reasonable to investigate withdrawal operations that do not always satisfy Recovery, but that, nevertheless, retain its desirable features. It is to this task that we now turn.

6.2 Basic withdrawal

In chapter 2 we saw that there is a distinction to be drawn between basic AGM theory contraction and AGM theory contraction (which satisfies the supplementary postulates as well). The latter, which also satisfies the supplementary AGM contraction postu-

lates, is more principled in the sense that it imposes restrictions on the relationship between belief sets resulting from the contraction by different wffs (of a fixed belief set K). And from a semantic point of view, we saw in chapters 3 and 5 that (principled) AGM contraction involves the use of the faithful layered preorders.

We shall see below that a similar distinction holds for withdrawal. We consider two versions of withdrawal that can be considered as basic, at least in the sense that they do not satisfy (K-7) and (K-8). In section 6.3 we switch our attention to more principled forms of withdrawal.

6.2.1 Saturatable withdrawal

For the purposes of constructing appropriate withdrawal operations, it is useful to start with a method for constructing all those withdrawals for which the withdrawal of every logically valid wff leaves the current belief set unaltered. That is, those belief removals satisfying (K-1) to (K-5), together with the following postulate:

(Failure) If $\models \alpha$ then $K - \alpha = K$

Definition 6.2.1 A withdrawal is called *proper* iff it satisfies Failure. □

Proper withdrawal can be characterised with the aid of Levi's saturatable contractions [1991].

Definition 6.2.2 A belief set K' is a *saturatable contraction with respect to K and α* iff $K' \subseteq K$ and $Cn(K' \cup \{\neg\alpha\}) \in L \perp \alpha$. We denote the set of saturatable contractions with respect to K and α by $sc(K, \alpha)$. □

Recall from definition 2.2.1, that $L \perp \alpha$ is the set of maximal subsets of L that do not entail α . So every element K' of $L \perp \alpha$ corresponds to an interpretation u , in the sense that $Th(u) = K'$. (But, in general, the same element of $L \perp \alpha$ might be determined by more than one interpretation — interpretations which are elementarily equivalent, but which might be non-isomorphic.) Note further that there are no saturatable contractions with respect to K and α if $\models \alpha$, and if $\alpha \notin K$, there will only be saturatable contractions in some cases.

To get a feel for the intuition underlying the use of the saturatable contractions, it is instructive to view them semantically. The set of saturatable contractions with respect to a belief set K and a wff α is obtained by adding single models of $\neg\alpha$, together

with any subset of models of α , to the models of K , and then taking the corresponding theory.

Proposition 6.2.3 *Suppose $\alpha \in K$ and $\not\models \alpha$.*

1. *If $x \in M(\neg\alpha)$ and $W \subseteq M(\alpha)$, then $Th(M(K) \cup W \cup \{x\}) \in sc(K, \alpha)$.*
2. *If $K' \in sc(K, \alpha)$, then there is an $x \in M(\neg\alpha)$ and a $W \subseteq M(\alpha)$ such that $K' = Th(M(K) \cup W \cup \{x\})$.*

Proof For the proof of (1), suppose that $x \in M(\neg\alpha)$ and $W \subseteq M(\alpha)$. It suffices to show that $Cn(Th(M(K) \cup W \cup \{x\}) \cup \{\neg\alpha\}) \in L\perp\alpha$. Since $M(K) \cup W \subseteq M(\alpha)$ and $x \in M(\neg\alpha)$, it follows by lemma 1.3.5 that

$$M(Th(M(K) \cup W \cup \{x\})) \cap M(\neg\alpha) = M(Th(M(K) \cup \{x\})) \cap M(\neg\alpha).$$

Since $\alpha \in K$, it follows from lemma 1.3.4 that $Th(x) = Th(M(K) \cup \{x\}) \cap M(\neg\alpha)$. And by proposition 3.2.1, there is an $X \in L\perp\alpha$ such that $Th(x) = X$. So $Cn(Th(M(K) \cup W \cup \{x\}) \cup \{\neg\alpha\}) = Th(x) \in L\perp\alpha$.

For the proof of (2), suppose that $K' \in sc(K, \alpha)$. So $Cn(K' \cup \{\neg\alpha\}) \in L\perp\alpha$. By proposition 3.2.1 there is an $x \in M(\neg\alpha)$ such that

$$Th(\{x\}) = Cn(K' \cup \{\neg\alpha\}) = Th(M(K') \cap M(\neg\alpha)).$$

Now let $W = M(K') \cap M(\alpha)$. We show that $K' = Th(M(K) \cup W \cup \{x\})$. For the left-to-right inclusion, note that $M(K) \subseteq M(K')$ and $x \in M(K')$, and so $M(K) \cup W \cup \{x\} \subseteq M(K')$. For the right-to-left inclusion, pick any $\beta \in Th(M(K) \cup W \cup \{x\})$. So $M(K) \cup W \cup \{x\} \subseteq M(\beta)$, and it suffices to show that $M(K') \setminus (M(K) \cup W \cup \{x\}) \subseteq M(\beta)$. Pick any $y \in M(K') \setminus (M(K) \cup W \cup \{x\})$. By the choice of W , $y \in M(\neg\alpha)$, and thus $y \in M(K' \cup \{\neg\alpha\})$. And since $Th(\{x\}) = Cn(K' \cup \{\neg\alpha\})$, it follows that $y \in M(\beta)$. \square

The removals permitted by Levi are those obtained by taking the intersection of any subset of the saturable contractions with respect to K and α , where $\alpha \in K \setminus Cn(\top)$. (Hansson and Olsson [1995] refer to these removals as *partial meet Levi-contraction operators*.) Semantically, this can be accomplished by using the saturable selection functions.

Definition 6.2.4 A function $ss_K : L \rightarrow \wp U$ is a *saturatable selection function* iff the following holds:

1. If $\alpha \equiv \beta$ then $ss_K(\alpha) = ss_K(\beta)$,
2. if $\models \alpha$ then $ss_K(\alpha) = \emptyset$,
3. if $\alpha \notin K$ then $ss_K(\alpha) \subseteq M(K)$, and
4. if $\not\models \alpha$ and $\alpha \in K$ then $ss_K(\alpha) \cap M(\neg\alpha) \neq \emptyset$.

□

There is a close correspondence between the semantic selection functions of definition 3.2.2, and the saturatable selection functions. As is the case with the semantic selection functions, we add the elements of $ss_K(\alpha)$ to the models of K to obtain the models of $K \sim \alpha$. The difference is that the saturatable selection functions, unlike the semantic selection functions, allow us to include, as models of $K \sim \alpha$, some countermodels of K that are also models of α .

Definition 6.2.5 A removal is a *saturatable withdrawal* iff it can be defined in terms of a saturatable selection function ss_K using (Def \sim from sm_K) (see section 3.2). □

Hansson and Olsson [1995] show that the proper withdrawals are precisely the saturatable withdrawals.⁴

Theorem 6.2.6 A removal \sim is a proper withdrawal iff it is a saturatable withdrawal.

6.2.2 Sensible withdrawal

It is generally acknowledged that the set of *all* withdrawals (and even the set of all proper withdrawals) allows for too much generality. And from the discussion in section 6.1, it seems reasonable to cut down on the set of all proper withdrawals by trying to weaken the Recovery postulate in some appropriate fashion. However, attempts to do so have proved to be quite difficult. For example, Hansson [1991] proposes the following two properties:

⁴Hansson and Olsson's constructions are phrased directly in terms of Levi's saturatable contractions, and not in terms of the saturatable selection functions, but by virtue of proposition 6.2.3, the required correspondence is easily established.

(Relevance) If $\beta \in K \setminus K - \alpha$, then there is an $X \subset K$ such that $K - \alpha \subseteq X$ and $\alpha \notin Cn(X)$, but $\alpha \in Cn(X) + \beta$

(Core-retainment) If $\beta \in K \setminus K - \alpha$, then there is an $X \subset K$ such that $\alpha \notin Cn(X)$, but $\alpha \in Cn(X) + \beta$

Core-retainment is clearly weaker than Relevance, and intuitively, it might seem as if Core-retainment, and perhaps Relevance as well, are weaker than Recovery. However, Hansson shows that they both imply Recovery in the presence of (K-1) to (K-5). Based on these results, Hansson conjectures that “a reasonable contraction operator without the Recovery property does not seem possible”. Indeed, the difficulty in constructing plausible withdrawal operations on belief sets that do not satisfy Recovery has led some researchers to view Recovery not necessarily as a fundamental postulate of theory contraction, but rather as an emerging property [Hansson and Rott, 1995].

Recently, Fermé and Rodriguez [1998] have succeeded in the provision of a weaker version of Recovery.

(Proxy Recovery) If $K \neq K - \alpha$ then there is a $\beta \in K$ such that $\beta \notin K - \alpha$ and $K \subseteq (K - \alpha) + \beta$

It is easily established that Proxy Recovery is a weaker version of Recovery. If Recovery is satisfied, Proxy Recovery holds by taking $\beta = \alpha$.

Definition 6.2.7 A withdrawal is called *sensible* iff it satisfies Failure and Proxy Recovery. \square

It is easily verified that the basic AGM contractions form a strict subset of the sensible withdrawals, which in turn, form a strict subset of the proper withdrawals. (Fermé [1998] provides an example proving the second strict inclusion.)

Fermé and Rodriguez characterise sensible withdrawal in terms of Fermé’s semi-contraction [1998]. The construction of semi-contractions is justified as follows. It is well-known, and easily verified, that if $-$ is a basic AGM contraction, then $\alpha \rightarrow \beta \in K - \alpha$ for every $\beta \in K \setminus K - \alpha$. But in some counterexamples to Recovery, this proves to be undesirable. Consider, for instance, example 6.1.2 again. One way of stating the problem with this example is that the wffs $(p \vee q) \rightarrow p$ and $(p \vee q) \rightarrow q$ are required to be in $K - (p \vee q)$. Fermé’s basic idea is to remove undesirable wffs such as these from the resulting belief set. This is done with the aid of semi-selection functions.

Definition 6.2.8 A *semi-selection* function is a function s from $\wp L$ to L such that $s(X) \in X$ if $X \neq \emptyset$, and $s(\emptyset) = \top$. \square

For any basic AGM contraction $-$, a semi-selection function picks out, for every $K \setminus K - \alpha$, the consequent β of a wff $\alpha \rightarrow \beta$ in $K - \alpha$, such that $\beta \in K \setminus K - \alpha$. This is equivalent to picking out the conjunction of a finite number of consequents β_i of wffs of the form $\alpha \rightarrow \beta_i$ in $K - \alpha$, where every β_i is in $K \setminus K - \alpha$. Semi-contraction is then defined as follows:

(Def \sim from $-$ and s) $K \sim \alpha = (K - \alpha) \cap (K - (\alpha \rightarrow s(K \setminus K - \alpha)))$

Definition 6.2.9 A removal function is a *semi-contraction* iff it can be defined in terms of a basic AGM contraction and a semi-selection function using (Def \sim from $-$ and s). \square

The following representation theorem of Fermé and Rodriguez [1998] establishes the relationship between sensible withdrawal and semi-contraction.

Theorem 6.2.10 A removal is a sensible withdrawal iff it is a semi-contraction.

While sensible withdrawal does indeed provide us with a withdrawal operation that is more permissive than basic AGM contraction, but not as permissive as proper withdrawal, there are indications that it is not principled enough to be regarded as an appropriate form of withdrawal. The following example shows that sensible withdrawal does not always satisfy the supplementary postulates, (K-7) and (K-8); not even when we restrict ourselves to the sensible withdrawals defined in terms of AGM contractions (which *do* satisfy the supplementary postulates).

Example 6.2.11 Let L be the propositional language generated by the three atoms p , q and r , and let (V, \models) be the valuation semantics for L where

$$V = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$

Let $K = Cn\{p, q, r\}$ and let \preceq be the faithful total preorder defined as follows:

$$x \preceq y \text{ iff } \begin{cases} y \in V \text{ if } x = 111, \\ y \in \{000, 001, 010, 011, 100, 101, 110\} \text{ if } x \in \{011, 101\}, \\ y \in \{000, 001, 010, 100, 110\} \text{ if } x = 110, \\ y \in \{000, 001, 010, 100\} \text{ if } x = 100, \text{ and} \\ y \in \{000, 001, 010\} \text{ if } x \in \{000, 001, 010\}. \end{cases}$$

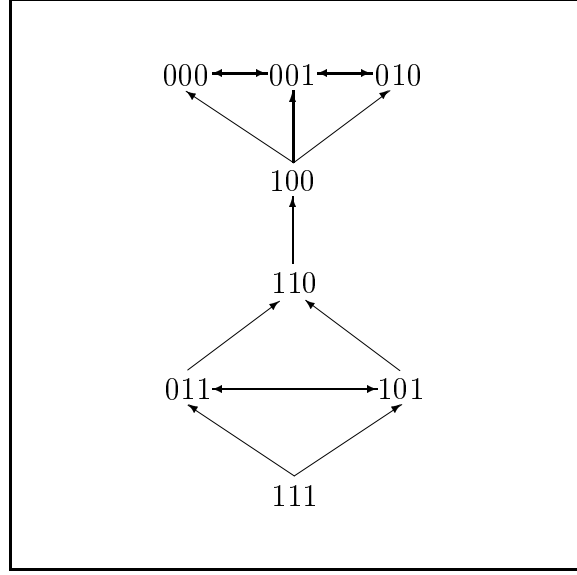


Figure 6.1: A graphical representation of the faithful total preorder \preceq used in example 6.2.11. The ordering is obtained from the reflexive transitive closure of the relation determined by the arrows.

Figure 6.1 contains a graphical representation of \preceq . Let $-$ be the AGM contraction defined in terms of \preceq using (Def \sim from \preceq), and let s be any semi-selection function such that

$$\begin{aligned}
 s(K \setminus K - (q \vee r)) &= q \vee r, \\
 s(K \setminus K - (\neg q \vee r)) &= \neg q \vee r, \\
 s(K \setminus K - r) &= q \leftrightarrow r, \\
 s(K \setminus K - (p \rightarrow q)) &= p \wedge q, \text{ and} \\
 s(K \setminus K - q) &= q.
 \end{aligned}$$

It is readily verified that such an s exists. Now, let \sim be the semi-contraction defined in terms of $-$ and s using (Def \sim from $-$ and s). By theorem 6.2.10, \sim is a sensible withdrawal.

To show that \sim violates (K-7), take $q \vee r$ as α , $\neg q \vee r$ as β , and observe that $K \sim (q \vee r) = Cn(p \wedge (q \leftrightarrow r))$, $K \sim (\neg q \vee r) = Cn(p \wedge q)$, and $K \sim ((q \vee r) \wedge (\neg q \vee r)) = K \sim r = Cn(p \wedge (q \vee r))$. To show that \sim violates (K-8), take q as α , $p \rightarrow q$ as β , and observe that $K \sim (p \rightarrow q) = Cn((p \vee q) \wedge r)$, and that

$$K \sim (q \wedge (p \rightarrow q)) = K \sim q = Cn(p \wedge r). \quad \square$$

The failure of (K-7) and (K-8) can be traced back to the semi-selection functions and the undesirable amount of freedom they allow in choosing a wff in $K \setminus K - \alpha$.

6.3 Principled withdrawal

In the light of the failure of sensible withdrawal to satisfy (K-7) and (K-8), the challenge before us is to define a type of withdrawal that is truly principled in nature, like AGM contraction, but without the requirement that Recovery should hold. To obtain such an approach to withdrawal, it is necessary to take a closer look at the intuition associated with AGM contraction.

As we have seen, the inclusion of the Recovery postulate in the AGM framework is justified by an appeal to the principle of Informational Economy [Gärdenfors, 1988]. When epistemic states are viewed as belief sets, this view dictates that informational economy should be measured in terms of set-inclusion, thus providing a restatement of the principle of Conservatism. If the principle of Informational Economy had been the overriding concern, it would have implied that the belief set resulting from an α -contraction of K should be a maximal subset of K that does not imply α ; that is, an α -remainder (see definition 2.2.1). But it is easily seen that this involves a restriction to maxichoice contraction (see page 22), a special case of AGM contraction which Alchourrón and Makinson [1982] have shown to be too strong for a general account of theory contraction.

Since AGM contraction is more than just maxichoice contraction, it follows that the principle of Informational Economy is not the only requirement in question, but rather one of several equally important guidelines. In particular, as Rott and Pagnucco [1999] argue in their excellent survey of withdrawal, the respective roles of the principles of Indifference and Preference in the construction of AGM contractions are as important as that of the principle of Informational Economy. We shall see below that in defining AGM contraction, the principle of Informational Economy has, to some degree, already given way to the principle of Indifference. It is our contention that, for a description of principled withdrawal, it is necessary for this process to take its full course. That is, we propose that both the principles of Indifference and Preference should take strict precedence over the principle of Informational Economy. We adopt an information-theoretic point of view, and use the faithful layered preorders on the infatoms of L as

the machinery for constructing withdrawal operations. The principles of Indifference, Preference, and Informational Economy are then applied, in a consistent manner, to different subsets of the faithful layered preorders, yielding different forms of principled withdrawal.

Let us first consider, in detail, the way in which these three principles are combined to obtain AGM contraction. In this case, the application of the principle of Informational Economy is twofold. Its influence is felt in the requirement that some mechanism should be used for distinguishing between the level of entrenchment of infatoms. This requirement is implemented by the use of a faithful total preorder. Secondly, the principle of Informational Economy restricts the application of the remaining two principles to content bits of α during an α -contraction. (This is where it still takes precedence over the principle of Preference and, to some extent, over the principle of Indifference.) The principle of Preference then ensures that any content bit i of α which is regarded as at most as entrenched as a content bit j of α , will receive at most as much consideration for removal from K as j . Consequently, only the worst content bits of α are considered for removal. And finally, since the worst content bits of α are all seen as equally entrenched, the principle of Indifference ensures that they will all be removed from K . So, in this sense at least, the principle of Indifference holds sway over the principle of Informational Economy.

It is our view that the role of the principle of Informational Economy should be reviewed in order for both the principles of Indifference and Preference to take complete precedence over it. In this view, its application only results in the use of the faithful layered preorders to distinguish between the level of entrenchment of infatoms. Guided by the two remaining principles, the set of infatoms removed from K then contains all the infatoms that are at most as entrenched as the worst content bits of α . We shall see that the application of these three principles in the manner described above, leads to the development of a number of different forms of principled withdrawal.

6.3.1 Severe withdrawal

Rott and Pagnucco [1999] use the faithful total preorders to define the set of severe withdrawals.⁵ Recall from section 5.1 that the downset of a wff α is defined in terms

⁵Actually, they use Grove's systems of spheres, but it is easily extendable to the slightly more general case that we consider.

of a faithful preorder using (Def ∇_{\preceq} from \preceq). The downset of α contains all the interpretations that are at least as low down in the ordering as the minimal models of α . Downsets are used to define severe withdrawal as follows:

(Def \sim from ∇_{\preceq}) $K \sim \alpha = Th(M(K) \cup \nabla_{\preceq}(\neg\alpha))$

Definition 6.3.1 A removal is a *severe withdrawal* iff it is defined in terms of a faithful total preorder using (Def \sim from ∇_{\preceq}). \square

Viewed information-theoretically, it should be apparent that (Def \sim from ∇_{\preceq}) is an application, in terms of the faithful total preorders, of the principles of Indifference, Preference, and Informational Economy in the manner described above.

Rott and Pagnucco show that severe withdrawal is characterised by the following set of postulates.⁶

(K $\ddot{-}$ 1) $K \ddot{-} \alpha = Cn(K \ddot{-} \alpha)$

(K $\ddot{-}$ 2) $K \ddot{-} \alpha \subseteq K$

(K $\ddot{-}$ 3) If $\alpha \notin K$ then $K \ddot{-} \alpha = K$

(K $\ddot{-}$ 4) If $\not\models \alpha$ then $\alpha \notin K \ddot{-} \alpha$

(K $\ddot{-}$ 5) If $\alpha \equiv \beta$ then $K \ddot{-} \alpha = K \ddot{-} \beta$

(K $\ddot{-}$ 6) If $\models \alpha$ then $K \ddot{-} \alpha = K$

(K $\ddot{-}$ 7) If $\not\models \alpha$ then $K \ddot{-} \alpha \subseteq K \ddot{-}(\alpha \wedge \beta)$

(K $\ddot{-}$ 8) If $\beta \notin K \ddot{-}(\alpha \wedge \beta)$ then $K \ddot{-}(\alpha \wedge \beta) \subseteq K \ddot{-} \beta$

Theorem 6.3.2 [Rott and Pagnucco, 1999] A removal $\ddot{-}$ is a severe withdrawal iff it satisfies (K $\ddot{-}$ 1) to (K $\ddot{-}$ 8).

The postulates for severe withdrawal differ from those for AGM contraction only on the sixth and seventh postulates; the remaining ones are identical to their AGM contraction counterparts. (K $\ddot{-}$ 6), which replaces Recovery, is the postulate we have referred to as Failure. (K $\ddot{-}$ 7) is a much stronger requirement than the corresponding AGM

⁶Pagnucco [1996] originally gave a different characterisation of severe withdrawal.

contraction postulate (K-7). It is a kind of monotonicity property, requiring that the removal of weaker wffs should always result in smaller belief sets. Rott and Pagnucco regard this as an intuitively plausible postulate which follows from the application of the principles of Indifference and Preference. In section 6.3.5, we argue against the inclusion of this postulate, showing that it has some undesirable consequences, and that (K-7) is a consequence of the principles of Indifference and Preference only when they are applied to the faithful *total* preorders.

6.3.2 Systematic withdrawal

In this section we introduce a set of withdrawals that are closely related to the severe withdrawals. Their construction is based on an application of the principles of Indifference, Preference and Informational Economy in a manner identical to that used in the construction of severe withdrawal. The only difference is that they are obtained using the faithful modular weak partial orders, instead of the faithful total preorders.

Definition 6.3.3 A belief removal \div is a *systematic withdrawal* iff it is defined in terms of a faithful modular weak partial order using (Def \sim from ∇_{\preceq}). \square

The difference between systematic withdrawal and severe withdrawal lies in the difference between the downset (see definition 5.1.1) of a wff α obtained from a total preorder and that obtained from a modular weak partial order. In the latter case, the downset consists of the minimal models of α as well as all the interpretations strictly below them (which are all, of course, countermodels of α). The former case includes all the interpretations mentioned above, as well as the countermodels of α on the same level as the minimal models of α . In section 6.3.5 we shall see that this seemingly minor technical difference accounts for some fundamental differences between these two forms of principled withdrawal. For the moment, we provide a characterisation of systematic withdrawal in terms of a set of postulates.

$$(K\div 1) \quad K \div \alpha = Cn(K \div \alpha)$$

$$(K\div 2) \quad K \div \alpha \subseteq K$$

$$(K\div 3) \quad \text{If } \alpha \notin K \text{ then } K \div \alpha = K$$

$$(K\div 4) \quad \text{If } \not\models \alpha \text{ then } \alpha \notin K \div \alpha$$

(K÷5) If $\alpha \equiv \beta$ then $K \div \alpha = K \div \beta$

(K÷6) If $\models \alpha$ then $K \div \alpha = K$

(K÷7) If $\gamma \in K \div (\alpha \wedge \gamma)$ then $\gamma \in K \div (\alpha \wedge \beta \wedge \gamma)$

(K÷8) If $\beta \notin K \div (\alpha \wedge \beta)$ then $K \div (\alpha \wedge \beta) \subseteq K \div \beta$

(K÷9) If $\alpha \in K$, $\alpha \vee \beta \in K \div \alpha$ and $\beta \notin K \div \alpha$ then $\alpha \in K \div (\alpha \wedge \beta)$

(K÷10) If $\not\models \alpha$ and $\beta \in K \div \alpha$ then $\alpha \notin K \div (\alpha \wedge \beta)$

Theorem 6.3.4 *A removal \div is a systematic withdrawal iff it satisfies (K÷1) to (K÷10).*

Proof The left-to-right direction follows from proposition B.1.2 in appendix B. For the converse, suppose that \sim satisfies (K÷1) to (K÷10). Now define $\ddot{\sim}$ in terms of \sim using (Def $\ddot{\sim}$ from \sim) on page 160. By lemma B.1.4 in appendix B, $\ddot{\sim}$ is a severe withdrawal. So there is a faithful total preorder \preceq from which $\ddot{\sim}$ can be obtained using (Def \sim from ∇_{\preceq}). Let \leq be the faithful modular weak partial order which is semantically related to \preceq . By proposition 6.3.20, the systematic withdrawal \div obtained from \leq using (Def \sim from ∇_{\preceq}) can also be defined in terms of $\ddot{\sim}$ using (Def \div from $\ddot{\sim}$) on page 162. And by lemma B.1.5 in appendix B, \div is identical to \sim . \square

The first five postulates coincide with the first five AGM contraction postulates, and the first six coincide with the first six postulates for severe withdrawal. (K÷7) is a much weaker version of (K $\ddot{\sim}$ 7). If a wff γ is entrenched enough in the belief set K so that it is retained when at least one of γ or α has to be discarded, then it should also be retained when at least one of γ or any wff logically stronger than α has to be discarded. (K÷8) is identical to (K–8) and (K $\ddot{\sim}$ 8). (K÷9) and (K÷10) both introduce more restrictions on the relationship between withdrawals by different wffs. (K÷9) gives conditions under which a wff α should be retained and (K÷10) gives conditions under which α should be discarded.

6.3.3 Revision-equivalence

With the definition of severe withdrawal and systematic withdrawal, we now have, together with AGM contraction, three types of principled withdrawal at our disposal

which, as it turns out, are very closely related. For a proper comparison of this relationship, it is instructive to commence with the description of a feature which Makinson [1987] refers to as revision-equivalence.

Definition 6.3.5 Two withdrawals \sim and \approx are *revision-equivalent* iff $(K \sim \neg\alpha) + \alpha = (K \approx \neg\alpha) + \alpha$. \square

In other words, two withdrawals are revision-equivalent iff the revisions they define in terms of the Levi identity (Def $*$ from \sim), are identical. From Makinson [1987] we obtain the following results concerning the revision-equivalence of basic AGM contraction and (basic) withdrawal.

Theorem 6.3.6 1. *A revision-equivalent class of withdrawals contains a unique basic AGM contraction.*

2. *The basic AGM contraction $-$ is the maximal element in the equivalence class $[-]$ of withdrawals that are revision-equivalent to $-$. That is, for every \sim in $[-]$, $K \sim \alpha \subseteq K - \alpha$ for every $\alpha \in L$.*

To bring severe withdrawal into the picture, we need to restrict ourselves to the revision-equivalent classes which contain the AGM contractions.

Definition 6.3.7 A revision-equivalent class is called *principled* iff it contains an AGM contraction. \square

Note that a withdrawal in a principled revision-equivalence class need not satisfy (K-7) and (K-8). A case in point is the sensible withdrawal in example 6.2.11.

Rott and Pagnucco [1999] provide the following results.

Theorem 6.3.8 1. *Every principled revision-equivalent class contains a unique severe withdrawal.*

2. *The severe withdrawal $\ddot{-}$ is the minimal element in the (principled) equivalence class $[\ddot{-}]$ of withdrawals that are revision-equivalent to $\ddot{-}$ and that satisfy (K-8). That is, for every \sim in $[\ddot{-}]$ that satisfies (K-8), $K \ddot{-} \alpha \subseteq K \sim \alpha$ for every $\alpha \in L$.⁷*

It should come as no surprise that the revision-equivalence of an AGM contraction and a severe withdrawal is closely tied to their semantic definitions in terms of faithful total preorders.

⁷This is a result derived from Observation 7 in [Rott and Pagnucco, 1999].

Definition 6.3.9 An AGM contraction and a severe withdrawal are *semantically related* iff they can be defined in terms of the same faithful total preorder using (Def \sim from \preceq) and (Def \sim from ∇_{\preceq}) respectively. \square

Theorem 6.3.10 *An AGM contraction and a severe withdrawal are semantically related iff they are revision-equivalent.*

Proof Follows from lemma 1.3.5 and the fact that every principled revision-equivalence class contains a unique AGM contraction and a unique severe withdrawal. \square

In fact, it is easily established that the very notion of principled revision-equivalence hinges on the use of minimal-equivalent faithful layered preorders (see definition 3.3.6).

Proposition 6.3.11 *Suppose \sim and \approx are two withdrawals which are in the same principled revision-equivalence class, and let $*$ be the AGM revision obtained in terms of \sim and \approx using (Def $*$ from \sim). Furthermore, let \preceq be any faithful layered preorder in terms of which $*$ is defined using (Def $*$ from \preceq). Then, for every $\alpha \in K \setminus Cn(\top)$, there is a $W_{\alpha}^{\sim} \subseteq M(\alpha)$ and a $W_{\alpha}^{\approx} \subseteq M(\alpha)$ such that*

$$\begin{aligned} K \sim \alpha &= Th(M(K) \cup Min_{\preceq}(\neg\alpha) \cup W_{\alpha}^{\sim}), \text{ and} \\ K \approx \alpha &= Th(M(K) \cup Min_{\preceq}(\neg\alpha) \cup W_{\alpha}^{\approx}). \end{aligned}$$

Proof Follows from lemma 1.3.5. \square

The significance of proposition 6.3.11 is that it enables us to regard a set of minimal-equivalent faithful layered preorders as the basis for obtaining a principled revision-equivalent class of withdrawals, and allows us to see every withdrawal in a principled revision-equivalence class as “independent” of the other members in the class. For example, Rott and Pagnucco show that the smallest withdrawal $\dot{-}$ in a principled revision-equivalent class $[\dot{-}]$ can be defined in terms of the severe withdrawal in $[\dot{-}]$ as follows:

$$(\text{Def } \dot{-} \text{ from } \ddot{-}) \quad K \dot{-} \alpha = \begin{cases} Cn(\alpha) \cap K \ddot{-} \alpha & \text{if } \alpha \in K \setminus Cn(\top), \\ K & \text{otherwise} \end{cases}$$

But $\dot{-}$ can also be defined, “independently” of $\ddot{-}$, in terms of a faithful total preorder \preceq as follows:

$$(\mathbf{Def} \dashv\!\!\!\vdash \mathbf{from} \preceq) K \dashv\!\!\!\vdash \alpha = \begin{cases} Th(M(K) \cup Min_{\preceq}(\neg\alpha) \cup M(\alpha)) \\ \text{if } \alpha \in K \setminus Cn(\top), \\ K \text{ otherwise} \end{cases}$$

Proposition 6.3.12 *Let $\dashv\!\!\!\vdash$ be the severe withdrawal defined in terms of a faithful total preorder \preceq using (Def \sim from ∇_{\preceq}). The withdrawal $\dashv\!\!\!\vdash$ defined in terms of $\dashv\!\!\!\vdash$ using (Def $\dashv\!\!\!\vdash$ from $\dashv\!\!\!\vdash$) can also be defined in terms of \preceq using (Def $\dashv\!\!\!\vdash$ from \preceq).*

Proof We only consider the case where $\alpha \in K \setminus Cn(\top)$. Then

$$\begin{aligned} & Cn(\alpha) \cap K \dashv\!\!\!\vdash \alpha \\ = & Th(M(K \dashv\!\!\!\vdash \alpha) \cup M(\alpha)) \\ = & Th(M(K) \cup \nabla_{\preceq}(\neg\alpha) \cup M(\alpha)) \text{ by (Def } \sim \text{ from } \nabla_{\preceq}) \\ = & Th(M(K) \cup Min_{\preceq}(\neg\alpha) \cup M(\alpha)) \text{ from (Def } \nabla_{\preceq}). \end{aligned}$$

□

The withdrawal $\dashv\!\!\!\vdash$ defined in terms of a faithful total preorder using (Def $\dashv\!\!\!\vdash$ from \preceq) is in gross violation of the principles of Indifference, Preference and Informational Economy. From an information-theoretic point of view, it removes *all* the content bits of $\neg\alpha$ from K during a withdrawal of α , regardless of how entrenched they are. As such, it is not an appropriate candidate for principled withdrawal. It is most likely examples such as these which prompted Lindström and Rabinowicz [1991] to advance the thesis that any reasonable withdrawal should lie somewhere between AGM contraction and severe withdrawal. To be more precise, in a principled revision-equivalence class containing the AGM contraction $-$ and the severe withdrawal $\dashv\!\!\!\vdash$, we should regard as reasonable, only those withdrawals \sim for which $K \dashv\!\!\!\vdash \alpha \subseteq K \sim \alpha \subseteq K - \alpha$ for every $\alpha \in L$. Following a suggestion by Rott [1992a, 1995], we refer to this proposal as the *LR interpolation thesis*.

Definition 6.3.13 A withdrawal is *reasonable* iff it satisfies the LR interpolation thesis. □

Note that being a reasonable withdrawal is not a guarantee of principled behaviour. Some such withdrawals, such as the sensible withdrawal in example 6.2.11, do not even satisfy (K-7) and (K-8).⁸ From an information-theoretic point of view, the LR

⁸It is easily verified that the sensible withdrawal in this example is indeed reasonable.

interpolation thesis requires an α -withdrawal to be effected by removing from K , any subset of the content bits of $\neg\alpha$ that are at most as entrenched as the least entrenched content bits of α , together with these least entrenched content bits of α . So, while it does not guarantee an adherence to the principles of Preference and Indifference with regard to $C(\neg\alpha)$ (the content bits of $\neg\alpha$), it ensures the satisfaction of these two principles in terms of $C(\alpha)$ (the content bits of α) and it goes some way towards satisfying these principles when comparing elements of $C(\alpha)$ and $C(\neg\alpha)$.

We are now in a position to bring systematic withdrawal into the picture as well. It is perhaps to be expected that every principled revision-equivalence class contains a unique systematic withdrawal. And this is indeed the case, as the next proposition shows.

Proposition 6.3.14 *Every principled revision-equivalence class contains a unique systematic withdrawal.*

Proof Pick any principled revision-equivalence class. By theorem 6.3.6, it contains a unique AGM contraction $-$ which, by proposition 5.7.3, can be defined in terms of a faithful modular weak partial order \leq . By lemma 1.3.5, the systematic withdrawal \div , defined in terms of \leq using (Def \sim from \leq), is revision-equivalent to $-$. Now assume there is a different systematic withdrawal \sim in this revision-equivalence class. By theorem 6.3.4, it can be defined in terms of a faithful modular weak partial order \preceq using (Def \sim from ∇_{\preceq}), where \preceq is *not* minimal-equivalent to \leq . And then \preceq defines an AGM contraction $\check{-}$ in terms of (Def \sim from \preceq) which, though revision-equivalent to \sim , differs from $-$. But this contradicts the uniqueness of $-$ in the given revision-equivalence class. \square

It is easily seen that systematic withdrawal is also reasonable (that is, it satisfies the LR interpolation thesis).

Proposition 6.3.15 *Every systematic withdrawal belongs to a principled revision-equivalence class, and is reasonable.*

Proof Consider any systematic withdrawal \div . By definition, there is a faithful modular weak partial order \leq in terms of which \div is defined using (Def \sim from ∇_{\leq}). By lemma 1.3.5, the AGM contraction defined in terms of \leq using (Def \sim from \preceq) is revision-equivalent to \div , and it thus follows that \div belongs to a principled revision-equivalence class. Furthermore, from theorem 6.3.6, $K \div \alpha \subseteq K - \alpha$ for every $\alpha \in L$.

Now consider the faithful total preorder \preceq obtained in terms of \leq using (Def \preceq from \leq), and let $\ddot{-}$ be the severe withdrawal defined in terms of \preceq using (Def \sim from ∇_{\preceq}). Then $K \ddot{-} \alpha \subseteq K \div \alpha$ for every $\alpha \in L$, and by lemma 1.3.5, $\ddot{-}$ is revision-equivalent to \div . So \div satisfies the LR-interpolation thesis; i.e. it is reasonable. \square

In the context of revision-equivalence, the relationship between AGM contraction, systematic withdrawal, severe withdrawal, and the faithful layered preorders defining these different forms of principled withdrawal, is summarised in the following corollary.

Corollary 6.3.16 *Consider a principled revision-equivalence class \mathcal{R} of withdrawals.*

1. There is a minimal-equivalence class \mathcal{M} of faithful layered preorders such that, for every faithful layered preorder \preceq in \mathcal{M} and every withdrawal \sim in \mathcal{R} , $K \sim \alpha = Th(M(K) \cup Min_{\preceq}(\neg\alpha) \cup W_{\alpha}^{\sim})$, where $W_{\alpha}^{\sim} \subseteq M(\alpha)$.
2. \mathcal{R} contains a unique AGM contraction $-$, a unique systematic withdrawal \div that is also reasonable, and a unique severe withdrawal $\ddot{-}$.
3. For every withdrawal \sim in \mathcal{R} , $K \sim \alpha \subseteq K - \alpha$ for every $\alpha \in L$.
4. For every withdrawal \sim in \mathcal{R} which satisfies (K-8), $K \ddot{-} \alpha \subseteq K \sim \alpha$ for every $\alpha \in L$.
5. The AGM contraction $-$ can be defined in terms of every faithful layered preorder \preceq in \mathcal{M} , using (Def \sim from \preceq).
6. The systematic withdrawal \div can be defined in terms of every faithful modular weak partial order \leq in \mathcal{M} , using (Def \sim from ∇_{\leq}).
7. The severe withdrawal $\ddot{-}$ can be defined in terms of every faithful total preorder \preceq in \mathcal{M} , using (Def \sim from ∇_{\preceq}).

Proof Follows from proposition 6.3.11, theorems 6.3.6 and 6.3.8, propositions 6.3.14, 6.3.15, and 5.7.3, theorem 6.3.4, and theorem 6.3.2. \square

6.3.4 Reasonable withdrawal

This section is devoted to an investigation of the relationship between various reasonable withdrawals, with particular emphasis on AGM contraction, systematic withdrawal and severe withdrawal. We have seen that AGM contraction and severe withdrawal both occupy special positions in the revision-equivalence classes. The former provides an upper bound for reasonable withdrawal, and the latter a lower bound. As a result both can be defined in terms of the remaining reasonable withdrawals. In particular, $-$ can be obtained from any revision-equivalent reasonable withdrawal \sim as follows

(Def – from \sim) $K - \alpha = K \cap ((K \sim \alpha) + \neg\alpha)$

And $\ddot{-}$ can be obtained from any revision-equivalent reasonable withdrawal \sim in one of two ways:⁹

(Def $\ddot{-}$ from \sim) $\beta \in K \ddot{-} \alpha$ iff $\begin{cases} \beta \in K \sim (\alpha \wedge \beta) \text{ if } \not\models \alpha, \\ \beta \in K \text{ otherwise} \end{cases}$

(Def $\ddot{-}$ from \sim (v2)) $K \ddot{-} \alpha = \begin{cases} \bigcap \{K \sim (\alpha \wedge \beta) \mid \beta \in L\} \text{ if } \not\models \alpha, \\ K \text{ otherwise} \end{cases}$

Proposition 6.3.17 *Let $-$ and $\ddot{-}$ be an AGM contraction and a severe withdrawal respectively, that are revision-equivalent. Suppose that \sim is a reasonable withdrawal that is revision-equivalent to $\ddot{-}$ (and $-$). Then*

1. $-$ can be defined in terms of \sim using (Def – from \sim),
2. $\ddot{-}$ can be defined in terms of \sim using (Def $\ddot{-}$ from \sim), and
3. $\ddot{-}$ can be defined in terms of \sim using (Def $\ddot{-}$ from \sim (v2)).

Proof Let \preceq be a faithful total preorder in terms of which $-$ is defined using (Def \sim from \preceq). By corollary 6.3.16, $\ddot{-}$ can be defined in terms of \preceq using (Def \sim from ∇_{\preceq}). Since \sim is reasonable, and therefore revision-equivalent to $-$, there is, by lemma 1.3.5, a $W_\alpha \subseteq M(\alpha)$ such that $K \sim \alpha = Th(M(K) \cup W_\alpha \cup Min_{\preceq}(\neg\alpha))$, for every $\alpha \in K \setminus Cn(\top)$. We only consider the cases where $\not\models \alpha$.

⁹Since (Def $\ddot{-}$ from \sim) and (Def $\ddot{-}$ from \sim (v2)) define the same severe withdrawal when applied to any reasonable withdrawal, any further results involving (Def $\ddot{-}$ from \sim) should be seen as results involving (Def $\ddot{-}$ from \sim (v2)) as well.

1. Follows from lemma 1.3.5.
2. If $\beta \notin K \sim (\alpha \wedge \beta)$ then $\beta \notin K \ddot{-}(\alpha \wedge \beta)$, since \sim is reasonable and revision-equivalent to $\ddot{-}$. So there is a $y \in M(K) \cup \nabla_{\preceq}(\neg(\alpha \wedge \beta))$ such that $z \in M(\neg\beta)$, and therefore $y \in \text{Min}_{\preceq}(\neg(\alpha \wedge \beta))$. Therefore $x \not\prec y$ for every $x \in \text{Min}_{\preceq}(\neg\alpha)$, and thus $\beta \notin K \ddot{-}\alpha$. Conversely, if $\beta \notin K \ddot{-}\alpha$ then $y \in M(\neg\beta)$ for some $y \in M(K) \cup \nabla_{\preceq}(\neg\alpha)$, and there is thus an $x \in \text{Min}_{\preceq}(\neg(\alpha \wedge \beta))$ such that $x \in M(\neg\beta)$. Therefore $\beta \notin K \sim (\alpha \wedge \beta)$.
3. If $\gamma \notin \bigcap \{K \sim (\alpha \wedge \beta) \mid \beta \in L\}$ then there is a $\beta \in L$ such that $\gamma \notin K \sim (\alpha \wedge \beta)$. And then $\gamma \notin K \ddot{-}(\alpha \wedge \beta)$, since \sim is reasonable and revision-equivalent to $\ddot{-}$. So there is a $z \in M(K) \cup \nabla_{\preceq}(\neg(\alpha \wedge \beta))$ such that $M(\neg\gamma)$. But then $\gamma \notin K \ddot{-}\alpha$, since $y \not\prec x$ for every $y \in \text{Min}_{\preceq}(\neg\alpha)$ and every $x \in \text{Min}_{\preceq}(\neg(\alpha \wedge \beta))$. Conversely, if $\gamma \notin K \ddot{-}\alpha$ then $\gamma \notin K \ddot{-}(\alpha \wedge \gamma)$ by part (2), from which the required result follows.

□

And as a corollary of proposition 6.3.17, the identities (Def $-$ from \sim) and (Def $\ddot{-}$ from \sim) are interchangeable when restricted to AGM contraction and severe withdrawal.¹⁰ That is, starting with an AGM contraction or a severe withdrawal, and applying these two identities in the appropriate sequence, brings us back to where we started.

Definition 6.3.18 An AGM contraction $-$, a systematic withdrawal \div , and a severe withdrawal $\ddot{-}$ are *semantically related* iff there is a faithful total preorder \preceq and a semantically related faithful modular weak partial order \leq such that

1. $-$ is defined in terms of \preceq (and \leq) using (Def \sim from \preceq),
2. \div is defined in terms of \leq using (Def \sim from ∇_{\preceq}), and
3. $\ddot{-}$ is defined in terms of \preceq using (Def \sim from ∇_{\preceq}).

□

¹⁰Part (1) of proposition 6.3.17 can be traced back to [Makinson, 1987]. Also, proposition 6.3.17, when restricted to AGM contraction and severe withdrawal, and corollary 6.3.19, albeit in a slightly different guise, can be found in [Rott and Pagnucco, 1999].

Corollary 6.3.19 *If an AGM contraction and a severe withdrawal are semantically related, then they can also be defined in terms of each other using (Def $-$ from \sim) and (Def $\ddot{-}$ from \sim).*

Proof Follows from proposition 6.3.17 and corollary 6.3.16. \square

Since systematic withdrawal is reasonable, it follows from proposition 6.3.17 that every systematic withdrawal \div defines a revision-equivalent AGM contraction $-$ using (Def $-$ from \sim), and a revision-equivalent severe withdrawal $\ddot{-}$ using (Def $\ddot{-}$ from \sim). Being reasonable, \div lies somewhere between $-$ and $\ddot{-}$, so to speak. (In fact, it lies much “closer” to severe withdrawal, in terms of set-inclusion.) Nevertheless, it is possible to define systematic withdrawal in terms of both AGM contraction and severe withdrawal. In particular, \div can be defined in terms of $-$ as follows:

$$(\text{Def } \div \text{ from } -) \quad \beta \in K \div \alpha \text{ iff } \begin{cases} \alpha \vee \beta \in K - (\alpha \wedge \beta) \text{ and } \alpha \notin K - (\alpha \wedge \beta) \\ \text{if } \not\models \alpha, \not\models \beta, \alpha \in K, \\ \beta \in K \text{ otherwise} \end{cases}$$

And \div can be defined in terms of $\ddot{-}$ as follows:

$$(\text{Def } \div \text{ from } \ddot{-}) \quad \beta \in K \div \alpha \text{ iff } \begin{cases} \alpha \vee \beta \in K \ddot{-} \alpha \text{ and } \alpha \notin K \ddot{-} \beta \\ \text{if } \not\models \alpha, \not\models \beta \text{ and } \alpha \in K, \\ \beta \in K \text{ otherwise} \end{cases}$$

Proposition 6.3.20 *Let $-$ be an AGM contraction, let \div be a systematic withdrawal, and let $\ddot{-}$ be a severe withdrawal. Suppose that $-$, \div and $\ddot{-}$ are semantically related.*

1. \div can also be defined in terms of $-$ using (Def \div from $-$).
2. \div can also be defined in terms of $\ddot{-}$ using (Def \div from $\ddot{-}$).

Proof Let \preceq be a faithful total preorder in terms of which $-$ and $\ddot{-}$ are defined using (Def \sim from \preceq) and (Def \sim from ∇_{\preceq}) respectively, and let \leq be the faithful modular weak partial order that is semantically related to \preceq . We only consider the case where $\not\models \alpha, \not\models \beta$ and $\alpha \in K$.

1. Suppose that $\beta \in K \div \alpha$. Then $\nabla_{\leq}(\neg\alpha) \subseteq M(\beta)$ and so $\text{Min}_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha \vee \beta)$ and $\text{Min}_{\leq}(\neg\alpha) = \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$. Therefore $\alpha \vee \beta \in K - (\alpha \wedge \beta)$ and $\alpha \notin K - (\alpha \wedge \beta)$. Conversely, suppose that $\alpha \vee \beta \in K - (\alpha \wedge \beta)$ and $\alpha \notin K - (\alpha \wedge \beta)$. So $\text{Min}_{\leq}(\neg\alpha) \subseteq \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ and thus $\nabla_{\leq}(\neg\alpha) \subseteq \nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha \vee \beta)$, from which it follows that $\beta \in K \div \alpha$.

2. Suppose that $\beta \in K \div \alpha$. Then $\nabla_{\leq}(\neg\alpha) \subseteq M(\beta)$, and so $\nabla_{\leq}(\neg\alpha) \subseteq M(\alpha \vee \beta)$, and thus $\alpha \vee \beta \in K \ddot{-} \alpha$. Furthermore, $y \not\prec x$ for every $x \in \text{Min}_{\leq}(\neg\alpha)$ and every $y \in \text{Min}_{\leq}(\neg\beta)$, and so $\alpha \notin K \ddot{-} \beta$. Conversely, suppose that $\alpha \vee \beta \in K \ddot{-} \alpha$ and $\alpha \notin K \ddot{-} \beta$. Then $\nabla_{\leq}(\neg\alpha) \subseteq M(\alpha \vee \beta)$, which means $\text{Min}_{\leq}(\neg\alpha) \subseteq M(\beta)$. Furthermore, $\nabla_{\leq}(\neg\beta) \not\subseteq M(\alpha)$, and so $y \not\prec x$ for every $x \in \text{Min}_{\leq}(\neg\alpha)$ and every $y \in \text{Min}_{\leq}(\neg\beta)$. Therefore $\nabla_{\leq}(\neg\alpha) \setminus \text{Min}_{\leq}(\neg\alpha) \subseteq M(\beta)$ and thus $\beta \in K \div \alpha$.

□

And as a corollary of propositions 6.3.17 and 6.3.20, the identities (Def $-$ from \sim) and (Def \div from $-$) are interchangeable when applied to AGM contraction and systematic withdrawal. Similarly, the identities (Def $\ddot{-}$ from \sim) and (Def \div from $\ddot{-}$) are interchangeable when applied to severe withdrawal and systematic withdrawal.

Corollary 6.3.21 *Let $-$ be an AGM contraction, let \div be a systematic withdrawal, and let $\ddot{-}$ be a severe withdrawal. Suppose that $-$, \div and $\ddot{-}$ are semantically related.*

1. $-$ and \div can also be defined in terms of one another using (Def $-$ from \sim) and (Def \div from $-$).
2. $\ddot{-}$ and \div can also be defined in terms of one another using (Def $\ddot{-}$ from \sim) and (Def \div from $\ddot{-}$).

6.3.5 Systematic withdrawal vs. severe withdrawal

Systematic withdrawal and severe withdrawal are motivated by similar concerns. Indeed, they apply the principles of Indifference, Preference and Informational Economy in the same manner, and the method of construction used is identical; they differ only in the choice of faithful layered preorders to apply to (Def \sim from ∇_{\leq}). As a result, they have many features in common. Firstly, both these forms of withdrawal are special cases of Cantwell's [1999] fallback-based withdrawal. Moreover, it is easily verified that systematic withdrawal and severe withdrawal satisfy (K-7), and that severe withdrawal, like systematic withdrawal, satisfies (K \div 7) and (K \div 10).

Proposition 6.3.22 *Every systematic and every severe withdrawal satisfies (K-7), and every severe withdrawal satisfies (K \div 7) and (K \div 10).*

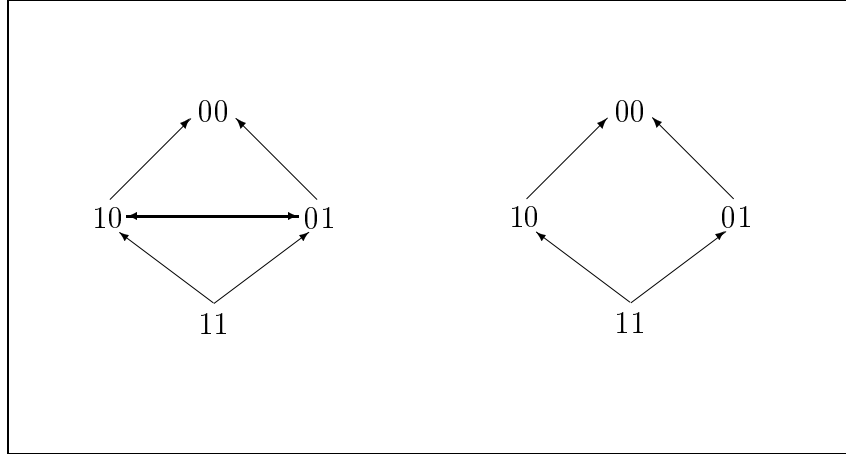


Figure 6.2: A graphical representation of the faithful total preorder \preceq and the semantically related faithful modular weak partial order \leq used in example 6.3.23. In both figures, two interpretations x and y are in the relevant faithful preorder iff (x, y) is in the reflexive transitive closure of the relation determined by the arrows.

Proof For severe withdrawal, note that (K-7) follows easily from (K-7), and that (K÷7) follows easily from (K-7) if $\not\models \alpha$. For the remaining part of (K÷7), suppose that $\models \alpha$ and that $\gamma \in K \ddot{-} \alpha \wedge \gamma$. By (K-5), $\gamma \in K \ddot{-} \gamma$ and thus $\models \gamma$ by (K-4). And then $\gamma \in K \ddot{-} \alpha \wedge \beta \wedge \gamma$ by (K-1). The proof that severe withdrawal satisfies (K÷10) is identical to the proof that systematic withdrawal satisfies (K÷10). It can be found in appendix B, proposition B.1.2.

Let \div be a systematic withdrawal defined in terms of the faithful modular weak partial order \leq using (Def \sim from ∇_{\leq}). To prove that \div satisfies (K-7), it suffices to show that $M(K) \cup \nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(K) \cup \nabla_{\leq}(\neg\alpha) \cup \nabla_{\leq}(\neg\beta)$. Pick any $x \in M(K) \cup \nabla_{\leq}(\neg(\alpha \wedge \beta))$. We only consider the case where $x \notin M(K)$. If $x \in M(\neg(\alpha \wedge \beta))$ then, from (Def ∇_{\leq}), $x \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$. Therefore either $x \in \text{Min}_{\leq}(\neg\alpha)$ or $x \in \text{Min}_{\leq}(\neg\beta)$. And since $\text{Min}_{\leq}(\neg\alpha) \subseteq \nabla_{\leq}(\neg\alpha)$ and $\text{Min}_{\leq}(\neg\beta) \subseteq \nabla_{\leq}(\neg\beta)$, it follows that $x \in \nabla_{\leq}(\neg\alpha) \cup \nabla_{\leq}(\neg\beta)$. On the other hand, if $x \in M(\alpha \wedge \beta)$ then there is an $y \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ such that $x \leq y$. Now, either $y \in \text{Min}_{\leq}(\neg\alpha)$ or $y \in \text{Min}_{\leq}(\neg\beta)$. In the former case, $x \in \nabla_{\leq}(\neg\alpha)$ and in the latter case $x \in \nabla_{\leq}(\neg\beta)$. \square

And at the risk of illustrating the obvious, the next example shows that neither systematic withdrawal nor severe withdrawal satisfies the Recovery postulate.

Example 6.3.23 Let L be the propositional language generated by the two atoms p and q , and let (V, \models) be the valuation semantics for L , with $V = \{00, 01, 10, 11\}$. Furthermore, let $K = Cn(\{p, q\})$. Now, let \preceq be the faithful total preorder defined as follows:

$$x \preceq y \text{ iff } \begin{cases} y \in V \text{ if } x = 11, \\ y \in \{01, 10, 00\} \text{ if } x \in \{01, 10\}, \text{ and} \\ y = 00 \text{ if } x = 00, \end{cases}$$

and let \leq be the associated faithful modular weak partial order defined in terms of \preceq using (Def \leq from \preceq). Figure 6.2 contains graphical representations of \preceq and \leq . Let $\ddot{-}$ be the severe withdrawal defined in terms of \preceq using (Def \sim from ∇_{\preceq}), and let \div be the systematic withdrawal defined in terms of \leq using (Def \sim from ∇_{\leq}). So $\nabla_{\preceq}(\neg(p \vee q)) = \nabla_{\leq}(\neg(p \vee q)) = V$ and thus $K \ddot{-}(p \vee q) = K \div(p \vee q) = Th(V) = Cn(\top)$. But $K \ddot{-}(p \vee q) + (p \vee q) = (K \div(p \vee q)) + (p \vee q) = Cn(p \vee q) \subset K$, thus invalidating Recovery. \square

The close relationship between systematic and severe withdrawal raises the question of whether the two notions ever coincide. Part of the answer to this question is easy. Whenever a faithful layered preorder \preceq is both a total preorder and a modular weak partial order, the severe withdrawal and the systematic withdrawal defined in terms of \preceq using (Def \sim from ∇_{\preceq}) are, by definition, identical. It is easy to see that this is the case only when \preceq is a K -linear order (see definition 5.5.8).

Proposition 6.3.24 *Let \preceq be any K -linear order. The belief removal defined in terms of \preceq using (Def \sim from ∇_{\preceq}) is a severe withdrawal as well as a systematic withdrawal.*

Proof Follows immediately from the fact that \preceq is both a faithful total preorder and a faithful modular weak partial order. \square

Furthermore, if a severe withdrawal cannot be defined in terms of a K -linear order using (Def \sim from ∇_{\preceq}), then it is not a severe withdrawal, and vice versa; at least for the finitely generated propositional languages.

Proposition 6.3.25 *Let L be a finitely generated propositional language with a valuation semantics (V, \models) .*

1. *Let $\ddot{-}$ be a severe withdrawal that cannot be defined in terms of a K -linear order using (Def \sim from ∇_{\preceq}). Then $\ddot{-}$ is not a systematic withdrawal.*

2. Let \div be a systematic withdrawal that cannot be defined in terms of a K -linear order using (Def \sim from ∇_{\preceq}). Then \div is not a severe withdrawal.

Proof 1. Assume that $\ddot{-}$ is a systematic withdrawal. Now, let \preceq be a faithful total preorder in terms of which $\ddot{-}$ is defined using (Def \sim from ∇_{\preceq}), and let \leq be the faithful modular weak partial order defined in terms of \preceq using (Def \sim from ∇_{\preceq}). By corollary 6.3.16, the systematic withdrawal \div defined in terms of \leq using (Def \sim from ∇_{\preceq}) is revision-equivalent to $\ddot{-}$, and it thus follows that \div is equal to $\ddot{-}$. By supposition, \preceq is not a K -linear order, which means there are at least two distinct countermodels, x and y , of K such that $x \equiv_{\preceq} y$ and $x \parallel_{\leq} y$. Now, let α be a wff such that $M(\alpha) = \{x\}$. (By our choice of L , there is such an α .) Then $\nabla_{\preceq}(\alpha) \neq \nabla_{\leq}(\alpha)$ and thus $K \ddot{-} \neg\alpha \neq K \div \neg\alpha$; a contradiction.

2. The proof is similar to that of part (1) and is omitted. □

Notwithstanding the similarities between systematic and severe withdrawal, there are fundamental differences between them as well. We now come to a number of properties that are indicative of the major differences. Interestingly enough, the intuitive plausibility of all these properties are, in some way or another, related to the following simple example.¹¹

Example 6.3.26 While reading about Cleopatra, I have come across one source claiming that she had a son, and another claiming that she had a daughter. Now consider the following three situations.

1. If I attend a talk about the life and times of Cleopatra, and the speaker, an expert on the subject, says something which prompts me to retract the belief that Cleopatra had a son, it seems reasonable to retain the belief that she had a daughter.
2. Similarly, if the speaker leads me to retract the belief that Cleopatra had a daughter, I should retain the belief that she had a son.
3. And finally, suppose that the speaker relates an incident which is specific enough to cast doubts on my belief that she had a son *and* a daughter, but is too vague

¹¹This is a variant of example 6.1.2.

to indicate whether she didn't have a son, didn't have a daughter, or perhaps, did not have any children at all. In these circumstances, intuition dictates that I should retain the belief that she had a child, without committing myself to a belief about it being a son or a daughter.

To formalise this example, let L be a propositional language generated by the two atoms p and q . Let p denote the assertion that Cleopatra had a son, and q the assertion that she had a daughter. Then $K = Cn(p, q)$. The three different situations described above are then formalised as follows:

$$K \sim p = Cn(q), K \sim q = Cn(p), \text{ and } K \sim (p \wedge q) = Cn(p \vee q).$$

It is easily verified that the systematic withdrawal in example 6.3.23 is able to accommodate example 6.3.26, but as we shall see below, the adherence to (K $\ddot{-}$ 7) ensures that severe withdrawal disallows this type of withdrawal. \square

Let us now consider each of the relevant properties indicating the differences between systematic withdrawal and severe withdrawal. The first one is the property expressed by (K $\ddot{-}$ 7). That it is not satisfied by systematic withdrawal, unlike severe withdrawal, is evident by considering the systematic withdrawal in example 6.3.23, and noting that $q \in K \div p$, but that $q \notin K \div (p \wedge q)$. Rott and Pagnucco [1999] argue in favour of (K $\ddot{-}$ 7) by making an appeal to the principles of Indifference and Preference. Observe that an $\alpha \wedge \beta$ -withdrawal forces us to get rid of at least one of α or β . If α is given up, they argue, we can obtain an α -withdrawal by abandoning the same beliefs as when withdrawing $\alpha \wedge \beta$. And if β is given up, we might have to remove even more beliefs. Information-theoretically, this can be justified as follows. If α is given up during an $\alpha \wedge \beta$ -withdrawal, the worst content bits of $\alpha \wedge \beta$ is at least as entrenched as the worst content bits of α . But the content bits of α are also content bits of $\alpha \wedge \beta$, and the worst content bits of $\alpha \wedge \beta$ can thus not be more entrenched than the worst content bits of α . From the principles of Indifference and Preference it then follows that an α -withdrawal should result in the removal of exactly the same infatoms as an $\alpha \wedge \beta$ -withdrawal. On the other hand, if β is given up during an $\alpha \wedge \beta$ -withdrawal, it follows by similar reasoning that the worst content bits of $\alpha \wedge \beta$ and of β are all equally entrenched, with the worst content elements of α at least as entrenched, and possibly more entrenched. Consequently, the principles of Indifference and Preference dictate that an α -withdrawal should remove at least as much infatoms as an $\alpha \wedge \beta$ -withdrawal.

A careful analysis of the argument advanced above makes it clear that it relies heavily on the assumption that two infatoms can never be *incomparable*. In other words, it assumes the existence of a faithful total preorder to measure the relative entrenchment of infatoms. But the moment this restriction is relaxed to, say, a faithful modular weak partial order, the postulate (K $\ddot{-}$ 7) is *not* sanctioned by the same appeal to principles of Indifference and Preference. This can, perhaps, best be illustrated by example 6.3.26. Even though both p and q are given up during a $p \wedge q$ -withdrawal, we don't want either a p -withdrawal or a q -withdrawal to remove as much information as a $p \wedge q$ -withdrawal.

The next property we consider is that expressed by the postulate (K \div 9). Unlike systematic withdrawal, it is not satisfied by severe withdrawal, a result which can be verified by noting that for the severe withdrawal $\ddot{-}$ in example 6.3.23, $p \in K$, $p \vee q \in K \ddot{-} p$ and $q \notin K \ddot{-} p$, but $p \notin K \ddot{-} (p \wedge q)$. Intuitively, we can justify (K \div 9) as follows. If $\alpha \vee \beta$, but not β , is retained after an α -withdrawal, it is an indication that β is more easily dislodged from K than α . Consequently, we should retain α , and discard β , when having to withdraw $\alpha \wedge \beta$.

Rott and Pagnucco [1999] show that severe withdrawal satisfies the following properties:

(Inclusion) Either $K \ddot{-} \alpha \subseteq K \ddot{-} \beta$ or $K \ddot{-} \beta \subseteq K \ddot{-} \alpha$

(Decomposition) Either $K \ddot{-} (\alpha \wedge \beta) = K \ddot{-} \alpha$ or $K \ddot{-} (\alpha \wedge \beta) = K \ddot{-} \beta$

(Converse conjunctive inclusion) If $\not\models \alpha$, $\not\models \beta$, and $K \ddot{-} (\alpha \wedge \beta) \subseteq K \ddot{-} \beta$ then $\beta \notin K \ddot{-} \alpha$

(Expulsiveness) If $\not\models \alpha$ and $\not\models \beta$ then either $\alpha \notin K \ddot{-} \beta$ or $\beta \notin K \ddot{-} \alpha$

Rott and Pagnucco regard it as regrettable that severe withdrawal satisfies Expulsiveness, in particular, and write as follows:

“Expulsiveness is an undesirable property since we do not necessarily want sentences that intuitively have nothing to do with one another to affect each other in belief contractions. This is the bitter pill we have to swallow if we want to adhere to the principles of Indifference and Preference.”

We contend that it is the use of the faithful *total* preorders, and not these two principles that are the problem. This is made abundantly clear by noting that systematic

withdrawal does not satisfy Expulsiveness. In fact, by considering the systematic withdrawal in example 6.3.23, and taking p as α , and q as β in the four properties above, we see that systematic withdrawal doesn't satisfy any of the four properties above. Example 6.3.26 is thus evidence of the undesirability of these properties.

An analysis of the properties above creates the impression that, at least in some respects, severe withdrawal removes too much information from a belief set. This impression is strengthened by noting that severe withdrawal, unlike systematic withdrawal, includes the following particularly severe instance of proper withdrawal:

$$(\text{Def } \dot{-}) \quad K \dot{-} \alpha = \begin{cases} Cn(\top) & \text{if } \alpha \in K \setminus Cn(\top), \\ K & \text{otherwise} \end{cases}$$

Proposition 6.3.27 *The belief removal $\dot{-}$ defined in (Def $\dot{-}$) is a severe withdrawal, but not a systematic withdrawal.*

Proof It is easily verified that $\dot{-}$ is defined in terms of the faithful total preorder \preceq using (Def \sim from \preceq), where \preceq is defined as follows:

$$x \preceq y \text{ iff } \begin{cases} y \in U \text{ if } x \in M(K), \\ y \in U \setminus M(K), \text{ otherwise} \end{cases}$$

and $\dot{-}$ is thus a severe withdrawal. Now assume that $\dot{-}$ is also a systematic withdrawal. Clearly $\dot{-}$ is revision-equivalent to itself, and by corollary 6.3.16 it then follows that there is no other systematic withdrawal that is revision-equivalent to $\dot{-}$. Now, let \leq be the faithful modular weak partial order that is semantically related to \preceq . Since \leq is minimal-equivalent to \preceq , it follows from corollary 6.3.16 that the systematic withdrawal \div defined in terms of \leq using (Def \sim from \preceq) is revision-equivalent to $\dot{-}$, and it is easily verified that \div is not equal to $\dot{-}$; a contradiction. \square

At the beginning of this section we saw that systematic withdrawal and severe withdrawal sometimes coincide. A related question is whether these two forms of withdrawal ever coincide with AGM contraction. It turns out that full meet contraction (see page 22) is the only case for which systematic withdrawal and AGM contraction are identical. (See section 3.3.2 for a semantic description of full meet contraction.)

Proposition 6.3.28 *Full meet contraction is the only AGM contraction that is a systematic withdrawal.*

Proof The full meet contraction $-$ can be defined in terms of the following faithful modular weak partial order using (Def \sim from \preceq): $x \leq y$ iff $x = y$, or $x \in M(K)$ and $y \notin M(K)$. It is therefore, by definition, a systematic withdrawal. Next we show that $-$ is the only belief removal that is both an AGM contraction and a systematic withdrawal. Pick any systematic withdrawal \div other than $-$. So $K - \alpha \neq K \div \alpha$ for some $\alpha \in K \setminus Cn(\top)$. If $K - \alpha \not\subseteq K \div \alpha$, then there is a $\beta \in K - \alpha$ (and thus $\beta \in K$) such that $\beta \notin K \div \alpha$. Since $K - \alpha = Th(M(K) \cup M(\neg\alpha))$, there is therefore an $x \in M(K \div \alpha)$ such that $x \in M(\alpha \wedge \neg\beta)$. And thus $\beta \notin K \div \alpha + \alpha$, which is a violation of Recovery. So suppose that $K - \alpha \subset K \div \alpha$. Now let \preceq be a faithful modular weak partial order in terms of which \div is defined using (Def \sim from ∇_{\preceq}). Since $K - \alpha = Th(M(K) \cup M(\neg\alpha))$, it follows from $K - \alpha \subset K \div \alpha$ that there is a $\beta \in Th(M(K) \cup \nabla_{\preceq}(\neg\alpha))$ such that $y \in M(\neg\beta)$ for some $y \in M(\neg\alpha)$. So $\not\models \alpha \vee \beta$, and since $Min_{\preceq}(\neg\alpha) \subseteq \nabla_{\preceq}(\neg\alpha)$, $Min_{\preceq}(\neg\alpha) \cap Min_{\preceq}(\neg(\alpha \vee \beta)) = \emptyset$, which means that $Min_{\preceq}(\neg\alpha) \prec Min_{\preceq}(\neg(\alpha \vee \beta))$.¹² By smoothness, $Min_{\preceq}(\neg\alpha) \neq \emptyset$, and there is thus an $x \in M(\neg\alpha \wedge \beta)$ such that $x \prec z$ for every $z \in Min_{\preceq}(\neg(\alpha \vee \beta))$. So $x \in \nabla_{\preceq}(\neg(\alpha \vee \beta))$ and thus $\alpha \notin K \div (\alpha \vee \beta) + (\alpha \vee \beta)$. So \div does not satisfy Recovery, and is therefore not an AGM contraction. \square

With the exception of some cases involving a few trivial belief sets, though, severe withdrawal and AGM contraction always produce different results.

Proposition 6.3.29 *Let K be such that for some $\alpha, \beta \in K$, $\not\models \alpha$, $\not\models \beta$ and $\alpha \not\equiv \beta$. Then severe withdrawal and AGM contraction never coincide.*

Proof Pick any severe withdrawal $\ddot{-}$ and let \preceq be a faithful total preorder in terms of which $\ddot{-}$ can be defined using (Def \sim from ∇_{\preceq}). If $Min_{\preceq}(\neg(\alpha \leftrightarrow \beta)) \subseteq M(\alpha)$ then $\nabla_{\preceq}(\neg(\alpha \vee \neg\beta)) \not\subseteq M(\beta)$ and so $\beta \notin K \ddot{-}(\alpha \vee \neg\beta) + (\alpha \vee \neg\beta)$, which is a violation of Recovery. The remaining two cases, where $Min_{\preceq}(\neg(\alpha \leftrightarrow \beta)) \subseteq M(\beta)$, and where $Min_{\preceq}(\neg(\alpha \leftrightarrow \beta)) \not\subseteq M(\alpha)$ and $Min_{\preceq}(\neg(\alpha \leftrightarrow \beta)) \not\subseteq M(\beta)$, are similar. \square

We conclude this section with a suggestion prompted by a remark from Hans Rott [personal communication] that it seems difficult to come up with yet more appropriate forms of principled withdrawal. It turns out that there *is* a semantic way to describe another set of reasonable withdrawals, all of which exhibit principled behaviour. The

¹²See section 1.3 for an explanation of the convention of applying \preceq , \prec and \equiv_{\preceq} to sets of interpretations.

method of constructing this set relies on the principles of Indifference, Preference and Informational Economy, and they are employed in a manner identical to that used in the construction of severe and systematic withdrawal. It differs only from the constructive modellings of systematic and severe withdrawal in the choice of permissible faithful preorders.

Definition 6.3.30 A belief removal is a *methodical withdrawal* iff it is defined in terms of a faithful layered preorder using $(\text{Def } \sim \text{ from } \nabla_{\preceq})$. \square

Since the set of faithful layered preorders includes the faithful modular weak partial orders and the faithful total preorders, methodical withdrawal includes both systematic and severe withdrawal. However, it excludes the AGM contractions which do not coincide with systematic withdrawal. It is our contention that methodical withdrawal constitutes a class of withdrawals that deserve further study. We provide a tentative first step in this direction with a result involving some properties of methodical withdrawal.

Proposition 6.3.31 *Methodical withdrawal satisfies $(K-1)$ to $(K-5)$, $(K-7)$, $(K-8)$, $(K \div 7)$ and $(K \div 10)$.*

Proof The proofs are similar to those for systematic and severe withdrawal, and are omitted. \square

From theorem 6.3.8 it thus follows that methodical withdrawal is also reasonable.

6.3.6 Informational value

In section 6.2.1 we saw that proper withdrawal, as characterised by Levi's saturatable withdrawals, is too general to be regarded as principled. In particular, it contains many removals which do not satisfy $(K-7)$ and $(K-8)$. These, of course, include the basic AGM contractions that are not AGM contractions.

Levi [1991] provides two methods for obtaining a more principled form of proper withdrawal. The basic tenet of the constructions is that it is *not* the loss of information that should be minimised, but rather the loss of informational *value*. In order to achieve this, it is necessary to provide a measure \mathcal{V} on the belief sets that are subsets of the current belief set K . He considers two monotonicity conditions of \mathcal{V} :

(Strong monotonicity) If $X \subset Y$ then $\mathcal{V}(X) < \mathcal{V}(Y)$

(Weak monotonicity) If $X \subset Y$ then $\mathcal{V}(X) \leq \mathcal{V}(Y)$

Levi sees strong monotonicity as too strong a requirement to impose on \mathcal{V} , arguing instead for the imposition of weak monotonicity. This is referred to as a measure of *undamped* informational value.¹³ The intuition is that some information may have *no* informational value, and that the addition of such information should leave the informational value of a belief set unchanged. His first method uses undamped informational value. To determine the belief set resulting from an α -withdrawal of K , he finds the saturatable contractions with respect to K and α that minimises the loss of informational value, and takes their intersection. That is done as follows with the aid of a measure of undamped informational value \mathcal{V} :

$$\text{(Def } \sim \text{ from } \mathcal{V}) \quad K \sim \alpha = \begin{cases} \bigcap \{X \in sc(K, \alpha) \mid \mathcal{V}(X) \geq \mathcal{V}(Y) \ \forall Y \in sc(K, \alpha)\} \\ \text{if } \alpha \in K \setminus Cn(\top), \\ K \text{ otherwise} \end{cases}$$

Definition 6.3.32 A withdrawal is called *informational valued* iff it defined in terms of a measure of undamped informational value \mathcal{V} using (Def \sim from \mathcal{V}) \square

As Levi observes, this method is problematic from a decision-theoretic point of view, since the belief set obtained from an α -withdrawal may not represent a minimal loss in informational value.

Example 6.3.33 Let L be the propositional language generated by the atoms p and q , and let (V, \models) be the valuation semantics for L . Now let $K = Cn(p)$, and let $\mathcal{V}(Cn(p)) = 1$, $\mathcal{V}(Cn(p \vee q)) = \mathcal{V}(Cn(p \vee \neg q)) = \frac{3}{4}$, and $\mathcal{V}(Cn(\top)) = 0$. It is easily seen that $K \sim p = Cn(\top)$ for the withdrawal \sim defined in terms of \mathcal{V} using (Def \sim from \mathcal{V}). And yet

$$\mathcal{V}(Cn(\top)) = 0 < \mathcal{V}(Cn(p \vee q)) = \mathcal{V}(Cn(p \vee \neg q)) = \frac{3}{4}.$$

Choosing either $Cn(p \vee q)$ or $Cn(p \vee \neg q)$ would thus have resulted in a loss of informational value of $\frac{1}{4}$, while the choice of $Cn(\top)$ represents a loss of informational value of 1. \square

¹³Actually, Levi's measure of undamped informational value, as proposed in [Levi, 1991], is required to be a probability measure. We stick to the watered-down version used by Hansson and Olsson [1995].

To rectify this undesirable behaviour, Levi switches to *damped* informational value. A measure \mathcal{V}_D of damped informational value is determined in terms of a measure \mathcal{V} of undamped informational value as follows:

$$(\text{Def } \mathcal{V}_D \text{ from } \mathcal{V}) \quad \mathcal{V}_D(X) = \min\{\mathcal{V}(Y) \in sc(K, \alpha) \mid Y \subseteq X\}$$

In other words, the damped informational value of a belief set $X \subseteq K$ is equal to the minimum undamped informational value of the saturatable contractions contained in X . Levi's second method then defines proper withdrawals in terms of damped informational value using (Def \sim from \mathcal{V}). It is easily established that the proper withdrawals defined in terms of \mathcal{V} and \mathcal{V}_D using (Def \sim from \mathcal{V}), where \mathcal{V}_D is obtained in terms of \mathcal{V} using (Def \mathcal{V}_D from \mathcal{V}), are identical. The advantage in using damped informational value is that it can be motivated from a decision-theoretic point of view. Hansson and Olsson [1995] show that informational valued withdrawal satisfies (K-7) and (K-8). In this sense, then, it is a principled form of withdrawal.

Levi [1998] has recently expressed some doubts about the appropriateness of informational valued withdrawal, as it has been presented thus far. He presents an example which is representative of a class of informational valued withdrawals satisfying Recovery, which he sees as counterintuitive [Levi, 1998, p. 37]. Furthermore, he points out that the undamped and damped informational value of some belief sets (such as the saturatable contractions) are the same, but that it differs for others. As a result, he proposes the use of a second version of damped informational value. The removals defined in terms of this version of damped informational value using (Def \sim from \mathcal{V}) is dubbed *mild contraction*. It turns out that mild contraction coincides exactly with severe withdrawal. This is one of the reasons why Levi favours severe withdrawal over systematic withdrawal. He argues that his construction of severe withdrawal (or mild contraction) in terms of undamped informational value (version 2) provides a decision-theoretic motivation; something that systematic withdrawal does not appear to possess.

6.4 Withdrawal and entrenchment

As discussed in chapter 5, entrenchment orderings on wffs are intended to provide a measure of the extent to which a particular belief of an agent is entrenched in its belief set. As such, these orderings can be useful in the construction of withdrawals. In

this regard, we have already seen how AGM contraction can be defined in terms of two forms of entrenchment; the EE-orderings of section 2.3, and the RE-orderings of section 5.5. In fact, we saw in section 5.8 that the EE-orderings, the RE-orderings and AGM contraction are interchangeable in terms of the relevant identities. In this section we show that this interchangeability can be extended to include systematic and severe withdrawal as well. It will be convenient to generalise the notion of semantic relatedness found in definitions 5.8.1 and 6.3.18.

Definition 6.4.1 An AGM contraction $-$, an AGM revision $*$, an EE-ordering \sqsubseteq_{EE} , a GE-ordering \sqsubseteq_{GE} , an RE-ordering \sqsubseteq_{RE} , an RG-ordering \sqsubseteq_{RG} , a systematic withdrawal \div , and a severe withdrawal $\ddot{-}$ are *semantically related* iff there is a faithful total preorder \preceq and a semantically related faithful modular weak partial order \leq such that

1. $-$ is defined in terms of \preceq (and \leq) using (Def \sim from \preceq),
2. $*$ is defined in terms of \preceq (and \leq) using (Def $*$ from \preceq),
3. \sqsubseteq_{EE} is defined in terms of \preceq using (Def \sqsubseteq_E from \preceq),
4. \sqsubseteq_{GE} is defined in terms of \preceq using (Def \sqsubseteq_G from \preceq),
5. \sqsubseteq_{RE} is defined in terms of \leq using (Def \sqsubseteq_E from \preceq),
6. \sqsubseteq_{RG} is defined in terms of \leq using (Def \sqsubseteq_G from \preceq),
7. \div is defined in terms of \leq using (Def \sim from ∇_{\preceq}), and
8. $\ddot{-}$ is defined in terms of \preceq using (Def \sim from ∇_{\preceq}).

□

Note that, for the remainder of this chapter, we shall frequently make use of lemma 5.2.1 without explicitly referring to it, as has been the convention in chapter 5.

Let us begin with sharper versions of results by Rott and Pagnucco [1999], showing that severe withdrawal and epistemic entrenchment are interdefinable by means of the following two identities:

$$(\text{Def } \ddot{-} \text{ from } \sqsubseteq_{EE}) \quad K \ddot{-} \alpha = \begin{cases} K \cap \{\beta \mid \alpha \sqsubseteq_{EE} \beta\} & \text{if } \alpha \in K \setminus Cn(\top), \\ K & \text{otherwise} \end{cases}$$

(Def \sqsubseteq_{EE} from $\ddot{-}$) $\alpha \sqsubseteq_{EE} \beta$ iff $\alpha \notin K \ddot{-} \beta$ or $\models \beta$

Proposition 6.4.2 *If an EE-ordering \sqsubseteq_{EE} and a severe withdrawal $\ddot{-}$ are semantically related, they can also be defined in terms of one another using (Def \sqsubseteq_{EE} from $\ddot{-}$) and (Def $\ddot{-}$ from \sqsubseteq_{EE}).*

Proof Let \preceq be a faithful preorder in terms of which \sqsubseteq_{EE} and $\ddot{-}$ are defined using (Def \sqsubseteq_{EE} from \preceq) and (Def \sim from ∇_{\preceq}). We only consider the case where $\alpha, \beta \in K \setminus Cn(\top)$. Observe that $\beta \in K \ddot{-} \alpha$ iff there is a $y \in Min_{\preceq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \preceq y$, iff $\beta \not\sqsubseteq_{EE} \alpha$, iff $\alpha \sqsubseteq_{EE} \beta$. And then observe that $\alpha \sqsubseteq_{EE} \beta$ iff for every $y \in Min_{\preceq}(\neg\beta)$ there is an $x \in M(\neg\alpha)$ such that $x \preceq y$, iff $\alpha \notin K \ddot{-} \beta$. \square

Proposition 6.4.2 thus also shows that the identities (Def $\ddot{-}$ from \sqsubseteq_{EE}) and (Def \sqsubseteq_{EE} from $\ddot{-}$) are interchangeable. Note that (Def \sqsubseteq_{EE} from $\ddot{-}$) and (Def $\ddot{-}$ from \sqsubseteq_{EE}) provide a very elegant method for moving between severe withdrawal and epistemic entrenchment. Barring some limiting cases, a wff β is in the belief set resulting from a severe α -withdrawal iff β is more entrenched than α .

Interestingly enough, Rott and Pagnucco [1999] show that (Def \sqsubseteq_{EE} from $\ddot{-}$) and (Def \sqsubseteq_{EE} from \sim) are equivalent when applied to severe withdrawals. This observation prompts us to show that the application of (Def \sqsubseteq_{EE} from \sim) to any two revision-equivalent reasonable withdrawals yields the same EE-ordering.

Proposition 6.4.3 *Let \sim and \approx be two reasonable withdrawals that are revision-equivalent. The EE-orderings defined in terms of \sim and \approx using (Def \sqsubseteq_{EE} from \sim) are identical.*

Proof Let $-$ and $\ddot{-}$ be the unique AGM contraction and severe withdrawal respectively, that are revision-equivalent to \sim and \approx . It suffices to show that the EE-orderings defined in terms of \sim and $-$ using (Def \sqsubseteq_{EE} from \sim) are identical. So let \sqsubseteq_{EE}^- be the EE-ordering defined in terms of $-$ using (Def \sqsubseteq_{EE} from \sim), and let $\sqsubseteq_{EE}^{\ddot{-}}$ be the EE-ordering defined in terms of $\ddot{-}$ using (Def \sqsubseteq_{EE} from $\ddot{-}$). Since (Def \sqsubseteq_{EE} from $\ddot{-}$) and (Def \sqsubseteq_{EE} from \sim) yield identical EE-orderings when applied to severe withdrawal, $\sqsubseteq_{EE}^{\ddot{-}}$ can also be defined in terms of $\ddot{-}$ using (Def \sqsubseteq_{EE} from \sim).

First we show that \sqsubseteq_{EE}^- and $\sqsubseteq_{EE}^{\ddot{-}}$ are identical. Let \preceq be a faithful total preorder in terms of which $-$ and $\ddot{-}$ can be defined using (Def \sim from \preceq) and (Def \sim from ∇_{\preceq}) respectively. By corollary 6.3.16 there is such a \preceq . By proposition 6.4.2, $\sqsubseteq_{EE}^{\ddot{-}}$ is the

EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq), and by proposition 3.3.4, \sqsubseteq_{EE}^- is the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq). So \sqsubseteq_{EE}^- and \sqsubseteq_{EE}^- are identical.

Now, let \sqsubseteq_{EE}^\sim be the relation on L defined in terms of \sim using (Def \sqsubseteq_{EE} from \sim). We only consider the case where $\not\models \alpha \wedge \beta$. If $\alpha \sqsubseteq_{EE}^- \beta$ then $\alpha \notin K - \alpha \wedge \beta$. Since \sim is reasonable, $\alpha \notin K \sim \alpha \wedge \beta$ and thus $\alpha \sqsubseteq_{EE}^\sim \beta$. Furthermore, if $\alpha \sqsubseteq_{EE}^\sim \beta$ then $\alpha \notin K \sim \alpha \wedge \beta$. Since \sim is reasonable, $\alpha \notin K \ddot{-} \alpha \wedge \beta$ and thus $\alpha \sqsubseteq_{EE}^- \beta$. And since \sqsubseteq_{EE}^- and \sqsubseteq_{EE}^\sim are identical, the required result follows. \square

When using (Def \sqsubseteq_{RE} from \sim), we obtain a result for reasonable withdrawal and refined entrenchment which is similar to proposition 6.4.3.

Proposition 6.4.4 *Let \sim and \approx be two reasonable withdrawals that are revision-equivalent. The RE-orderings defined in terms of \sim and \approx using (Def \sqsubseteq_{RE} from \sim) are identical.*

Proof Let $-$ be the unique AGM contraction and $\ddot{-}$ the unique severe withdrawal that are both revision-equivalent to \sim and \approx . Furthermore, let \preceq be a faithful total preorder in terms of which $-$ and $\ddot{-}$ can be defined using (Def \sim from \preceq) and (Def \sim from ∇_{\preceq}) respectively. By corollary 6.3.16 there is such a \preceq . Moreover, let \leq be the faithful modular weak partial order that is semantically related to \preceq . By theorem 5.5.15 we know that the RE-ordering \sqsubseteq_{RE}^- defined in terms of \leq using (Def \sqsubseteq_E from \preceq) can also be defined in terms of $-$ using (Def \sqsubseteq_{RE} from \sim). Below we show that the binary relation $\sqsubseteq_{RE}^\ddot{-}$ on L defined in terms of $\ddot{-}$ using (Def \sqsubseteq_{RE} from \sim) is identical to \sqsubseteq_{RE}^- . The required result then follows in a manner that is similar to the proof of proposition 6.4.3.

Suppose that $\alpha \sqsubseteq_{RE}^- \beta$. Then $\alpha \rightarrow \beta \in K - \alpha \wedge \beta$ and so $M(K) \cup \text{Min}_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha \rightarrow \beta)$. From this it follows that for every $x \preceq y$, where $y \in \text{Min}_{\preceq}(\neg(\alpha \wedge \beta))$, $x \in M(\alpha \rightarrow \beta)$. That is, $\nabla_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha \rightarrow \beta)$, and thus $\alpha \rightarrow \beta \in K \ddot{-}(\alpha \wedge \beta)$, from which it follows that $\alpha \sqsubseteq_{RE}^\ddot{-} \beta$. Conversely, if $\alpha \sqsubseteq_{RE}^\ddot{-} \beta$ then $\alpha \rightarrow \beta \in K \ddot{-}(\alpha \wedge \beta)$. But this means that $\alpha \rightarrow \beta \in K - (\alpha \wedge \beta)$ and so $\alpha \sqsubseteq_{RE}^- \beta$. \square

A result similar to proposition 6.4.2 holds for severe withdrawal and refined entrenchment when (Def \sqsubseteq_{RE} from \sim) and the identity below are used:

$$(\text{Def } \ddot{-} \text{ from } \sqsubseteq_{RE}) \quad K \ddot{-} \alpha = \begin{cases} K \cap \{\beta \mid \beta \not\sqsubseteq_{RE} \alpha \text{ and } \beta \rightarrow \alpha \sqsubseteq_{RE} \beta\} & \text{if } \not\models \alpha, \\ K & \text{otherwise} \end{cases}$$

Proposition 6.4.5 *If an RE-ordering \sqsubseteq_{RE} and a severe withdrawal $\ddot{-}$ are semantically related, then they can also be defined in terms of one another using (Def \sqsubseteq_{RE} from \sim) and (Def $\ddot{-}$ from \sqsubseteq_{RE}).*

Proof Let \preceq be a faithful total preorder in terms of which $\ddot{-}$ can be defined using (Def $\ddot{-}$ from ∇_{\preceq}), and let \leq be the semantically related faithful modular weak partial order. The validity of the application of (Def $\ddot{-}$ from \sqsubseteq_{RE}) follows easily from proposition 6.4.2 and theorem 5.5.7. For (Def \sqsubseteq_{RE} from \sim) we only consider the case where $\not\models \alpha$. Observe that $\alpha \sqsubseteq_{RE} \beta$ iff for every $y \in \text{Min}_{\preceq}(\neg\beta)$ there is an $x \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \leq y$, iff $\nabla_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha \rightarrow \beta)$, iff $\alpha \rightarrow \beta \in K \ddot{-}(\alpha \wedge \beta)$. \square

Next is a similar result for systematic withdrawal and the EE-orderings, when using (Def \sqsubseteq_{EE} from \sim) and the identity below:

$$(\text{Def } \div \text{ from } \sqsubseteq_{EE}) \quad \beta \in K \div \alpha \text{ iff } \begin{cases} \alpha \sqsubseteq_{EE} \alpha \vee \beta \text{ and } \alpha \sqsubseteq_{EE} \beta \\ \text{if } \alpha \in K \setminus \text{Cn}(\top), \\ \beta \in K \text{ otherwise} \end{cases}$$

Proposition 6.4.6 *If the EE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related, then they can also be defined in terms of one another using (Def \div from \sqsubseteq_{EE}) and (Def \sqsubseteq_{EE} from \sim).*

Proof Let \preceq be a faithful total preorder in terms of which \sqsubseteq_{EE} can be defined using (Def \sqsubseteq_E from \preceq), and let \leq be the semantically related faithful modular weak partial order. For (Def \div from \sqsubseteq_{EE}) we only consider the case where $\alpha \in K \setminus \text{Cn}(\top)$. Suppose that $\beta \in K \div \alpha$. So, there is a $y \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \in M(\beta) \subseteq M(\alpha \vee \beta)$ for every $x \leq y$. Furthermore, since $\text{Min}_{\preceq}(\neg\alpha) \subseteq M(\beta)$, $x \in M(\alpha \vee \beta)$ for every x such that $y \not\prec x$. So $x \in M(\alpha \vee \beta)$ for every $x \preceq y$, which means that $\alpha \vee \beta \not\sqsubseteq_{EE} \alpha$ and thus that $\alpha \sqsubseteq_{EE} \alpha \vee \beta$. And since $x \in M(\beta)$ for every $x \prec y$, we have that $\alpha \sqsubseteq_{EE} \beta$. Conversely, suppose that $\beta \notin K \div \alpha$. So there is a $y \in M(K) \cup \nabla_{\preceq}(\neg\alpha)$ such that $y \in M(\neg\beta)$. Suppose further that $\alpha \sqsubseteq_{EE} \beta$. Then there is an $x \in M(\neg\alpha)$ such that $x \preceq y$ and thus $y \in \text{Min}_{\preceq}(\neg\alpha)$. So $y \in M(\neg(\alpha \vee \beta))$ and $y \preceq z$ for every $z \in M(\neg\alpha)$. That is, $\alpha \vee \beta \sqsubseteq_{EE} \alpha$, which means that $\alpha \not\sqsubseteq_{EE} \alpha \vee \beta$.

For (Def \sqsubseteq_{EE} from \sim), suppose that $\alpha \not\sqsubseteq_{EE} \beta$. So there is a $y \in \text{Min}_{\preceq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \preceq y$. Then $\not\models \alpha \wedge \beta$, $y \in \text{Min}_{\preceq}(\neg(\alpha \wedge \beta))$, and so $M(K) \cup \nabla_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$. That is $\alpha \in K \div (\alpha \wedge \beta)$. Conversely, suppose that

$\alpha \in K \div \alpha \wedge \beta$. We only consider the case where $\not\models \alpha \wedge \beta$. Then $M(K) \cup \nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$, and there is thus a $y \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta)) = \text{Min}_{\leq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \leq y$. Furthermore, since $\text{Min}_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$, $x \in M(\alpha)$ for every x such that $y \not\leq x$. So $x \in M(\alpha)$ for every $x \preceq y$ and thus $\alpha \not\sqsubseteq_{EE} \beta$. \square

Finally, we obtain a related result for systematic withdrawal and refined entrenchment in terms of $(\text{Def } \sqsubseteq_{RE} \text{ from } \sim)$ and the identity below.

$$(\text{Def } \div \text{ from } \sqsubseteq_{RE}) \ K \div \alpha = \begin{cases} \{\beta \mid \beta \not\sqsubseteq_{RE} \alpha \text{ and } \beta \rightarrow \alpha \sqsubseteq_{RE} \alpha\} \\ \text{if } \alpha \in K \setminus \text{Cn}(\top), \\ K \text{ otherwise} \end{cases}$$

Proposition 6.4.7 *If the RE-ordering \sqsubseteq_{RE} and the systematic withdrawal \div are semantically related, then they can also be defined in terms of one another using $(\text{Def } \div \text{ from } \sqsubseteq_{RE})$ and $(\text{Def } \sqsubseteq_{RE} \text{ from } \sim)$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using $(\text{Def } \sqsubseteq_E \text{ from } \preceq)$ and $(\text{Def } \sim \text{ from } \nabla_{\leq})$. The proof for $(\text{Def } \sqsubseteq_{RE} \text{ from } \sim)$ is identical to the part of the proof of proposition 6.4.5 concerning $(\text{Def } \sqsubseteq_{RE} \text{ from } \sim)$. For $(\text{Def } \div \text{ from } \sqsubseteq_{RE})$, we only consider the case where $\alpha \in K \setminus \text{Cn}(\top)$. Suppose that $\beta \in K \div \alpha$. So there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \leq y$. That is, $\beta \not\sqsubseteq_{RE} \alpha$. Note further that for every $z \in \text{Min}_{\leq}(\neg\alpha)$, $z \in M(\beta)$, and so $\beta \rightarrow \alpha \sqsubseteq_{RE} \alpha$. Conversely, suppose that $\beta \not\sqsubseteq_{RE} \alpha$ and $\beta \rightarrow \alpha \sqsubseteq_{RE} \alpha$. From $\beta \not\sqsubseteq_{RE} \alpha$ there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \leq y$, and from $\beta \rightarrow \alpha \sqsubseteq_{RE} \alpha$ it follows that $y \in M(\beta)$ for every $y \in \text{Min}_{\leq}(\neg\alpha)$. And therefore, $\beta \in K \div \alpha$. \square

6.5 Systematic withdrawal and entrenchment

Section 6.4 contains a plethora of results, providing strong links between severe withdrawal, systematic withdrawal, the EE-orderings and the RE-orderings, in terms of appropriate identities. But with the exception of the connection between severe withdrawal and the EE-orderings, it is difficult to view these identities as intuitively plausible descriptions of how these constructions relate to each other. This is not unlike the connection between AGM contraction and epistemic entrenchment provided by the identities $(\text{Def } - \text{ from } \sqsubseteq_{EE})$ and $(\text{Def } \sqsubseteq_{EE} \text{ from } \sim)$ in section 2.3. $(\text{Def } - \text{ from } \sqsubseteq_{EE})$

in particular, has been the subject of some scrutiny in the theory change literature. With the exception of some limiting cases, it shows that a wff β is in the α -contraction of K iff $\alpha \vee \beta$ is more entrenched than α . It is, of course, the use of $\alpha \vee \beta$, instead of β above, that is the source of concern. Blackburn et al. [1997], for example, attribute the use of $\alpha \vee \beta$ to “technical reasons”. Gärdenfors and Makinson [1988] provide a motivation for its use, but it is somewhat difficult to understand, and depends on the acceptance of the Recovery postulate. More recently, Gärdenfors has admitted that the identity is somewhat counterintuitive [1992,p. 19].

In section 5.1 we gave a different characterisation of AGM contraction in terms of the EE-orderings; one that, in our opinion, provides a closer match with the Gärdenfors intuition that contraction in terms of epistemic entrenchment is based on the idea of being “forced to choose” between the removal of two wffs. In this section we intend to provide an analogous match between systematic withdrawal and refined entrenchment. We start by showing that the EE-orderings have too coarse a grainsize to provide a suitable intuitive description of systematic withdrawal. This is followed by another description of systematic withdrawal in terms of the RE-orderings; one which differs from the one given in proposition 6.4.7. Finally, we show that for the finitely generated propositional case, there is a graph based procedure for defining systematic withdrawal in terms of refined entrenchment.

6.5.1 Systematic withdrawal and the EE-orderings

The reason that (Def – from \sqsubseteq_{EE}) is seen as a somewhat counterintuitive definition of AGM contraction in terms of the EE-orderings is that wffs that are *less* entrenched than a wff α are sometimes retained during an α -contraction, as the next example shows.

Example 6.5.1 Let L be the propositional language generated by the two atoms p and q , and let (V, \models) , with $V = \{00, 01, 10, 11\}$, be the valuation semantics for L . Let $K = Cn(p)$ and define the EE-ordering \sqsubseteq_{EE} as follows:

$$\alpha \sqsubseteq_{EE} \beta \text{ iff } \begin{cases} \beta \in L \text{ if } \alpha \notin K, \\ p \models \beta \text{ if } \alpha \equiv p \text{ or } \alpha \equiv p \vee \neg q, \\ p \vee q \models \beta \text{ if } \alpha \equiv p \vee q, \text{ and} \\ \models \beta \text{ if } \alpha \equiv \top. \end{cases}$$

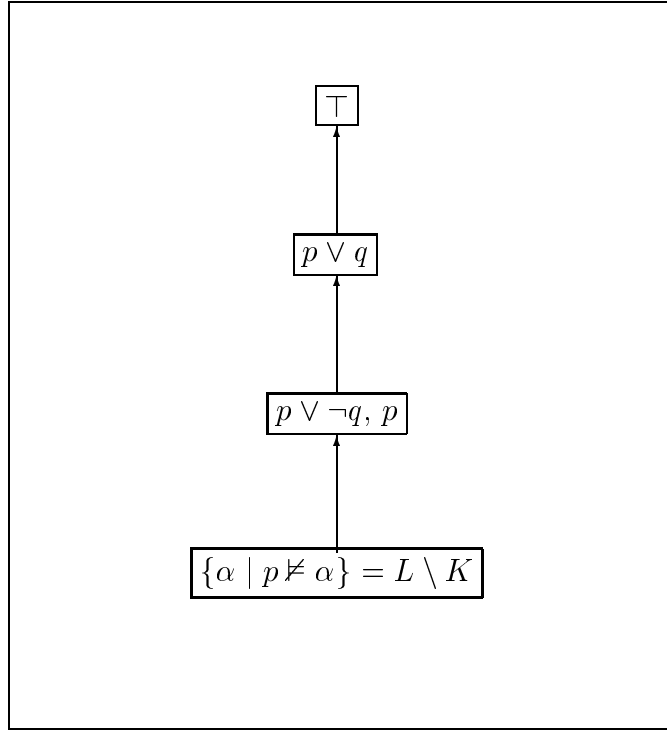


Figure 6.3: A graphical representation of the EE-ordering \sqsubseteq_{EE} with respect to the belief set $K = Cn(\{p\})$. This EE-ordering is used in example 6.5.1. For every $\alpha, \beta \in L$, $\alpha \sqsubseteq_{EE} \beta$ iff (α, β) is in the reflexive transitive closure of the relation determined by the arrows. Each wff in this figure is a canonical representative of the set of wffs which are logically equivalent to it.

It is easily verified that \sqsubseteq_{EE} is indeed an EE-ordering. Figure 6.3 contains a graphical representation of \sqsubseteq_{EE} . It can be verified that the AGM contraction – defined in terms of \sqsubseteq_{EE} using (Def – from \sqsubseteq_{EE}) yields $K - (p \vee q) = Cn(p \vee \neg q)$. So $K - (p \vee q)$ contains the wff $p \vee \neg q$, a wff that is *less* entrenched than $p \vee q$. \square

Our first result shows that, unlike AGM contraction, *none* of the wffs that are less entrenched than α are in the belief set resulting from a systematic α -withdrawal.

Proposition 6.5.2 *Suppose that the EE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. If $\not\models \alpha$ and $\beta \sqsubseteq_{EE} \alpha$ then $\beta \notin K \div \alpha$.*

Proof Let \preceq be a faithful total preorder in terms of which \sqsubseteq_{EE} can be defined using (Def \sqsubseteq_E from \preceq) and let \leq be its semantically related faithful modular weak partial

order. Suppose that $\not\models \alpha$ and $\beta \sqsubseteq_{EE} \alpha$. Then $\alpha \not\sqsubseteq_{EE} \beta$ and there is thus a $y \in \text{Min}_{\preceq}(\neg\beta)$ such that $x \in M(\alpha)$ for every $x \preceq y$. So $y \prec z$ for every $z \in \text{Min}_{\preceq}(\neg\alpha)$. Therefore $\nabla_{\preceq}(\neg\alpha) \not\subseteq M(\beta)$ and thus $\beta \notin K \div \alpha$. \square

Proposition 6.5.2 describes the fate of the wffs that are less entrenched than α , but it gives no indication of what happens to the remaining wffs. The next result gives a partial answer to this question.

Proposition 6.5.3 *Suppose that the EE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. If $\alpha \sqsubseteq_{EE} \beta$ then $\beta \in K \div \alpha$.*

Proof Let \preceq be a faithful total preorder in terms of which \sqsubseteq_{EE} can be defined using (Def \sqsubseteq_E from \preceq), and let \leq be its semantically related faithful modular weak partial order. Suppose that $\alpha \sqsubseteq_{EE} \beta$. We only consider the case where $\not\models \beta$. By (EE2), $\not\models \alpha$, and from $\beta \not\sqsubseteq_{EE} \alpha$ it follows that there is a $y \in \text{Min}_{\preceq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \preceq y$. So $M(K) \cup \nabla_{\preceq}(\neg\alpha) \subseteq M(\beta)$ and thus $\beta \in K \div \alpha$. \square

The wffs that are more entrenched than α will thus all be retained after a systematic α -withdrawal. It therefore only remains to be seen what systematic withdrawal does with the wffs that are *as* entrenched as α . Unfortunately it seems that the EE-orderings are too coarse to account for a proper description of how to handle these wffs.

Example 6.5.4 Consider the language L generated by the two atoms p and q , and let (V, \models) be the valuation semantics for L , with $V = \{00, 01, 10, 11\}$. Let $K = \text{Cn}(\{p, q\})$, define the faithful total preorder \preceq as follows

$$x \preceq y \text{ iff } \begin{cases} y \in V \text{ if } x = 11, \\ y \in \{00, 01, 10\} \text{ if } x \in \{01, 10\}, \text{ and} \\ y = 00 \text{ if } x = 00 \end{cases}$$

and let \leq be the associated faithful modular weak partial order defined in terms of \preceq using (Def \leq from \preceq). Now, let \sqsubseteq_{EE} be the EE-ordering defined in terms of \preceq using (Def \sqsubseteq_E from \preceq) and let \div be the systematic withdrawal defined in terms of \leq using (Def \sim from ∇_{\preceq}). Figure 6.4 contains a graphical representation of \preceq and \sqsubseteq_{EE} . An inspection of \sqsubseteq_{EE} in figure 6.4 shows that the status of the wffs which are exactly as entrenched as the wff we want to withdraw is somewhat ambiguous. To see this, note that $K \div p = \text{Cn}(\{q\})$. So although the wffs $p \vee \neg q$, $p \leftrightarrow q$, $p \wedge q$, $\neg p \vee q$, and q are exactly as entrenched as p , some of them are in $K \div p$, while others are not. \square

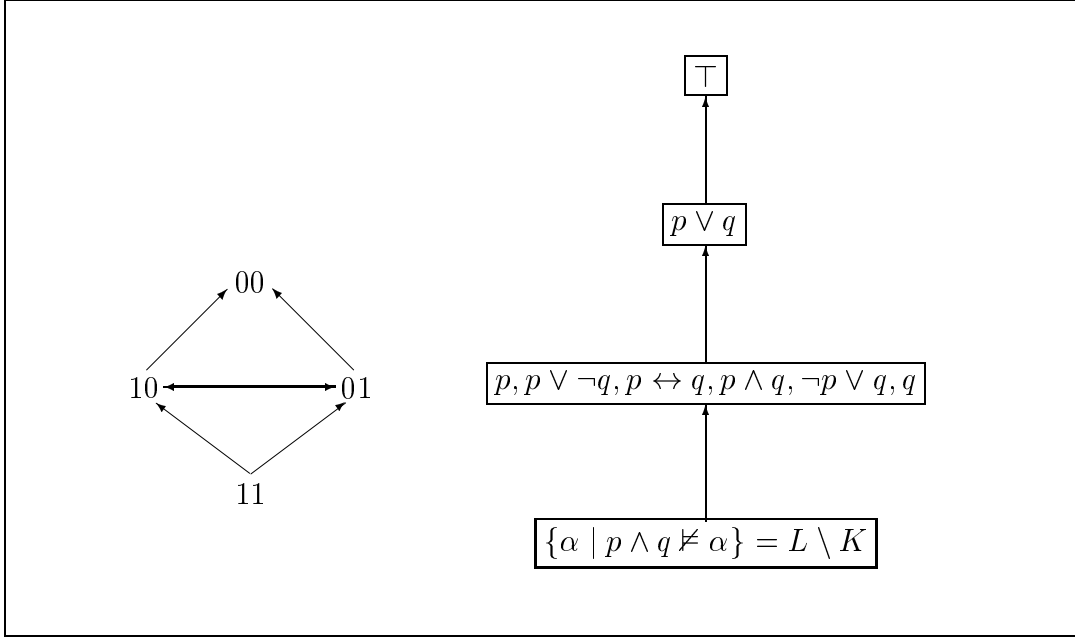


Figure 6.4: A graphical representation of the EE-ordering \sqsubseteq_{EE} with respect to the belief set $K = Cn(\{p, q\})$, and the faithful total preorder from which it is obtained using (Def \sqsubseteq_E from \preceq). These orderings are used in example 6.5.4. For every two interpretations x and y , $x \preceq y$ iff (x, y) is in the reflexive transitive closure of the relation determined by the arrows. Similarly, for every $\alpha, \beta \in L$, $\alpha \sqsubseteq_{EE} \beta$ iff (α, β) is in the reflexive transitive closure of the relation determined by the arrows. Each wff in the graphical representation of the EE-ordering is a canonical representative of the set of wffs which are logically equivalent to it.

Interestingly enough, this example does not represent a phenomenon that is unique to systematic withdrawal. The next proposition shows that systematic withdrawal and AGM contraction differ only on those wffs that are less entrenched than the wff α to be withdrawn. In other words, the type of problem highlighted in example 6.5.4 is one that has been carried over from AGM contraction.

Proposition 6.5.5 *Suppose that the EE-ordering \sqsubseteq_{EE} , the systematic withdrawal \div , and the AGM contraction $-$ are semantically related. If $\beta \not\sqsubseteq_{EE} \alpha$ then $\beta \in K - \alpha$ iff $\beta \in K \div \alpha$.*

Proof Let \preceq be a faithful total preorder in terms of which \sqsubseteq_{EE} and $-$ can be defined

using (Def \sqsubseteq_E from \preceq) and (Def \sim from \preceq), and let \leq be its semantically related faithful modular weak partial order. Suppose that $\beta \not\sqsubseteq_{EE} \alpha$. The right-to-left direction follows from corollary 6.3.16. Now suppose that $\beta \in K - \alpha$. Pick any $x \in \nabla_{\leq}(\neg\alpha)$. If $x \in M(K) \cup \text{Min}_{\leq}(\neg\alpha)$ then $x \in M(\beta)$, so we suppose otherwise. Then $x \in M(\alpha)$ and $x \prec y$ for every $y \in \text{Min}_{\leq}(\neg\alpha)$. Assume that $x \notin M(\beta)$. Then $x \in M(\alpha \wedge \neg\beta)$, and since $\alpha \sqsubseteq_{EE} \beta$, there is a $z \preceq x$ such that $z \in M(\neg\alpha)$, contradicting the fact that $x \prec y$ for every y in $\text{Min}_{\leq}(\neg\alpha)$. \square

Example 6.5.4 gives an indication that the EE-orderings have too coarse a grainsize to provide an intuitively satisfactory description of systematic withdrawal. This undesirable behaviour can be traced back to the fact that the EE-orderings are total preorders; a feature that has already been discussed at length in chapter 5.

6.5.2 Systematic withdrawal and the RE-orderings

We now come to an alternative description of systematic withdrawal in terms of refined entrenchment. It turns out that refined entrenchment retains the intuitively desirable results of section 6.5.1, and eliminates the counterintuitive results associated with the EE-orderings described in that section. First, we show that the result of proposition 6.5.2 carries over to the RE-orderings.

Proposition 6.5.6 *Suppose that the RE-ordering \sqsubseteq_{RE} and the systematic withdrawal \div are semantically related. If $\not\sqsubseteq \alpha$ and $\beta \sqsubseteq_{RE} \alpha$ then $\beta \notin K \div \alpha$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\leq}). Suppose that $\not\sqsubseteq \alpha$ and $\beta \sqsubseteq_{RE} \alpha$. So, for every $y \in \text{Min}_{\leq}(\neg\alpha)$ there is an $x \in \text{Min}_{\leq}(\neg\beta)$ such that $x \leq y$. This means that $\nabla_{\leq}(\neg\alpha) \not\subseteq M(\beta)$, and therefore that $\beta \notin K \div \alpha$. \square

So during an α -withdrawal, systematic withdrawal does not just guarantee the removal of all the wffs that are less entrenched than α , but also those that are *as* entrenched as α . It remains to be seen what happens to the remaining wffs; those are *not* at most as entrenched as the wff α to be withdrawn. Note firstly that AGM contraction and systematic withdrawal treat these wffs in exactly the same manner.

Proposition 6.5.7 *Suppose that the RE-ordering \sqsubseteq_{RE} and the systematic withdrawal \div are semantically related. If $\beta \not\sqsubseteq_{RE} \alpha$ then $\beta \in K - \alpha$ iff $\beta \in K \div \alpha$.*

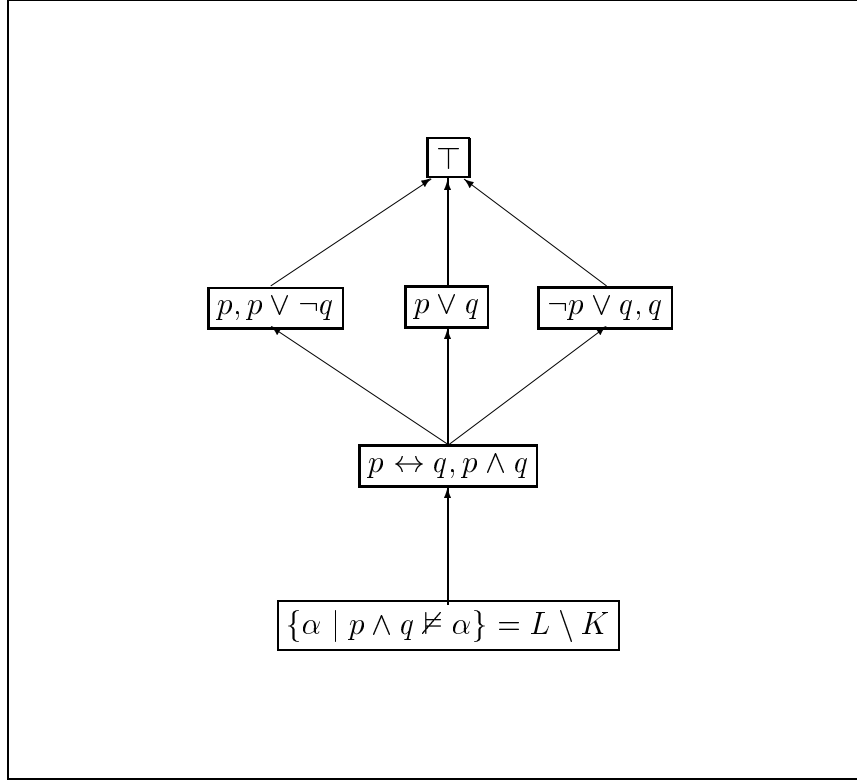


Figure 6.5: A graphical representation of the RE-ordering \sqsubseteq_{RE} with respect to the belief set $K = Cn(\{p \wedge q\})$. This RE-ordering is used in example 6.5.8. For every $\alpha, \beta \in L$, $\alpha \sqsubseteq_{RE} \beta$ iff (α, β) is in the reflexive transitive closure of the relation determined by the arrows. Each wff in this figure is a canonical representative of the set of wffs which are logically equivalent to it.

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \leq) and (Def \sim from ∇_{\leq}). Suppose that $\beta \not\sqsubseteq_{RE} \alpha$. By corollary 6.3.16 we already have that $K \div \alpha \subseteq K - \alpha$. So suppose that $\beta \in K - \alpha$. Then $M(K) \cup Min_{\leq}(\neg\alpha) \subseteq M(\beta)$ and it thus suffices to show that $\nabla_{\leq}(\neg\alpha) \setminus Min_{\leq}(\neg\alpha) \subseteq M(\beta)$. Now, since $\beta \not\sqsubseteq_{RE} \alpha$, there is a $y \in Min_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \leq y$. It then follows easily that $\nabla_{\leq}(\neg\alpha) \setminus Min_{\leq}(\neg\alpha) \subseteq M(\beta)$. \square

The next example shows that the wffs that are more entrenched than α are not always retained after a systematic α -withdrawal.

Example 6.5.8 Let L be the propositional language generated by the two atoms p and q and let (V, \models) be the valuation semantics for L , with $V = \{00, 01, 10, 11\}$. Now, let $K = Cn(\{p, q\})$, define the faithful modular weak partial order \leq as follows

$$x \leq y \text{ iff } \begin{cases} y \in V \text{ if } x = 11, \\ y \in \{01, 00\} \text{ if } x = 01, \\ y \in \{10, 00\} \text{ if } x = 10, \text{ and} \\ y = 00 \text{ if } x = 00, \end{cases}$$

and let \sqsubseteq_{RE} be the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq). Figure 6.2 contains a graphical representation of \leq , and figure 6.5 contains a graphical representation of \sqsubseteq_{RE} .

Now let \div be the systematic withdrawal defined in terms of \leq using (Def \sim from ∇_{\preceq}). It is easily verified that $K \div (p \leftrightarrow q) = Cn(p \vee q)$. Furthermore, an inspection of figure 6.5 shows that p , q , $\neg p \vee q$ and $p \vee \neg q$ are all *more* entrenched than $p \leftrightarrow q$. And yet, none of these wffs are in $K \div (p \leftrightarrow q)$. \square

An inspection of the RE-ordering \sqsubseteq_{RE} in figure 6.5 gives a clue as to why wffs that are more entrenched than α are sometimes not retained when performing an α -withdrawal. Observe in figure 6.5 that both p and q are more entrenched than $p \wedge q$. Retaining both of them in $K \div (p \wedge q)$ is out of the question (because it would then follow that $p \wedge q \in K \div (p \wedge q)$). Furthermore, \sqsubseteq_{RE} does not allow us to choose between p and q , since they are incomparable in terms of \sqsubseteq_{RE} . The prudent course of action is then to remove both. This argument can be formulated as a general principle involving sets of wffs.

Proposition 6.5.9 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. Now, suppose that $\alpha \sqsubseteq_{RE} \beta$, $X \cup \{\beta\} \models \alpha$, and both $\alpha \sqsubseteq_{RE} \gamma$ and $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ for every $\gamma \in X$. Then $\beta \notin K \div \alpha$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). Assume that $\beta \in K \div \alpha$ and pick a $\gamma \in X$. Since $\alpha \sqsubseteq_{RE} \gamma$, there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\gamma)$ for every $x \leq y$. Furthermore, $y \in M(\beta)$ because $\beta \in K \div \alpha$. Now, since $X \cup \{\beta\} \models \alpha$ and $y \in M(\neg\alpha \wedge \beta)$, there is a $\delta \in X$ such that $y \notin M(\delta)$. And because $\beta \parallel_{\sqsubseteq_{RE}} \delta$, there is a $v \in \text{Min}_{\leq}(\neg\beta)$ such that $u \in M(\delta)$ for every $u \leq v$. Therefore $y \not\leq v$. Furthermore,

from $\beta \in K \div \alpha$ it follows that $v \not\leq y$. And because $\alpha \sqsubseteq_{RE} \beta$, there is a $w \leq v$ such that $w \in M(\neg\alpha)$. Since $y \in \text{Min}_{\leq}(\neg\alpha)$ and $y \parallel_{\leq} v$, it has to be the case that $w = v$, and therefore $v \in \text{Min}_{\leq}(\neg\alpha)$. And since $v \notin M(\beta)$, $\beta \notin K \div \alpha$; a contradiction. \square

Note that proposition 6.5.9 only guarantees that β is not in $K \div \alpha$, and makes no such claim about the wffs in X as well, even though these wffs are all incomparable with β and more entrenched than α , just as β is. This can be explained by observing that the wffs in X need not be incomparable with one another. In the special case in which they *are* incomparable, it follows easily from proposition 6.5.9 that none of the wffs in X are in $K \div \alpha$ either.

A related result, and one that is of some importance for the results presented in the rest of this section, holds for the set of wffs that includes not only those that are more entrenched than α , but also those that are incomparable with α . For this result we need the following lemma.

Lemma 6.5.10 *Let \leq be a faithful modular weak partial order, \sqsubseteq_{RE} the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq), and \div the systematic withdrawal defined in terms of \leq using (Def \sim from ∇_{\leq}). If $\alpha \in K$, $\beta \not\sqsubseteq_{RE} \alpha$ and $\beta \notin K \div \alpha$, then there is a $y \in M(\neg\alpha \wedge \beta)$ and a $z \in M(\neg\alpha \wedge \neg\beta)$ such that $x \in M(\alpha \wedge \beta)$ for every $x < y$, and $x \in M(\alpha \wedge \beta)$ for every $x < z$.*

Proof Suppose that $\alpha \in K$, $\beta \not\sqsubseteq_{RE} \alpha$ and $\beta \notin K \div \alpha$. It follows from $\beta \not\sqsubseteq_{RE} \alpha$ that there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \leq y$. And therefore $y \in M(\neg\alpha \wedge \beta)$ and $x \in M(\alpha \wedge \beta)$ for every $x < y$. Furthermore, because $\beta \notin K \div \alpha$, there is a $z \in M(K) \cup \nabla_{\leq}(\neg\alpha)$ such that $z \in M(\neg\beta)$. If $z \notin \text{Min}_{\leq}(\neg\alpha)$ then, since $\alpha \in K$, $z < y$, which violates the result that all interpretations strictly below y are models of β . So $z \in \text{Min}_{\leq}(\neg\alpha)$, and because y is also a minimal model of $\neg\alpha$ it follows that $z \in M(\neg\alpha \wedge \neg\beta)$, and that $x \in M(\alpha \wedge \beta)$ for every $x < z$. \square

Proposition 6.5.11 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related.*

1. *If $\alpha \in K$, $\beta \not\sqsubseteq_{RE} \alpha$ and $\beta \notin K \div \alpha$, then there is a $\gamma \not\sqsubseteq_{RE} \alpha$ such that $\{\beta, \gamma\} \models \alpha$.*
2. *If $\beta \not\sqsubseteq_{RE} \alpha$, $\gamma \not\sqsubseteq_{RE} \alpha$, and $\{\beta, \gamma\} \models \alpha$, then $\beta \notin K \div \alpha$ and $\gamma \notin K \div \alpha$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). For the proof of (1), suppose that $\beta \not\sqsubseteq_{RE} \alpha$ and $\beta \notin K \div \alpha$. Now consider the wff $\beta \rightarrow \alpha$. It is clear that $\{\beta, \beta \rightarrow \alpha\} \models \alpha$. Since $\alpha \in K$, it follows from lemma 6.5.10 that there is a $y \in M(\neg\alpha \wedge \beta)$ and a $z \in M(\neg\alpha \wedge \neg\beta)$ such that $x \in M(\alpha \wedge \beta)$ for every $x < y$, and $x \in M(\alpha \wedge \beta)$ for every $x < z$. So z is a model of $\neg\alpha$ such that $x \in M(\beta \rightarrow \alpha)$ for every $x \leq z$. That is, $(\beta \rightarrow \alpha) \not\sqsubseteq_{RE} \alpha$, and we have the desired result.

To prove (2), suppose that $\beta \not\sqsubseteq_{RE} \alpha$, $\gamma \not\sqsubseteq_{RE} \alpha$ and $\{\beta, \gamma\} \models \alpha$. Because $\beta \not\sqsubseteq_{RE} \alpha$, there is a $y \in Min_{\leq}(\neg\alpha)$ such that $x \in M(\beta)$ for every $x \leq y$, and because $\{\beta, \gamma\} \models \alpha$, $y \in M(\neg\gamma)$. Similarly, from $\gamma \not\sqsubseteq_{RE} \alpha$ there is a $z \in Min_{\leq}(\neg\alpha)$ such that $x \in M(\gamma)$ for every $x \leq z$, and $z \in M(\neg\beta)$. It thus follows that $\gamma \notin K \div \alpha$ and $\beta \notin K \div \alpha$. \square

So proposition 6.5.11 tells us exactly which of the wffs that are *not* at most as entrenched as a wff α in K will be retained when withdrawing α from K , and which of these wffs will be discarded. It therefore places us in a position to formalise the relationship between the systematic withdrawals and the RE-orderings.

Theorem 6.5.12 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. Then*

$$\beta \notin K \div \alpha \text{ iff } \begin{cases} \beta \notin K \text{ and } \models \alpha, \text{ or} \\ \beta \notin K \text{ and } \alpha \notin K, \text{ or} \\ \beta \sqsubseteq_{RE} \alpha \text{ and } \not\models \alpha, \text{ or} \\ \beta \not\sqsubseteq_{RE} \alpha \text{ and } \exists \gamma \not\sqsubseteq_{RE} \alpha \text{ such that } \{\beta, \gamma\} \models \alpha, \end{cases} \quad (6.1)$$

or equivalently,

$$\beta \in K \div \alpha \text{ iff } \begin{cases} \beta \in K \text{ and } \models \alpha, \text{ or} \\ \beta \in K \text{ and } \alpha \notin K, \text{ or} \\ \beta \not\sqsubseteq_{RE} \alpha \text{ and for every } \gamma \not\sqsubseteq_{RE} \alpha, \{\beta, \gamma\} \not\models \alpha. \end{cases} \quad (6.2)$$

Proof The proof is mostly a combination of the results in propositions 6.5.11 and 6.5.6. It can be found in appendix B. \square

In fact, we can do better. The next proposition enables us to sharpen the relationship between systematic withdrawal and refined entrenchment.

Proposition 6.5.13 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related.*

1. If $\alpha \in K$, $\alpha \sqsubset_{RE} \beta$, and $\beta \notin K \div \alpha$, then $\alpha \sqsubset_{RE} \gamma$, $\gamma \parallel_{\sqsubseteq_{RE}} \beta$, and $\{\beta, \gamma\} \models \alpha$ for some $\gamma \in L$.
2. If $\alpha \in K$, $\alpha \parallel_{\sqsubseteq_{RE}} \beta$, and $\beta \notin K \div \alpha$, then $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$, $\gamma \parallel_{\sqsubseteq_{RE}} \beta$, and $\{\beta, \gamma\} \models \alpha$ for some $\gamma \in L$.¹⁴

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). The proofs are similar to those of part (1) of proposition 6.5.11. For the proof of (1), suppose that $\alpha \in K$, $\alpha \sqsubset_{RE} \beta$, and $\beta \notin K \div \alpha$. Now consider the wff $\beta \rightarrow \alpha$. It is clear that $\{\beta, \beta \rightarrow \alpha\} \models \alpha$. By lemma 6.5.10 there is a $y \in M(\neg\alpha \wedge \beta)$ and a $z \in M(\neg\alpha \wedge \neg\beta)$ such that $x \in M(\alpha \wedge \beta)$ for every $x < y$, and $x \in M(\alpha \wedge \beta)$ for every $x < z$. So $z \in M(\neg\alpha)$ and $x \in M(\beta \rightarrow \alpha)$ for every $x \leq z$. That is, $(\beta \rightarrow \alpha) \not\sqsubseteq_{RE} \alpha$. Since $M(\neg(\beta \rightarrow \alpha)) \subseteq M(\neg\alpha)$, it also clearly follows that $\alpha \sqsubseteq_{RE} (\beta \rightarrow \alpha)$, and so $\alpha \sqsubset_{RE} (\beta \rightarrow \alpha)$. To show that $(\beta \rightarrow \alpha) \parallel_{\sqsubseteq_{RE}} \beta$, note firstly that $z \in M(\neg\beta)$ and $x \in M(\beta \rightarrow \alpha)$ for every $x \leq z$. That is, $(\beta \rightarrow \alpha) \not\sqsubseteq_{RE} \beta$. And then note that $y \in M(\neg(\beta \rightarrow \alpha))$ and $x \in M(\beta)$ for every $x \leq y$. That is, $\beta \not\sqsubseteq_{RE} (\beta \rightarrow \alpha)$.

For the proof of (2), suppose that $\alpha \in K$, $\alpha \parallel_{\sqsubseteq_{RE}} \beta$, and $\beta \notin K \div \alpha$. Now consider the wff $\beta \leftrightarrow \alpha$. It is clear that $\{\beta, \beta \leftrightarrow \alpha\} \models \alpha$. By lemma 6.5.10, there is a $y \in M(\neg\alpha \wedge \beta)$ and a $z \in M(\neg\alpha \wedge \neg\beta)$ such that $x \in M(\alpha \wedge \beta)$ for every $x < y$ and $M(\alpha \wedge \beta)$ for every $x < z$. To show that $\beta \leftrightarrow \alpha \parallel_{\sqsubseteq_{RE}} \alpha$, note firstly $z \in M(\neg\alpha)$ and $x \in M(\beta \leftrightarrow \alpha)$ for every $x \leq z$. That is, $\beta \leftrightarrow \alpha \not\sqsubseteq_{RE} \alpha$. And then note that since $\alpha \not\sqsubseteq_{RE} \beta$, there is a $v \in M(\neg\beta)$ such that $u \in M(\alpha)$ for every $u \leq v$. Therefore $v \parallel_{\leq} y$, and so $u \in M(\alpha \wedge \beta)$ for every $u < v$. So $v \in M(\neg(\beta \leftrightarrow \alpha))$ and $u \in M(\alpha)$ for every $u \leq v$. That is, $\alpha \not\sqsubseteq_{RE} \beta \leftrightarrow \alpha$. Then, to show that $\beta \leftrightarrow \alpha \parallel_{\sqsubseteq_{RE}} \beta$, note that $z \in M(\neg\beta)$ and $x \in M(\beta \leftrightarrow \alpha)$ for $x \leq z$. That is, $\beta \leftrightarrow \alpha \not\sqsubseteq_{RE} \beta$. And then observe that $y \in M(\neg(\beta \leftrightarrow \alpha))$ and $x \in M(\beta)$ for every $x \leq y$. That is, $\beta \not\sqsubseteq_{RE} (\beta \leftrightarrow \alpha)$. \square

We are now in a position to state the main result of this section.

¹⁴It can also be shown that if $\alpha \in K$, $\alpha \parallel_{\sqsubseteq_{RE}} \beta$ and $\beta \notin K \div \alpha$, then $\alpha \sqsubset_{RE} \gamma$, $\gamma \parallel_{\sqsubseteq_{RE}} \beta$, and $\{\beta, \gamma\} \models \alpha$ for some γ . The proof is essentially the same as for part (1) of the proposition. While such a result does not offer much insight from an epistemological point of view, it might be useful for computational purposes.

Theorem 6.5.14 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. Then*

$$\beta \notin K \div \alpha \text{ iff } \left\{ \begin{array}{l} \beta \notin K \text{ and } \models \alpha, \text{ or} \\ \beta \notin K \text{ and } \alpha \notin K, \text{ or} \\ \beta \sqsubseteq_{RE} \alpha \text{ and } \not\models \alpha, \text{ or} \\ \alpha \sqsubseteq_{RE} \beta \text{ and } \exists \gamma \in L \text{ such that} \\ \alpha \sqsubseteq_{RE} \gamma, \beta \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \{\beta, \gamma\} \models \alpha, \text{ or} \\ \alpha \parallel_{\sqsubseteq_{RE}} \beta \text{ and } \exists \gamma \in L \text{ such that} \\ \alpha \parallel_{\sqsubseteq_{RE}} \gamma, \beta \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \{\beta, \gamma\} \models \alpha, \end{array} \right. \quad (6.3)$$

or equivalently,

$$\beta \in K \div \alpha \text{ iff } \left\{ \begin{array}{l} \beta \in K \text{ and } \models \alpha, \text{ or} \\ \beta \in K \text{ and } \alpha \notin K, \text{ or} \\ \alpha \sqsubseteq_{RE} \beta \text{ and } \forall \gamma \in L \text{ such that} \\ \alpha \sqsubseteq_{RE} \gamma \text{ and } \beta \parallel_{\sqsubseteq_{RE}} \gamma, \{\beta, \gamma\} \not\models \alpha, \text{ or} \\ \alpha \parallel_{\sqsubseteq_{RE}} \beta \text{ and } \forall \gamma \in L \text{ such that} \\ \alpha \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \beta \parallel_{\sqsubseteq_{RE}} \gamma, \{\beta, \gamma\} \not\models \alpha. \end{array} \right. \quad (6.4)$$

Proof The proof is mostly a combination of the results in propositions 6.5.11, 6.5.6 and 6.5.13. It can be found in appendix B. \square

From theorem 6.5.14 it emerges that, barring the limiting cases where α is logically valid or not in K , a wff $\beta \in K$ will *only* be removed during a systematic α -withdrawal for one of the following reasons:

1. The wff β is at most as entrenched as α .
2. The wff β is irrelevant with respect to α (i.e. β is not comparable with α) but there is another wff γ , which is irrelevant with respect to both α and β , and whose inclusion in the resulting belief set together with β , will force us to include α as well.
3. The wff β is more entrenched than α but there is another wff γ , also more entrenched than α , and irrelevant with respect to β , whose inclusion in the resulting belief set together with β , will force us to include α as well.

6.5.3 Representing systematic withdrawal graphically

Sections 6.4 and 6.5.2 provide interesting formal relationships between systematic withdrawal and refined entrenchment, but they provide little insight into the possible use of the RE-orderings to actually perform systematic withdrawal. In this section we show that for the special case of the finitely generated propositional languages, it is possible to define a process for constructing systematic withdrawal from the RE-orderings. This is important for computational purposes, but it is also of some epistemological importance.

The first result shows that those wffs which happen to be at most as entrenched as some discarded wff will also be discarded.

Proposition 6.5.15 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. If $\beta \notin K \div \alpha$ then $\gamma \notin K \div \alpha$ for every $\gamma \sqsubseteq_{RE} \beta$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). Suppose that $\beta \notin K \div \alpha$ and pick any $\gamma \sqsubseteq_{RE} \beta$. So there is a $y \in M(K \div \alpha)$ such that $y \in M(\neg\beta)$. And since $\gamma \sqsubseteq_{RE} \beta$ there is an $x \in M(\neg\gamma)$ such that $x \leq y$. So $x \in M(K \div \alpha)$ and therefore $\gamma \notin K \div \alpha$. \square

In section 6.5.2 it was shown that systematic withdrawal requires a good reason for removing a wff from K during an α -withdrawal. The next result is similar, providing a different kind of justification for the removal of some of these wffs.

Proposition 6.5.16 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. If $\alpha \in K$, $\gamma \parallel_{\sqsubseteq_{RE}} \alpha$ and $\gamma \notin K \div \alpha$, then there is a $\beta \notin K \div \alpha$ such that $\alpha \sqsubset_{RE} \beta$ and $\gamma \sqsubseteq_{RE} \beta$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). Pick any $\alpha \in K$ and any γ such that $\gamma \parallel_{\sqsubseteq_{RE}} \alpha$ and $\gamma \notin K \div \alpha$. We show that $\alpha \sqsubset_{RE} \alpha \vee \gamma$, $\gamma \sqsubseteq_{RE} \alpha \vee \gamma$ and $\alpha \vee \gamma \notin K \div \alpha$. Since $M(\alpha) \subseteq M(\alpha \vee \gamma)$ and $M(\gamma) \subseteq M(\alpha \vee \gamma)$, it immediately follows that $\alpha \sqsubseteq_{RE} \alpha \vee \gamma$ and $\gamma \sqsubseteq_{RE} \alpha \vee \gamma$. Since $\gamma \parallel_{\sqsubseteq_{RE}} \alpha$, there is a $y \in \text{Min}_{\leq}(\neg\alpha)$ such that $x \in M(\gamma)$ for every $x \leq y$, and there is a $v \in \text{Min}_{\leq}(\neg\gamma)$ such that $u \in M(\alpha)$ for every $u \leq v$. So $y \parallel_{\leq} v$, $x \in M(\gamma)$ for every $x < y$, and $u \in M(\alpha)$ for every $u < v$. Because $\gamma \notin K \div \alpha$, it then follows that there is a $z \in \text{Min}_{\leq}(\neg\alpha) \cap M(\neg\gamma)$. And this

means that $\alpha \vee \gamma \notin K \div \alpha$. Finally, note that for every $x \leq y$, $x \in M(\gamma)$ and thus $x \in M(\alpha \vee \gamma)$. Therefore $\alpha \vee \gamma \not\sqsubseteq_{RE} \alpha$, and then $\alpha \sqsubset_{RE} \alpha \vee \gamma$ (because $\alpha \sqsubseteq_{RE} \alpha \vee \gamma$). \square

Our next result relates specifically to the wffs that are *more* entrenched than a wff α to be withdrawn, but that are nevertheless removed from K . It shows that, for the finitely generated propositional languages, the structure of an RE-ordering can be used in a natural way to find these wffs. To do so, we need the notion of a closest upper gate with respect to a preorder on wffs.

$$(\text{Def } \text{cug}_{\sqsubseteq}) \quad \beta \in \text{cug}_{\sqsubseteq}(\alpha) \text{ iff } \begin{cases} 1. \alpha \sqsubset \beta, \\ 2. \forall \gamma \text{ such that } \alpha \sqsubset \gamma, \gamma \sqsubseteq \beta \text{ or } \beta \sqsubseteq \gamma, \text{ and} \\ 3. \forall \gamma \text{ such that } \alpha \sqsubset \gamma \sqsubset \beta, \\ \quad \exists \delta \text{ such that } \alpha \sqsubset \delta \sqsubset \beta \text{ and } \gamma \parallel_{\sqsubseteq} \delta \end{cases}$$

Definition 6.5.17 Let \sqsubseteq be a preorder on L . The *closest upper gate* $\text{cug}_{\sqsubseteq}(\alpha)$ of a wff α , with respect to \sqsubseteq , is defined in terms of \sqsubseteq using (Def cug_{\sqsubseteq}). \square

Roughly speaking, the closest upper gate of a wff α (with respect to a preorder \sqsubseteq on wffs) is the first equivalence class (modulo \sqsubseteq) of wffs encountered when moving “upwards” from α , which are not incomparable with respect to any of the wffs “above” α . We shall also have occasion to use the upset of a wff bounded by its closest upper gate.

$$(\text{Def } \diamond_{\sqsubseteq}) \quad \diamond_{\sqsubseteq}(\alpha) = \{\gamma \mid \alpha \sqsubset \gamma \sqsubset \beta \text{ for some } \beta \in \text{cug}_{\sqsubseteq}(\alpha)\}$$

Definition 6.5.18 Let \sqsubseteq be a preorder on L . The *upset* $\diamond_{\sqsubseteq}(\alpha)$ of a wff α *bounded by* $\text{cug}_{\sqsubseteq}(\alpha)$ is defined in terms of \sqsubseteq using (Def \diamond_{\sqsubseteq}). \square

The next lemma contains useful results about closest upper gates for RE-orderings. It shows that every closest upper gate (except for the empty set) is indeed an equivalence class modulo the RE-ordering, and that in the finitely generated propositional case, every wff, except for the logically valid ones, has a non-empty closest upper gate.

Lemma 6.5.19 *Let \sqsubseteq_{RE} be an RE-ordering.*

1. *If $\beta \in \text{cug}_{\sqsubseteq_{RE}}(\alpha)$ then $\text{cug}_{\sqsubseteq_{RE}}(\alpha) = [\beta]_{\sqsubseteq_{RE}}$.*¹⁵

¹⁵See section 1.3 for an explanation of the meaning of $[\beta]_{\sqsubseteq_{RE}}$.

2. If L is a finitely generated propositional language, then $cug_{\sqsubseteq_{RE}}(\alpha) = \emptyset$ iff $\models \alpha$.

Proof For the proof of (1), suppose that $\beta \in cug_{\sqsubseteq_{RE}}(\alpha)$, pick a $\gamma \in cug_{\sqsubseteq_{RE}}(\alpha)$, and assume that $\gamma \notin [\beta]_{\sqsubseteq_{RE}}$. From (1) and (2) in (Def cug_{\sqsubseteq}), $\beta \sqsubseteq_{RE} \gamma$ or $\gamma \sqsubseteq_{RE} \beta$, and therefore $\beta \sqsubset_{RE} \gamma$ or $\gamma \sqsubset_{RE} \beta$. We suppose that $\gamma \sqsubset_{RE} \beta$. By (3) in (Def cug_{\sqsubseteq}) there is thus a δ such that $\alpha \sqsubset_{RE} \delta$ and $\gamma \parallel_{\sqsubseteq_{RE}} \delta$. But since $\gamma \in cug_{\sqsubseteq_{RE}}(\alpha)$ this contradicts (2) in (Def cug_{\sqsubseteq}). A similar argument holds if $\beta \sqsubset_{RE} \gamma$, and so $cug_{\sqsubseteq_{RE}}(\alpha) \subseteq [\beta]_{\sqsubseteq_{RE}}$. Now pick any $\gamma \in [\beta]_{\sqsubseteq_{RE}}$. It then follows easily from (Def cug_{\sqsubseteq}) that $\gamma \in cug_{\sqsubseteq_{RE}}(\alpha)$, and so $[\beta]_{\sqsubseteq_{RE}} \subseteq cug_{\sqsubseteq_{RE}}(\alpha)$.

For the proof of (2), note firstly that if $\models \alpha$ then $\alpha \not\sqsubset_{RE} \gamma$ for every γ , and by (1) in (Def cug_{\sqsubseteq}), $cug_{\sqsubseteq_{RE}}(\alpha) = \emptyset$. On the other hand, suppose that $\not\models \alpha$ and assume that $cug_{\sqsubseteq_{RE}}(\alpha) = \emptyset$. We show that for every β satisfying (1) and (2) in (Def cug_{\sqsubseteq}), there is a $\gamma \sqsubset_{RE} \beta$ also satisfying (1) and (2) in (Def cug_{\sqsubseteq}), thus contradicting the fact that L is a finitely generated language. To do so, note firstly that every β that satisfies (1) and (2) in (Def cug_{\sqsubseteq}) does not satisfy (3). That is, for every β that satisfies (1) and (2) in (Def cug_{\sqsubseteq}), there is a γ such that $\alpha \sqsubset_{RE} \gamma \sqsubset_{RE} \beta$ and for every δ for which $\alpha \sqsubset_{RE} \delta \sqsubset_{RE} \beta$, either $\gamma \sqsubseteq_{RE} \delta$ or $\delta \sqsubseteq_{RE} \gamma$. So, if we can show that $\gamma \sqsubseteq_{RE} \varphi$ for every φ such that $\alpha \sqsubset_{RE} \varphi$ but $\varphi \not\sqsubset_{RE} \beta$, we will have shown that γ satisfies (1) and (2) in (Def cug_{\sqsubseteq}). Pick any φ such that $\alpha \sqsubset_{RE} \varphi$ but $\varphi \not\sqsubset_{RE} \beta$. Then either $\beta \sqsubseteq_{RE} \varphi$ or $\varphi \parallel_{\sqsubseteq_{RE}} \beta$. But since $\alpha \sqsubset_{RE} \varphi$, and since β satisfies (2) in (Def cug_{\sqsubseteq}), it cannot be the case that $\varphi \parallel_{\sqsubseteq_{RE}} \beta$. So $\beta \sqsubseteq_{RE} \varphi$, and since $\gamma \sqsubset_{RE} \beta$, it follows that $\gamma \sqsubset_{RE} \varphi$, which means we are done. \square

Before we can prove our next result, we need the following technical lemma.

Lemma 6.5.20 *Let L be a finitely generated propositional language, \leq a faithful modular weak partial order and \sqsubseteq_{RE} the RE-ordering defined in terms of \leq using (Def \sqsubseteq_E from \preceq) by \leq . Now let α and β be such that $M(\beta) = \{x \mid \forall y \in Min_{\leq}(\neg\alpha), y \not\leq x\}$. For every γ such that $\alpha \sqsubset_{RE} \gamma \sqsubset_{RE} \beta$, there is a $y \in M(\neg\alpha \wedge \beta \wedge \gamma)$ and a $z \in M(\neg\alpha \wedge \beta \wedge \neg\gamma)$ such that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < y$, and $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < z$.*

Proof If $\models \alpha$ the result holds vacuously, and we thus suppose that $\not\models \alpha$. Pick a γ such that $\alpha \sqsubset_{RE} \gamma \sqsubset_{RE} \beta$. Since $\gamma \not\sqsubseteq_{RE} \alpha$ there is a $y \in Min_{\leq}(\neg\alpha)$ such that $x \in M(\gamma)$ for every $x \leq y$. Combined with the definition of β it thus follows that $y \in M(\neg\alpha \wedge \beta \wedge \gamma)$ and $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < y$. Furthermore, since $\beta \not\sqsubseteq_{RE} \gamma$

there is a $z \in \text{Min}_{\leq}(\neg\gamma)$ such that $x \in M(\beta)$ for every $x \leq z$. And since $\alpha \sqsubseteq_{RE} \gamma$ there is a $w \in M(\neg\alpha)$ such that $w \leq z$. But then $w \in M(\beta)$ and so it follows from the definition of β that $w \in \text{Min}_{\leq}(\neg\alpha)$. Therefore $w = z$ and $v \in M(\alpha)$ for every $v < z$. So $z \in M(\neg\alpha \wedge \beta \wedge \neg\gamma)$, and because we have already seen that $x \in M(\gamma)$ for every $x \leq y$, it follows that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < z$. \square

The next result shows that for the finitely generated propositional languages, the wffs that are more entrenched than a wff α to be withdrawn, but that are not in the resulting belief set, are precisely those that lie between α and the closest upper gate of α .

Proposition 6.5.21 *Let L be a finitely generated propositional language and suppose that the RE -ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. If $\alpha \sqsubset_{RE} \gamma$ then $\gamma \notin K \div \alpha$ iff $\gamma \in \Diamond_{\sqsubseteq_{RE}}(\alpha)$.*

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). If $\models \alpha$ the result follows vacuously and so we suppose that $\not\models \alpha$. It suffices to show that for some $\beta \in \text{cug}_{\sqsubseteq_{RE}}(\alpha)$ and every γ such that $\alpha \sqsubset_{RE} \gamma$, $\gamma \notin K \div \alpha$ iff $\gamma \sqsubset_{RE} \beta$. Now pick any β such that $M(\beta) = \{x \mid \forall y \in \text{Min}_{\leq}(\neg\alpha), y \not\leq x\}$. Since L is a finitely generated propositional language, there is indeed such a β . We start by showing that $\beta \in \text{cug}_{\sqsubseteq_{RE}}(\alpha)$.

1. Pick any $y \in M(\neg\beta)$. By the definition of β , $z < y$ for every $z \in \text{Min}_{\leq}(\neg\alpha)$, and there is thus an $x \in M(\neg\alpha)$ such that $x \leq y$. Therefore $\alpha \sqsubseteq_{RE} \beta$. On the other hand, pick any $y \in \text{Min}_{\leq}(\neg\alpha)$. By the definition of β it follows that $x \in M(\beta)$ for every $x \leq y$, and thus $\beta \not\sqsubseteq_{RE} \alpha$.
2. Pick any $\gamma \in L$ such that $\alpha \sqsubset_{RE} \gamma$ and suppose that $\beta \not\sqsubseteq_{RE} \gamma$. So there is a $y \in M(\neg\gamma)$ such that $x \in M(\beta)$ for every $x \leq y$. Therefore $y \in M(\beta \wedge \neg\gamma)$. By the definition of β , $u < v$ for every $v \in M(\neg\beta)$ and $u \in M(\beta)$, and so $y < v$ for every $v \in M(\neg\beta)$. That is, for every $v \in M(\neg\beta)$ there is a $u \in M(\neg\gamma)$ such that $u \leq v$, which means that $\gamma \sqsubseteq_{RE} \beta$. So for every $\gamma \in L$ such that $\alpha \sqsubset_{RE} \gamma$, $\gamma \sqsubseteq_{RE} \beta$ or $\beta \sqsubseteq_{RE} \gamma$.
3. We show that $\alpha \sqsubset_{RE} \alpha \leftrightarrow \gamma \sqsubset_{RE} \beta$ and $\gamma \parallel_{\sqsubseteq_{RE}} \alpha \leftrightarrow \gamma$ for every γ such that $\alpha \sqsubset_{RE} \gamma \sqsubset_{RE} \beta$. Pick any $y \in M(\neg(\alpha \leftrightarrow \gamma))$. If $y \in M(\neg\alpha)$ then clearly there is an $x \in M(\neg\alpha)$ such that $x \leq y$. Otherwise $y \in M(\neg\gamma)$ and since $\alpha \sqsubseteq_{RE} \gamma$ there is an $x \in M(\neg\alpha)$ such that $x \leq y$. So $\alpha \sqsubseteq_{RE} \alpha \leftrightarrow \gamma$. On

the other hand, by lemma 6.5.20 there is a $z \in M(\neg\alpha \wedge \beta \wedge \neg\gamma)$ such that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < z$. So $z \in M(\neg\alpha)$ and for every $x \leq z$, $x \in M(\alpha \leftrightarrow \gamma)$. That is, $\alpha \leftrightarrow \gamma \not\sqsubseteq_{RE} \alpha$. And therefore $\alpha \sqsubseteq_{RE} \alpha \leftrightarrow \beta$. Pick any $w \in M(\neg\beta)$. By lemma 6.5.20 there is a $y \in M(\neg\alpha \wedge \beta \wedge \gamma)$ such that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < y$. By the definition of β it follows that $u < w$ for every $u \in M(\beta)$. So $y < w$ and because $y \in M(\neg(\alpha \leftrightarrow \gamma))$ it follows that there is an $x \in M(\neg(\alpha \leftrightarrow \gamma))$ such that $x \leq w$. That is, $\alpha \leftrightarrow \gamma \sqsubseteq_{RE} \beta$. On the other hand, $y \in M(\neg(\alpha \leftrightarrow \gamma))$ and for every $x \leq y$, $x \in M(\beta)$. That is, $\beta \not\sqsubseteq_{RE} \alpha \leftrightarrow \gamma$. And therefore $\alpha \leftrightarrow \beta \sqsubseteq_{RE} \beta$. It remains to be shown that $\gamma \parallel_{\sqsubseteq_{RE}} \alpha \leftrightarrow \gamma$. By lemma 6.5.20 there is a $y \in M(\neg\alpha \wedge \beta \wedge \gamma)$ such that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < y$. So $y \in M(\neg(\alpha \leftrightarrow \gamma))$ and for every $x \leq y$, $x \in M(\gamma)$. That is, $\gamma \not\sqsubseteq_{RE} \alpha \leftrightarrow \gamma$. Furthermore, by lemma 6.5.20 there is a $z \in M(\neg\alpha \wedge \beta \wedge \neg\gamma)$ such that $x \in M(\alpha \wedge \beta \wedge \gamma)$ for every $x < z$. So $z \in M(\neg\gamma)$ and for every $x \leq z$, $x \in M(\alpha \leftrightarrow \gamma)$. That is, $\alpha \leftrightarrow \gamma \not\sqsubseteq_{RE} \gamma$. So, for every $\gamma \in L$ such that $\alpha \sqsubseteq_{RE} \gamma \sqsubseteq_{RE} \beta$, $\alpha \sqsubseteq_{RE} \alpha \leftrightarrow \gamma \sqsubseteq_{RE} \beta$ and $\gamma \parallel_{\sqsubseteq_{RE}} \alpha \leftrightarrow \gamma$.

We have thus shown that $\beta \in \text{cug}_{\sqsubseteq_{RE}}(\alpha)$. Now we show that for every γ such that $\alpha \sqsubseteq_{RE} \gamma$, $\gamma \notin K \div \alpha$ iff $\gamma \sqsubseteq_{RE} \beta$. Pick any γ such that $\alpha \sqsubseteq_{RE} \gamma$. For the left-to-right direction, suppose that $\gamma \notin K \div \alpha$. That is, there is a $z \in M(K \div \alpha)$ such that $z \in M(\neg\gamma)$. Since $M(K \div \alpha) \subseteq M(\beta)$, it follows that $x \in M(\beta)$ for every $x \leq z$, and thus that $\beta \not\sqsubseteq_{RE} \gamma$. Now pick any $y \in M(\neg\beta)$. By the definition of β , $z < y$. And since $z \in M(\neg\gamma)$, there is an $x \in M(\neg\gamma)$ such that $x \leq y$. So $\gamma \sqsubseteq_{RE} \beta$. For the right-to-left direction, suppose that $\gamma \sqsubseteq_{RE} \beta$. Since $\beta \not\sqsubseteq_{RE} \gamma$ there is a $y \in M(\neg\gamma)$ such that $x \in M(\beta)$ for every $x \leq y$. So $y \in M(\beta)$ and by the definition of β , $v \in M(\alpha)$ for every $v < y$. And since $\alpha \sqsubseteq_{RE} \gamma$, there is a $w \in M(\neg\alpha)$ such that $w \leq y$, from which it then follows that $w = y$ and thus $y \in M(\neg\alpha)$. So $y \in M(\neg\alpha \wedge \beta \wedge \neg\gamma)$ and by the definition of β , $y \in \text{Min}_{\leq}(\neg\alpha)$. Therefore $\gamma \notin K \div \alpha$. \square

We are now in a position to prove the main result of this section.

Theorem 6.5.22 *Let L be a finitely generated propositional language and suppose that the RE-ordering \sqsubseteq_{RE} and the systematic withdrawal \div are semantically related. Then*

$$\beta \notin K \div \alpha \text{ iff } \begin{cases} \beta \notin K \text{ and either } \alpha \notin K \text{ or } \models \alpha, \text{ or} \\ \text{there is a } \gamma \in \Diamond_{\sqsubseteq_{RE}}(\alpha) \cup \{\alpha\} \text{ such that } \beta \sqsubseteq_{RE} \gamma. \end{cases}$$

Proof Let \leq be a faithful modular weak partial order in terms of which \sqsubseteq_{RE} and \div can be defined using (Def \sqsubseteq_E from \preceq) and (Def \sim from ∇_{\preceq}). For the proof in the left-to-right direction, suppose that $\beta \notin K \div \alpha$, and assume that either $\beta \in K$, or $\alpha \in K$ and $\not\models \alpha$. If $\beta \in K$ then $K \neq K \div \alpha$ (because $\beta \notin K \div \alpha$), and it thus follows from (K \div 3) and (K \div 6) that $\alpha \in K$ and $\not\models \alpha$. So in both cases, $\alpha \in K$ and $\not\models \alpha$. Now, if $\beta \sqsubseteq_{RE} \alpha$ then there is indeed a $\gamma \in \Diamond_{\sqsubseteq_{RE}}(\alpha) \cup \{\alpha\}$ such that $\beta \sqsubseteq_{RE} \gamma$. So we suppose that $\beta \not\sqsubseteq_{RE} \alpha$. This means that either $\beta \parallel_{\sqsubseteq_{RE}} \alpha$ or $\alpha \sqsubset_{RE} \beta$. In the latter case it follows from proposition 6.5.21 that $\beta \in \Diamond_{\sqsubseteq_{RE}}(\alpha)$. In the former case it follows from proposition 6.5.16 that there is a $\gamma \notin K \div \alpha$ such that $\alpha \sqsubset_{RE} \gamma$ and $\beta \sqsubseteq_{RE} \gamma$. And by proposition 6.5.21, $\gamma \in \Diamond_{\sqsubseteq_{RE}}(\alpha)$.

For the proof in the right-to-left direction, note firstly that if $\alpha \notin K$ and $\beta \notin K$ then $\beta \notin K \div \alpha$ by (K \div 3), and if $\models \alpha$ and $\beta \notin K$ then $\beta \notin K \div \alpha$ by (K \div 6). So we suppose that $\not\models \alpha$, $\alpha \in K$ and that there is a $\gamma \in \Diamond_{\sqsubseteq_{RE}}(\alpha) \cup \{\alpha\}$ such that $\beta \sqsubseteq_{RE} \gamma$. If $\gamma = \alpha$, it follows from proposition 6.5.6 that $\beta \notin K \div \alpha$. Otherwise $\alpha \sqsubset_{RE} \gamma \sqsubset_{RE} \text{cug}_{\sqsubseteq_{RE}}(\alpha)$. By proposition 6.5.21, $\gamma \notin K \div \alpha$ and by proposition 6.5.15 we then have that $\beta \notin K \div \alpha$. \square

Theorem 6.5.22 provides us with the following description of systematic withdrawal in terms of refined entrenchment. Consider the non-trivial case where α is in K , but is not logically valid. To obtain the belief set resulting from an α -withdrawal, we partition L into three sets; those that are at most as entrenched α , those that are more entrenched than α , and those that are incomparable with α .

1. None of the wffs that are at most as entrenched as α are in $K \div \alpha$.
2. The wffs that are more entrenched than α , but that aren't in $K \div \alpha$ are precisely those that are between α and the closest upper gate of α . These wffs are clustered right above α , and are strictly less entrenched than the wffs above α that are in $K \div \alpha$.
3. The only wffs that are incomparable with α , but that are *not* in $K \div \alpha$, are those that are less entrenched than one of the wffs which are removed from K even though it is more entrenched than α .

A particularly attractive feature obtained from this analysis is that the wffs that aren't in the resulting belief set when withdrawing α from K , are all clustered together in

the refined entrenchment ordering. We conclude this section with a brief description of how the results of theorem 6.5.22 can be used to define a graph-based procedure for computing systematic withdrawal. The basic idea is to view every refined entrenchment ordering \sqsubseteq_{RE} as a directed acyclic graph (DAG), with the equivalence classes of wffs (modulo \sqsubseteq_{RE}) as the vertices of the DAG, and with the arcs between vertices obtained from \sqsubseteq_{RE} . Consider the non-trivial case where $\alpha \in K \setminus Cn(\top)$. The wffs *not* in $K \div \alpha$ are obtained from the DAG associated with \sqsubseteq_{RE} as follows: Start from the vertex v containing α and follow all the paths leading out of v to the first vertex w where these paths all meet. The vertex w contains $cug_{\sqsubseteq_{RE}}(\alpha)$. Now consider all the vertices that were visited before reaching w (including v but excluding w) and do a backward traversal of the paths leading into these vertices. The wffs not in $K \div \alpha$ are precisely those contained in the vertices visited on these backward traversals. This process is made concrete in the following example.

Example 6.5.23 Let L be the propositional language generated by the two atoms p and q , let $K = Cn(p \wedge q)$, and let \sqsubseteq_{RE} be RE-ordering defined as follows:

$$\alpha \sqsubseteq_{RE} \beta \text{ iff } \begin{cases} \beta \in L \text{ if } \alpha \notin K, \text{ and} \\ \alpha \models \beta \text{ if } \alpha \in K. \end{cases}$$

Figure 6.6 contains a graphical representation of \sqsubseteq_{RE} . Let \div be the systematic withdrawal defined in terms of \sqsubseteq_{RE} using (Def \div from \sqsubseteq_{RE}). It can be verified that $K \div (p \leftrightarrow q) = Cn(p \vee q)$. This result can also be obtained by viewing figure 6.6 as a DAG. We start from the vertex v containing $p \leftrightarrow q$ and follow all the paths leading out of v until we reach the first vertex w where all these paths meet. The vertex w is the one containing the wff \top , and it also contains all the wffs in $cug_{\sqsubseteq_{RE}}(p \leftrightarrow q)$. The vertices visited before reaching w are the vertex v itself and the vertices containing the wffs $p \vee \neg q$ and $\neg p \vee q$. Now we do a backward traversal of all the paths leading into these three vertices. The vertices visited on these backward traversals are the boxed ones. The only remaining vertices are w (which contains the wff \top) and the vertex containing $p \vee q$. So the wffs in $K \div \alpha$ are precisely those that are logically equivalent to $p \vee q$ and to \top . \square

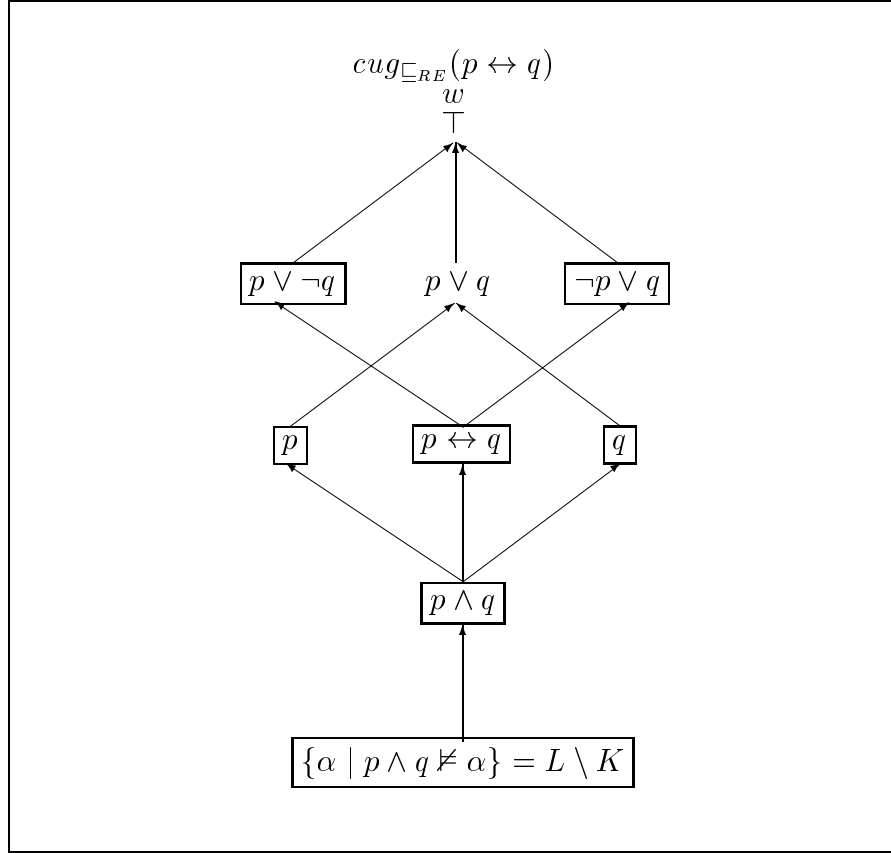


Figure 6.6: A graphical representation of the RE-ordering \subseteq_{RE} with respect to the belief set $K = Cn(p \wedge q)$. This RE-ordering is used in example 6.5.23. For every $\alpha, \beta \in L$, $\alpha \subseteq_{RE} \beta$ iff (α, β) is in the reflexive transitive closure of the relation determined by the arrows. Each wff in this figure is a canonical representative of the set of wffs that are logically equivalent to it. This graphical representation can also be seen as the directed acyclic graph (DAG) obtained from \subseteq_{RE} .

6.6 Summary

We close this chapter with a graphical summary of the connections between various forms of principled withdrawal and entrenchment orderings. It can be found in figure 6.7. As in previous chapters, it is difficult to escape the conclusion that they all have a semantic basis, and more particularly, are rooted in some subset of the faithful layered preorders. Finally, these semantic constructions bear testimony to two important principles, one of which was already noted by Rott and Pagnucco [1999, p. 33]. The

underlying semantic structures employed to define a particular form of withdrawal or entrenchment are obviously important. But whether two different structures define the same construction depends, to some extent, on the way these structures are used. And conversely, the same structure might very well be used in a number of different ways, depending on the precedence given to different principles, resulting in distinctly different constructions.

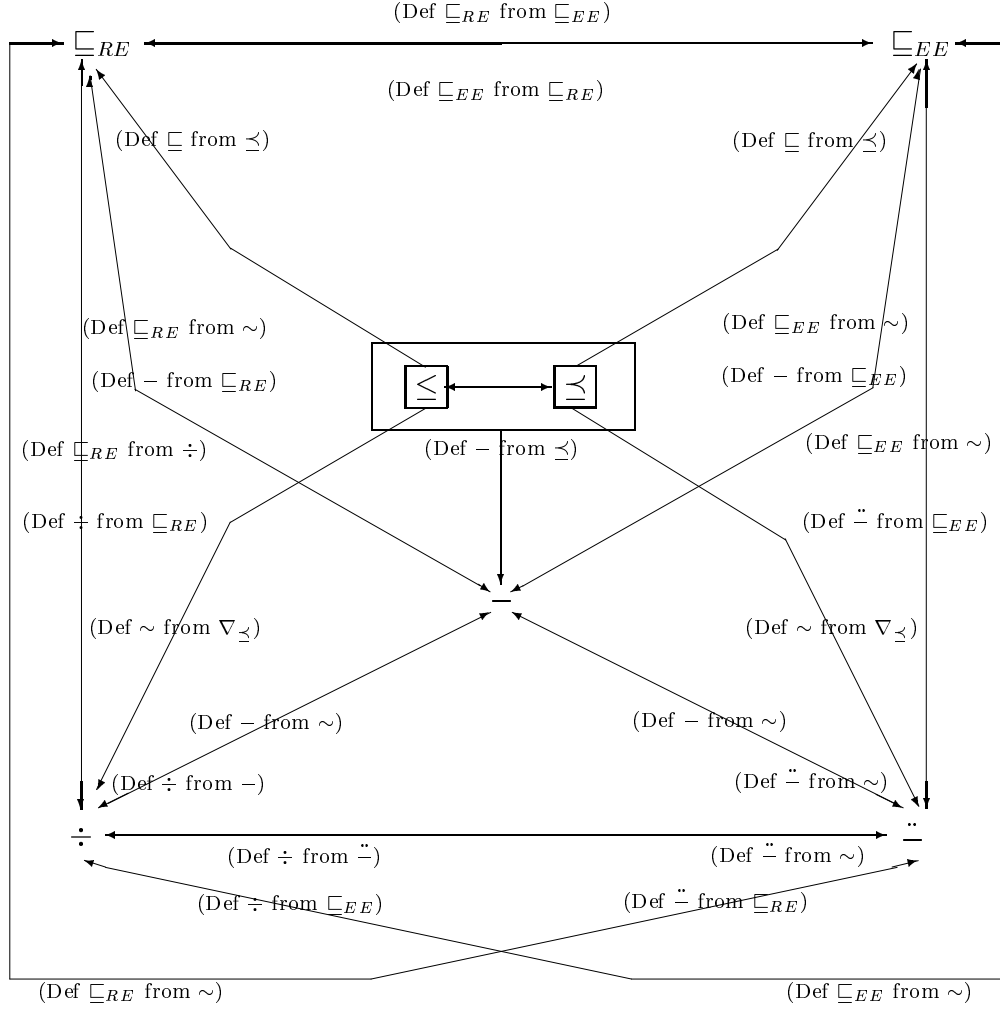


Figure 6.7: The relationship between minimal-equivalent faithful layered preorders, AGM contraction, the EE-orderings, the RE-orderings, systematic withdrawal and severe withdrawal.

Chapter 7

Iterated belief change

*After people have repeated a phrase a great number of times,
they begin to realize it has meaning and may even be true.*

H.G. Wells (1866-1946)

AGM theory change has proved to be very useful as an abstract account of effecting change in the epistemic state of an agent. As such, it provides a good platform from which to launch investigations into aspects of belief change which are not dealt with in the AGM framework. Makinson [1997] states this viewpoint as follows:

“But it is through such simple, idealized representations of belief sets that we have begun to obtain the insights needed to tackle more complex ones without getting lost in intricacies and overheads. Having acquired a fairly good understanding of the former over the decade since the AGM account appeared in 1985, we can now profitably give more attention to the latter.”

In the light of this statement, it should come as no surprise that investigations into various extensions of AGM theory change have become more frequent in recent years. *Iterated belief change*, the problem of dealing with a sequence of changes to the epistemic state of an agent, is an aspect of belief change which falls into this category, and is the focus of this chapter. Since most recent advances in this area have focused on finitely generated propositional languages, we shall, for the rest of this chapter, assume L to be such a language with a valuation semantics (V, \models) . We discuss the general frameworks for iterated revision provided by Williams [1994], Darwiche and Pearl

[1994, 1997], and Lehmann [1995], and take a brief look at a framework for iterated withdrawal. Furthermore, we consider revision operations proposed by Boutilier [1993, 1996], Williams [1994] and Papini [1998, 1999].

One of the most important lessons to be learnt from the study of iterated belief change is that belief change operations are performed on the level of epistemic states, and not belief sets. Inspired by this insight, we discuss a generalised version of revision in which epistemic states are *merged*. We propose some basic properties for such merging operations, and consider a few particular merging operations. Amongst those we consider are Nayak's version of iterated revision [1994b, 1996] and the framework for arbitration proposed by Liberatore and Schaerf [1998].

7.1 Transmutation

Recent advances in iterated belief change have benefitted substantially from ideas initially proposed by Spohn [1988, 1991], and generalised by Williams [1994]. It is thus appropriate that we commence with a discussion of these. Williams [1994] proposes a framework for belief change based on Spohn's ordinal conditional functions (see section 5.3). It is a generalisation of withdrawal and revision in two respects.

Firstly, the informational inputs are not wffs, but ordered pairs of the form (α, n) , where $\alpha \in L \setminus \{\beta \mid \beta \equiv \perp \text{ or } \beta \equiv \top\}$ and n is a natural number. To be more precise, Williams proposes to use pairs of the form (W, n) , where $\emptyset \subset W \subset V$ and n is an ordinal, but since we assume L to be a finitely generated propositional language, every such a W is axiomatisable by a single wff (which is satisfiable but not logically valid). Furthermore, since V is finite, we restrict ourselves to those OCFs with ranges consisting of subsets of ω , the set of natural numbers.¹

Secondly, transmutations are operations on ordinal conditional functions (OCFs), and not on belief sets. Recall from section 5.3 that, for an OCF κ , K_κ denotes the set $Th(\{v \mid \kappa(v) = 0\})$, and that K_κ can therefore be regarded as the belief set associated with κ . Furthermore, recall that an OCF assigns the number 0 to at least one valuation, from which it follows that K_κ is satisfiable. And moreover, recall that

¹Having made these simplifications, it is tempting to augment the domain of OCFs to include the empty set as well, and set $\kappa(\emptyset) = \omega$ for every OCF κ . In such a case we would be dealing with Spohn's [1991] *natural conditional functions*. However, for the sake of simplicity we shall resist this temptation.

a wff in K_κ is said to be believed with a firmness of n iff $\kappa(M(\neg\alpha)) = n$.

Williams defines a *transmutation* of the OCF κ as any function

$$* : (L \setminus (Cn(\top) \cup \{\beta \mid \beta \equiv \perp\})) \times \omega \rightarrow \mathcal{K},$$

where \mathcal{K} is the set of all OCFs (with ranges consisting of subsets of ω) such that:

1. $(\kappa * (\alpha, n))(M(\neg\alpha)) = n$
2. $K_{\kappa * (\alpha, n)} = \begin{cases} Th(\{v \in M(\alpha) \mid \kappa(v) = \kappa(\alpha)\}) & \text{if } n > 0, \\ Th(\{v \mid \kappa(v) = 0 \text{ or } (v \in M(\neg\alpha) \text{ and } \kappa(v) = \kappa(\neg\alpha))\}) & \text{otherwise} \end{cases}$

So a transmutation of the current OCF κ by (α, n) yields a new OCF κ' in which α is believed with a firmness of n . Furthermore, if $n > 0$, we can think of a transmutation as a revision, while it can be regarded as a contraction if $n = 0$. This view is justified by noting that $K_{\kappa'}$, the belief set associated with κ' , is generated by the minimal models of α (with regard to κ) if $n > 0$, and by the minimal models of α , together with the models of K_κ , if $n = 0$.

Williams considers the construction of two transmutations. The first one is Wolfgang Spohn's [1988] *conditionalisation*, which has turned out to be a particularly influential contribution to the enterprise of iterated belief change. Indeed, as mentioned above, transmutation was proposed as a generalisation of Spohn's conditionalisation, and the latter has also served as inspiration for the general framework of Darwiche and Pearl discussed in section 7.3. The OCF $\kappa * (\alpha, n)$, referred to as the (α, n) -*conditionalisation* of κ , is defined as follows:

$$(\mathbf{Def} * \mathbf{from} \kappa) \kappa * (\alpha, n)(v) = \begin{cases} \kappa(v) - \kappa(M(\alpha)) & \text{if } v \in M(\alpha), \\ \kappa(v) - \kappa(M(\neg\alpha)) + n & \text{otherwise} \end{cases}$$

In other words, the models of α are shifted “downwards” without affecting the distances between them, so that the minimal models of α are assigned the number 0, while the models of $\neg\alpha$ are shifted “upwards” without affecting the distances between them, so that the minimal models of $\neg\alpha$ are assigned the number n . It is easily established that for a fixed OCF $\kappa \in \mathcal{K}$ and a fixed $n > 0$, the revision $*$ defined in terms of κ using (Def $*$ from κ) below is an AGM revision:²

²As Darwiche and Pearl [1997, p. 15] have noted though, viewing a revision $*$ as a function from $Bel \times L$ to Bel means that (Def $*$ from κ) may yield different results for different OCFs corresponding to the same belief set, thus violating the functionality of $*$.

$$(\mathbf{Def} * \mathbf{from} \kappa) \left[\begin{array}{l} K_{\kappa} * \alpha = \begin{cases} K_{\kappa} & \text{if } \models \alpha, \\ K_{\kappa * (\alpha, n)} & \text{if } \alpha \not\equiv \perp \text{ and } \not\models \alpha, \\ Cn(\perp) & \text{otherwise} \end{cases} \\ Cn(\perp) * \alpha = K_{\kappa} * \alpha \end{array} \right]$$

Similarly, for a fixed OCF $\kappa \in \mathcal{K}$, the removal \sim defined in terms of κ using (Def \sim from κ) below is an AGM contraction:

$$(\mathbf{Def} \sim \mathbf{from} \kappa) \left[\begin{array}{l} K_{\kappa} \sim \alpha = \begin{cases} K_{\kappa} & \text{if } \models \alpha \text{ or } \alpha \equiv \perp \\ K_{\kappa * (\alpha, 0)} & \text{otherwise} \end{cases} \quad \text{and} \\ Cn(\perp) \sim \alpha = K_{\kappa * (\alpha, 1)} \end{array} \right]$$

The second transmutation that Williams considers is known as *adjustment*. The basic idea is that an (α, n) -adjustment should be the transmutation that leaves the current OCF as undisturbed as possible; an appeal to the Principle of Minimal Change. The (α, n) -*adjustment* \star of the OCF κ is defined as follows:

$$(\mathbf{Def} \star \mathbf{from} \kappa) \kappa \star (\alpha, n)(v) = \begin{cases} 0 & \text{if } n = 0, v \in M(\neg\alpha), \text{ and } \kappa(v) = \kappa(\neg\alpha), \\ \kappa(v) & \text{if } n = 0 \text{ and } (v \in M(\alpha) \text{ or } \kappa(v) \neq \kappa(\neg\alpha)), \\ 0 & \text{if } 0 < n, v \in M(\alpha), \text{ and } \kappa(v) = \kappa(\alpha), \\ \kappa(v) & \text{if } 0 < n, v \in M(\alpha), \text{ and } \kappa(v) \neq \kappa(\alpha), \\ n & \text{if } 0 < n < \kappa(\neg\alpha), v \in M(\neg\alpha), \\ & \text{and } \kappa(v) = \kappa(\neg\alpha), \\ \kappa(v) & \text{if } 0 < n < \kappa(\neg\alpha), v \in M(\neg\alpha), \\ & \text{and } \kappa(v) \neq \kappa(\neg\alpha), \\ n & \text{if } 0 < n, n \geq \kappa(\neg\alpha), v \in M(\neg\alpha), \\ & \text{and } \kappa(v) \leq n, \\ \kappa(v) & \text{if } 0 < n, n \geq \kappa(\neg\alpha), v \in M(\neg\alpha), \\ & \text{and } \kappa(v) > n \end{cases}$$

This definition looks quite complicated, but it can be broken down into three mutually exclusive cases:

1. If $n = 0$, the only difference between κ and $\kappa \star (\alpha, n)$ is that the minimal models of $\neg\alpha$ (with respect to κ), are all assigned the number 0 in $\kappa \star (\alpha, n)$.

2. If $n > 0$ and the number that κ assigns to the minimal models of $\neg\alpha$ (with respect to κ) is greater than n , then the only difference between κ and $\kappa \star (\alpha, n)$ is that the minimal models of $\neg\alpha$ (with respect to κ) are all assigned the number n in $\kappa \star (\alpha, n)$.
3. If $n > 0$ and the number that κ assigns to the minimal models of $\neg\alpha$ (with respect to κ) is less than or equal to n , then the only difference between κ and $\kappa \star (\alpha, n)$ is that the models of $\neg\alpha$ for which κ assigns numbers less than n are all assigned the number n in $\kappa \star (\alpha, n)$.

We shall see below that it only takes a small modification to apply the intuitions underlying conditionalisation and adjustment to iterated belief change based on AGM revision.

7.2 AGM and iterated belief change

It is by now widely accepted that AGM theory change is not able to deal with issues of iterated belief change in an adequate manner [Alchourrón and Makinson, 1985, Gärdenfors, 1988, Levi, 1988, Boutilier, 1993, 1996, Nayak, 1994b, Nayak et al., 1996]. This statement can be interpreted in at least two ways. In the *static* view adopted by Freund and Lehmann [1994], theory revision³ is described as an operation with two arguments; a belief set K , and a wff α with which to revise K .⁴ So the operation $*$ represents a process of revision which is fixed right from the start, so that an α -revision of a belief set K will always yield the same result, regardless of how an agent arrived at K . The static view therefore dooms an agent to picking, for every belief set K , a single epistemic state to associate with K , and to using only that epistemic state to guide its reasoning whenever its set of beliefs corresponds to K . Thus, for example, if the two belief sets K and $(K * \alpha) * \beta$ happen to be identical, it will always be the case that $K * \gamma = ((K * \alpha) * \beta) * \gamma$. It is, essentially, the postulate (K*5) which requires of us to treat iterated revision in such a static manner. In this view, a proper account of iterated revision is just the natural next step in the move from basic AGM theory revision to AGM theory revision. While basic AGM revision fixes both arguments of the revision

³Since most researchers restrict themselves to treatments of revision when it comes to iterated belief change, we shall do the same, for the most part.

⁴See also [Areces and Becher, 1998]

operation $*$, AGM revision (which also satisfies the supplementary postulates) fixes the first argument, the belief set, and allows the second argument, the wff with which to revise, to vary. And iterated revision is then seen as the next step, where the first argument is also allowed to vary. The static view thus advocates the introduction of additional postulates in the style of the AGM revision postulates in order to obtain an appropriate account of iterated revision.

There is a *dynamic* view of iterated revision as well, in which the revision process depends on more than just the belief set to be revised. In this view, the revision *procedure* used to revise the belief set $K * \alpha$ may very well differ from the one used when revising K . As a result, for example, $K * \gamma$ and $((K * \alpha) * \beta) * \gamma$ need not be identical when K and $(K * \alpha) * \beta$ are. Strictly speaking, this view is incompatible with AGM revision, and more particularly, with (K*5). But this is merely because the notation used in the AGM postulates does not reflect the fact that revision is an operation on epistemic sets, and not on belief sets (or stated differently, that belief sets do not have enough structure to serve as appropriate representations of epistemic states). And as we shall see, it only requires a slight reformulation of the AGM revision postulates to do away with the incompatibility brought on by (K*5).

Although the static view of revision might serve as a first approximation, it seems reasonable to conclude that a proper rational account of iterated revision can only be found by embracing the dynamic view, and more particularly, the move from revision as an operation on belief sets, to revision as an operation on epistemic states. That revision ought to be seen as an operation on epistemic states, becomes apparent when observing that AGM theory change is in clear violation of the principle of Categorical Matching. It (AGM theory change) delivers a belief set as a result of a change operation, but requires an epistemic state, consisting of a belief set together with some kind of selection mechanism (such as a faithful layered preorder) to perform these change operations. Furthermore, it is easy to construct examples demonstrating that two agents with different epistemic states containing the same belief set, will sometimes follow different revision strategies. Below we provide such an example, which is a slight modification of an example presented in [Darwiche and Pearl, 1997].⁵

Example 7.2.1 Two jurors in a murder trial possess different biases; both jurors

⁵The example of Darwiche and Pearl assumes that the two jurors have the same belief set, even though juror number one believes that C is definitely innocent and B might be guilty, while juror number two believes that B is definitely innocent while C might be guilty.

believe that A is the murderer, and both believe that only A, B or C could have committed the crime (a classic case of an alphabet murder mystery). But whereas juror number one would more easily regard C as the murdering type than B, juror number two would more easily consider B to be guilty than C. Both jurors thus have the same belief set. Some surprising evidence now comes to light; A has produced a reliable alibi. Clearly juror number one would now believe C to be the murderer, while juror number two would believe that B is the murderer. \square

Finally, it is worth observing that even when adopting the dynamic view, AGM revision is not completely noncommittal when it comes to iterated belief change. Indeed, by way of the two supplementary postulates (K*7) and (K*8), it does place *some* constraints on the way iterated theory revision may be performed, although these constraints are fairly mild. Observe that (K*3) and (K*4) ensure that an α -expansion and an α -revision are identical whenever $\neg\alpha$ is not in the belief set K . Consequently, as Freund and Lehmann [1994] have shown, the following is a property derived from (K*3), (K*4), (K*7) and (K*8).

(K*9) If $\neg\beta \notin K * \alpha$ then $(K * \alpha) * \beta = K * (\alpha \wedge \beta)$

So AGM theory revision, in the form of (K*9), provides us with a sufficient condition for insisting that the belief set resulting from a simultaneous revision of two wffs α and β (that is, an $\alpha \wedge \beta$ -revision) be identical to the belief set obtained from an α -revision followed by a β -revision.

7.3 Iterated DP-revision

In two influential recent papers Darwiche and Pearl [1994, 1997] have made an important contribution to the study of iterated belief change. Of particular significance is the shift they make in [Darwiche and Pearl, 1997] from revision as an operation on belief sets to an operation on epistemic states. Although they do not define the notion of an epistemic state explicitly, they work on the assumption that we can extract from every epistemic state a belief set $K(\Phi)$. Formally, they see a revision $*$ as a function from $\mathcal{E} \times L$ to \mathcal{E} , where \mathcal{E} is the set of all epistemic states. To accommodate the move to epistemic states, the AGM revision postulates are modified appropriately.⁶

⁶Actually, Darwiche and Pearl modify the postulates of Katsuno and Mendelson (see section 3.2.1), but our account here is the obvious translation to the AGM revision postulates.

$$(E*1) \quad K(\Phi * \alpha) = Cn(K(\Phi * \alpha))$$

$$(E*2) \quad \alpha \in K(\Phi * \alpha)$$

$$(E*3) \quad K(\Phi * \alpha) \subseteq K(\Phi) + \alpha$$

$$(E*4) \quad \text{If } \neg\alpha \notin K(\Phi), \text{ then } K(\Phi) + \alpha \subseteq K(\Phi * \alpha)$$

$$(E*5) \quad \text{If } \Phi = \Psi \text{ and } \alpha \equiv \beta \text{ then } K(\Phi * \alpha) = K(\Psi * \beta)$$

$$(E*6) \quad K(\Phi * \alpha) = Cn(\perp) \text{ iff } \models \neg\alpha$$

$$(E*7) \quad K(\Phi * \alpha \wedge \beta) \subseteq K(\Phi * \alpha) + \beta$$

$$(E*8) \quad \text{If } \neg\beta \notin K(\Phi * \alpha), \text{ then } K(\Phi * \alpha) + \beta \subseteq K(\Phi * \alpha \wedge \beta)$$

With the exception of (E*5), these postulates are just obvious translations of the corresponding AGM revision postulates. (E*5) is an appropriate weakening of (K*5). It requires a revision by two pieces of logically equivalent evidence to yield identical belief sets when the epistemic states to be revised are identical, and not merely when the belief sets contained in these epistemic states are identical. Note that (E*5) does not require a revision by two logically equivalent wffs to yield the same epistemic state; it only insists that the belief sets associated with these epistemic states be identical. This is quite surprising, especially since, in the words of Darwiche and Pearl [1997, p. 2], an epistemic state contains, in addition to the belief set, "...the entire information needed for coherent reasoning, including, in particular, the very strategy which the agent wishes to employ at that given time". Moreover, since (K*5) is a formal expression of the principle of the Irrelevance of Syntax, one would expect (E*5) to be an expression of the same principle in the more general context of revision on epistemic states. It thus seems as if the following postulate would have been more appropriate:

$$(E*9) \quad \text{If } \Phi = \Psi \text{ and } \alpha \equiv \beta \text{ then } \Phi * \alpha = \Psi * \beta$$

We shall not pursue this matter further, except to note that replacing (E*5) with (E*9) is compatible with the results in the remainder of this section.

Darwiche and Pearl provide a representation result that is analogous to theorem 3.2.6.

Theorem 7.3.1 *Suppose we associate with every epistemic state Φ , a $K(\Phi)$ -faithful total preorder \preceq_Φ , and let $*$ be any revision such that $K(\Phi * \alpha)$ can be defined in terms of \preceq_Φ using (Def * from \preceq), for every $\Phi \in \mathcal{E}$. Then $*$ satisfies (E*1) to (E*8). Conversely, suppose that $*$ is a revision which satisfies (E*1) to (E*8). Then every epistemic state Φ can be associated with a $K(\Phi)$ -faithful total preorder \preceq_Φ so that, for every $\Phi \in \mathcal{E}$, $K(\Phi * \alpha)$ can be defined in terms of \preceq_Φ using (Def * from \preceq).*

Observe that if the antecedent in (E*5) had been the requirement that $K(\Phi) = K(\Psi)$, we would have been obliged to consider only those sets of faithful preorders for which $\preceq_\Phi = \preceq_\Psi$ whenever $K(\Phi) = K(\Psi)$. As it stands, though, we are free to associate with an epistemic state Φ , any $K(\Phi)$ -faithful total preorder.

Since we are dealing with the finitely generated propositional case, it is easily verified that for a given revision $*$ satisfying (E*1)–(E*8), every epistemic state Φ is associated with a unique $K(\Phi)$ -faithful total preorder.

Proposition 7.3.2 *Let $*$ be a revision that satisfies (E*1) to (E*8) and pick any $\Phi \in \mathcal{E}$. There is a unique $K(\Phi)$ -faithful total preorder \preceq_Φ in terms of which $K(\Phi * \alpha)$ can be defined using (Def * from \preceq).*

Proof By theorem 7.3.1, \preceq_Φ exists. Assume there is a different $K(\Phi)$ -faithful total preorder \leq_Φ in terms of which $K(\Phi * \alpha)$ can be defined using (Def * from \preceq). That is, for some $u, v \in V$, either $u \preceq_\Phi v$ and $u \not\leq_\Phi v$, or $u \leq_\Phi v$ and $u \not\preceq_\Phi v$. Now pick an α such that $M(\alpha) = \{u, v\}$. (Since L is finitely generated, there is such an α .) Then $Min_{\preceq_\Phi}(\alpha) \neq Min_{\leq_\Phi}(\alpha)$, contradicting the assumption that $K(\Phi * \alpha)$ can be defined in terms of both \preceq_Φ and \leq_Φ using (Def * from \preceq). \square

Armed with proposition 7.3.2, we shall deviate slightly from the presentation of Darwiche and Pearl by taking an epistemic state Φ to be an ordered pair of the form $(K(\Phi), \preceq_\Phi)$. This is a potentially dangerous move, since it is at odds with the possibility that different epistemic states may be associated with the same belief set and faithful total preorder; a possibility that Darwiche and Pearl make provision for. Nevertheless, it will aid in the readability of the results discussed below.

Having made the move to revision operations on epistemic states, Darwiche and Pearl argue that (E*1)–(E*8) are too weak to provide a satisfactory account of iterated revision. Their argument is based on the application of the principle of Minimal Change which provides the underlying rationale for AGM revision. Where iterated revision

on epistemic states is concerned, it seems necessary to apply this principle to more than just the object-level beliefs of an agent. Semantically speaking, the faithful total preorder \preceq_Φ determines the models of $K(\Phi * \alpha)$ uniquely, but does not place any restrictions on the relative ordering of the countermodels of $K(\Phi * \alpha)$. Darwiche and Pearl propose that the principle of Minimal Change should be brought into play to minimise any change in the relative ordering of interpretations in the epistemic state resulting from a revision. From an information-theoretic point of view, this can be interpreted as an attempt to retain the relative credibility (or entrenchment) of infatoms. Their proposal involves the addition of the following postulates (the DP-postulates) to (E*1)–(E*8):

(DP*1) If $\alpha \models \beta$ then $K((\Phi * \beta) * \alpha) = K(\Phi * \alpha)$

(DP*2) If $\alpha \models \neg\beta$ then $K((\Phi * \beta) * \alpha) = K(\Phi * \alpha)$

(DP*3) If $\beta \in K(\Phi * \alpha)$ then $\beta \in K((\Phi * \beta) * \alpha)$

(DP*4) If $\neg\beta \notin K(\Phi * \alpha)$ then $\neg\beta \notin K((\Phi * \beta) * \alpha)$

Definition 7.3.3 A revision on epistemic states is a *DP-revision* iff it satisfies (E*1)–(E*8) and (DP*1)–(DP*4). \square

(DP*1) states that if an agent obtains more specific information after learning that β is the case, then β should be ignored. (DP*2) requires that any information contradicting newly obtained information should be ignored. On a contrapositive reading, (DP*3) insists that if an agent obtains the information β , but loses it immediately when acquiring the new information α , then β would never have formed part of the beliefs of the agent if it had acquired α immediately. And if an agent hasn't completely ruled out β after obtaining α , then (DP4) requires that first obtaining β and then α would also mean that β is not completely ruled out. In other words, as Darwiche and Pearl put it, information cannot contribute towards its own demise.

That the DP-postulates do indeed minimise changes in the relative ordering of interpretations can be seen from the following representation theorem, courtesy of Darwiche and Pearl. They prove that each one of the postulates (DP*1) to (DP*4) can be represented semantically as follows:

(DPR*1) If $u \models \alpha$ and $v \models \alpha$ then $u \preceq_\Phi v$ iff $u \preceq_{\Phi * \alpha} v$

(DPR*2) If $u \models \neg\alpha$ and $v \models \neg\alpha$ then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi*\alpha} v$

(DPR*3) If $u \models \alpha$ and $v \models \neg\alpha$ then $u \prec_{\Phi} v$ only if $u \prec_{\Phi*\alpha} v$

(DPR*4) If $u \models \alpha$ and $v \models \neg\alpha$ then $u \preceq_{\Phi} v$ only if $u \preceq_{\Phi*\alpha} v$

Theorem 7.3.4 [Darwiche and Pearl, 1997] *Let $*$ be a revision that satisfies (E*1) to (E*8). Then $*$*

$$\text{satisfies } \left\{ \begin{array}{l} (\text{DP} * 1) \\ (\text{DP} * 2) \\ (\text{DP} * 3) \\ (\text{DP} * 4) \end{array} \right\} \text{ iff it satisfies } \left\{ \begin{array}{l} (\text{DPR} * 1) \\ (\text{DPR} * 2) \\ (\text{DPR} * 3) \\ (\text{DPR} * 4) \end{array} \right\}.$$

(DPR*1) ensures that the relative ordering of the models of α is preserved after an α -revision; an application of the principle of Minimal Change to the models of α . Similarly, (DPR*2) requires that the relative ordering of the countermodels of α is preserved after an α -revision, which is a case of applying the principle of Minimal Change to the countermodels of α . (DPR*3) and (DPR*4) together ensure that any change in the relative ordering of a model u of α and a countermodel v of α will involve u moving lower down than v . As such, they also involve, to some extent, an application of the principle of Minimal Change.

A DP-revision by a wff α thus involves a “downward shift” of the models of α , while maintaining the relative orderings of the models of α and the countermodels of α respectively. DP-revision can therefore be seen as a qualitative version of Spohn’s conditionalisation. Indeed, Darwiche and Pearl mention that the inspiration for these postulates came from Spohn’s conditionalisation.

7.3.1 Minimal change

While (DPR*1)–(DPR*4) together impose considerable restrictions on the permissible ways of performing iterated revision, it is not in absolute accordance with the principle of Minimal Change. This is evident from the observation that there is a remaining case which is not covered by (DPR*1)–(DPR*4); disallowing the upward shift of a model of α relative to a countermodel of α . A blanket restriction of this kind would, of course, be incompatible with (E*1)–(E*8), since the minimal models of α will then not always

be permitted to occupy the lowest level in the ordering resulting from an α -revision. The closest we can come to an absolute adherence to the principle of Minimal Change is to preserve the relative ordering of all interpretations, except for those in $Min_{\preceq_{\Phi}}(\alpha)$. This idea is expressed by the following property:

(CBR*) If $u \notin Min_{\preceq_{\Phi}}(\alpha)$ and $v \notin Min_{\preceq_{\Phi}}(\alpha)$ then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \alpha} v$

Darwiche and Pearl show that this property is a semantic expression of the following postulate:

(CB*) If $\neg\alpha \in K(\Phi * \beta)$ then $K((\Phi * \beta) * \alpha) = K(\Phi * \alpha)$

Theorem 7.3.5 [Darwiche and Pearl, 1997] *Let $*$ be a revision that satisfies (E*1) to (E*8). Then $*$ satisfies (CB*) iff it satisfies (CBR*).*

It is easily seen that (CBR*) implies (DPR*1)–(DPR*4) but that the converse doesn't hold. In fact, when added to (E*1)–(E*8), (CB*) describes a unique revision, having the following semantic definition:

$$\textbf{(Def } *) \left[\begin{array}{l} K(\Phi * \alpha) = Th(Min_{\preceq_{\Phi}}(\alpha)) \text{ and} \\ u \preceq_{\Phi * \alpha} v \text{ iff } \left\{ \begin{array}{l} v \in V \text{ if } u \in Min_{\preceq_{\Phi}}(\alpha), \\ u \preceq_{\Phi} v \text{ and } v \notin Min_{\preceq_{\Phi}}(\alpha) \text{ otherwise} \end{array} \right. \end{array} \right]$$

The revision defined in terms of (Def *) was first proposed by Boutilier under the names “natural revision” [Boutilier, 1993] and “minimal conditional revision” [Boutilier, 1996]. From theorem 7.3.1 it follows that minimal conditional revision satisfies (E*1)–(E*8), and from theorem 7.3.4 that it satisfies (DP*1)–(DP*4). It can also be seen as a qualitative version of adjustment, one of the transmutation methods of Williams which was discussed in section 7.1.

7.3.2 Conditional beliefs

Darwiche and Pearl also justify the DP-postulates in terms of conditional beliefs. An agent is said to hold a *conditional belief* $\alpha \triangleright \beta$ iff the belief β is in the set of beliefs that the agent holds after an α -revision. Note that while α and β are taken as wffs of the language L , the conditional belief $\alpha \triangleright \beta$ is not, and \triangleright should thus be seen as a *meta-connective*. As Boutilier [1993, 1996] has shown, epistemic states can also

be represented as (appropriately chosen) sets of conditional beliefs. It is simply a matter of associating with an epistemic state Φ , the following set of conditional beliefs: $\{\alpha \triangleright \beta \mid \beta \in K(\Phi * \alpha)\}$. We say that the conditional belief $\alpha \triangleright \beta$ is in the epistemic state Φ , written as $\alpha \triangleright \beta \in \Phi$, iff $\beta \in K(\Phi * \alpha)$. The DP-postulates can be rephrased in terms of conditional beliefs as follows:

(CDP*1) If $\alpha \models \beta$ then $\alpha \triangleright \gamma \in (\Phi * \beta)$ iff $\alpha \triangleright \gamma \in \Phi$

(CDP*2) If $\alpha \models \neg\beta$ then $\alpha \triangleright \gamma \in (\Phi * \beta)$ iff $\alpha \triangleright \gamma \in \Phi$

(CDP*3) If $\alpha \triangleright \beta \in \Phi$ then $\alpha \triangleright \beta \in (\Phi * \beta)$

(CDP*4) If $\alpha \triangleright \neg\beta \notin \Phi$ then $\alpha \triangleright \neg\beta \notin (\Phi * \beta)$

In this reading, the DP-postulates can be justified by an application of the principle of Minimal Change to conditional beliefs. (CDP*1) and (CDP*2) ensure that certain sets of conditional beliefs will remain unchanged, (CDP*3) requires that certain conditional beliefs be retained, and (CDP*4) forbids the addition of certain conditional beliefs. More precisely, (CDP*1) requires that the conditional beliefs in Φ with antecedents that are logically stronger than a wff β , should be exactly those in the epistemic state obtained from Φ by a β -revision. Similarly, (CDP*2) requires that the conditional beliefs in Φ with antecedents that contradict a wff β , should be exactly those in the epistemic state obtained from Φ by a β -revision. And (CDP*3) requires that a conditional belief should not be given up after a revision by its consequent, while (CDP*4) insists that a conditional not in the current epistemic state should not be added after a revision by the negation of its consequent.

As the name suggests, Boutilier's minimal conditional revision can also be justified by reference to conditional beliefs. Observe firstly that (CB*) can be given the following reading in terms of conditional beliefs:

(CCB*) If $\neg\alpha \in \Phi * \beta$ then $\alpha \triangleright \gamma \in \Phi$ iff $\alpha \triangleright \gamma \in \Phi * \beta$

In other words, (CCB*) states that if α is incompatible with $\Phi * \beta$ then Φ and $\Phi * \beta$ should contain exactly the same conditional beliefs with α as antecedent. Boutilier [1996, pp. 277-278] has shown that minimal conditional revision is the revision satisfying (E*1)–(E*8), which causes the minimum disturbance with regard to conditional beliefs. With such a strict adherence to the principle of Minimal Change, it is thus well worth considering whether minimal conditional revision should be regarded as *the* way to perform iterated revision. We discuss this issue in section 7.3.3.

7.3.3 Is iterated DP-revision rational?

Darwiche and Pearl provide a number of convincing examples in justification of their account of iterated revision. Some of these serve as counterexamples, indicating that (E*1)–(E*8) do not rule out all counterintuitive forms of iterated revision, thus paving the way for the introduction of additional postulates. Others are used as evidence in corroboration of the more abstract claims intended as justification for adding the four DP-postulates. While the latter examples form part of a powerful case in favour of regarding all DP-revisions as rational, they cannot be used as part of an argument that the *only* rational iterated revisions are DP-revisions. And indeed, there are indications that (DP*2), in particular, will eliminate some perfectly plausible forms of iterated revision.⁷ Cantwell [1999] shows that the following variant of the controversial Recovery postulate for AGM contraction (see chapter 6) is a derived property of any revision satisfying (E*1)–(E*8) and (DP*2):

(Revision Recovery) If $K(\Phi) \neq Cn(\perp)$ and $\alpha \in K(\Phi)$ then $K((\Phi * \neg\alpha) * \alpha) = K(\Phi)$

As a result, the counterexamples levelled against Recovery can also be used to argue against the inclusion of (DP*2). Here, for instance, is a modified version of example 6.1.2 to show that Revision Recovery is counterintuitive.

Example 7.3.6 I read a book about Cleopatra, in which the claim is made that she had a son and a daughter. I subsequently discover that the book is fictional, which leads me to adopt the belief that Cleopatra did *not* have a child. However, on consulting a history book I discover that Cleopatra indeed had a child, and I thus revise my belief set with this assertion.

Let L be a propositional language generated by the two atoms p and q . Let p denote the assertion that Cleopatra had a son, and q the assertion that she had a daughter. Then $K(\Phi) = Cn(p, q)$. The adoption of the belief that she did not have a child is formalised as $\Phi * \neg(p \vee q)$. Since $p \vee q \in K(\Phi)$, Revision Recovery requires that $K((\Phi * \neg(p \vee q)) * (p \vee q)) = K(\Phi)$. So revising with the assertion that Cleopatra did, after all, have a child, will ensure that I again entertain the belief that she had a son and the belief that she had a daughter; a conclusion which seems unreasonable in this context. \square

⁷In section 8.4.1, we show that recent developments concerning base change also call the appropriateness of (DP*1) into question.

In fact, as Cantwell [1999] observes, Revision Recovery seems to be even more problematic than Recovery, since revision usually involves greater changes in epistemic states than contraction.

We now turn to a different kind of question regarding the rationality of DP-revision. We have seen that Boutilier's minimal conditional revision causes the minimum disturbance in the current epistemic state, resulting in the strongest possible adherence to the principle of Minimal Change. This raises the question of whether minimal conditional revision should, perhaps, be regarded as the *only* rational form of revision on epistemic states. Darwiche and Pearl [1997] provide a convincing argument against such a view, indicating that the importance of the principle of Minimal Change should not be overestimated. It is based on the following example.

Example 7.3.7 We encounter a strange new animal and it appears to be a bird, so we believe the animal is a bird. As it comes closer to our hiding place, we see clearly that the animal is red, so we believe that it is a red bird. To remove any further doubts, we call in a bird expert who takes it for examination and concludes that it is not really a bird, but some sort of mammal. The question is now whether we should still believe that the animal is red. Intuitively, it seems that we should, but minimal conditional revision rules that we may *not* believe that the animal is red. This can be verified by using the propositional language generated by the two atoms b and r to represent our knowledge. Let b represent the assertion that the animal is a bird, let r represent the assertion that it is red, let (V, \models) be the valuation semantics for L with $V = \{00, 01, 10, 11\}$, and let $*$ be the minimal conditional function defined using (Def $*$).⁸ Let $\Phi = (K(\Phi), \preceq_\Phi)$ be the epistemic state representing the situation *before* we see the bird. Then $K(\Phi) = Cn(\top)$ and $\preceq_\Phi = V \times V$. Furthermore, it can be verified that $K(((\Phi * b) * r) * \neg b) = Cn(\neg b)$. \square

The problem can be approached from various angles, but Darwiche and Pearl provide a particularly convincing analysis in terms of conditional beliefs. It is easily verified that the conditional belief $\neg b \triangleright r$ is not, and should not be, in the epistemic state $\Phi * b$. Bearing in mind that minimal conditional revision effects the minimal permissible change on the conditional beliefs in an epistemic state, $\neg b \triangleright r$ will not be in the epistemic state $(\Phi * b) * r$ either. But this is counterintuitive. Since the colour of the

⁸We assume that the first digit in the pairs of zeroes and ones denotes the truth value of b , and the second one the truth value of r .

animal is independent of it being a bird or not, we should persist in believing that it is red. In general then, the requirement of minimal conditional revision that the set of conditional beliefs should remain as stable as possible, may impact on the persistence of some object-level beliefs, yielding counterintuitive results.

Although minimal conditional revision is clearly too restrictive to be the only form of rational revision on epistemic states, it is perhaps worth considering a special case under which it *does* seem reasonable to ensure that minimal changes occur in an epistemic state Φ ; when the wff with which to revise is compatible with Φ . This can be expressed as a weakened form of the postulate (CB).

(WCB) If $\neg\beta \notin K(\Phi)$ and $\neg\alpha \in K(\Phi * \beta)$, then $K((\Phi * \beta) * \alpha) = K(\Phi * \alpha)$

And predictably, (WCB) can be expressed semantically as follows:

(WCBR) If $\neg\beta \notin K(\Phi)$, $u \notin \text{Min}_{\preceq_{\Phi}}(\beta)$ and $v \notin \text{Min}_{\preceq_{\Phi}}(\beta)$, then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \beta} v$

Proposition 7.3.8 *Let $*$ be a revision satisfying (E*1)–(E*8). Then $*$ satisfies (WCB) iff it satisfies (WCBR).*

Proof Observe firstly that by theorem 7.3.1, $K(\Phi * \alpha) = \text{Th}(\text{Min}_{\preceq_{\Phi}}(\alpha))$. Now suppose that (WCB) holds, that $\neg\beta \notin K(\Phi)$, and pick any $u, v \in V$ such that $u \notin \text{Min}_{\preceq_{\Phi}}(\beta)$ and $v \notin \text{Min}_{\preceq_{\Phi}}(\beta)$. Let α be such that $M(\alpha) = \{u, v\}$. (Since L is finitely generated, there is such an α .) Then $\neg\alpha \in K(\Phi * \beta)$, and $K((\Phi * \beta) * \alpha) = K(\Phi * \alpha)$ by (WCB). But then $\text{Min}_{\preceq_{\Phi * \beta}}(\alpha) = \text{Min}_{\preceq_{\Phi}}(\alpha)$, from which it follows that $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \beta} v$. Conversely, suppose that (WCBR) holds, that $\neg\beta \notin K(\Phi)$, and that $\neg\alpha \in K(\Phi * \beta)$. By (WCBR) it follows that $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \beta} v$ for every $u, v \in V$ such that $u \notin \text{Min}_{\preceq_{\Phi}}(\beta)$ and $v \notin \text{Min}_{\preceq_{\Phi}}(\beta)$. And since $M(\alpha) \cap M(K(\Phi * \beta)) = \emptyset$, we have that $\text{Min}_{\preceq_{\Phi}}(\alpha) = \text{Min}_{\preceq_{\Phi * \beta}}(\alpha)$. Therefore $K(\Phi * \alpha) = K((\Phi * \beta) * \alpha)$. \square

While the weakened form of minimal conditional revision, obtained by replacing (CB) with (WCB), might seem appealing at first, it does not escape the problems associated with full minimal conditional revision, as one might have hoped. Example 7.3.7 in particular, is also applicable to any revision satisfying (E*1)–(E*8) and (WCB).

7.3.4 Iterated DP-withdrawal

Recall from page 7 that Levi's commensurability thesis sees revision as a two-step process involving a removal followed by an expansion. Taking this view seriously requires

of us to provide an account of iterated removal and iterated expansion on epistemic states as well. Unlike the case for revision, it seems reasonable to require of an expansion \oplus on epistemic states to be in strict adherence to the principle of Minimal Change. The following semantic property is a rephrasing of (CBR) for expansion.

(**ER** \oplus) If $u \notin M(K(\Phi \oplus \alpha))$ and $v \notin M(K(\Phi \oplus \alpha))$, then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi \oplus \alpha} v$

Combined with the obvious requirement that $K(\Phi \oplus \alpha) = K(\Phi) + \alpha$, we thus have the following unique method for expanding epistemic states.

$$(\mathbf{Def} \oplus) \left[\begin{array}{l} K(\Phi \oplus \alpha) = K(\Phi) + \alpha \text{ and} \\ u \preceq_{\Phi \oplus \alpha} v \text{ iff } \begin{cases} v \in V \text{ if } u \in M(K(\Phi) + \alpha), \\ u \preceq_{\Phi} v \text{ and } v \notin M(K(\Phi) + \alpha) \text{ otherwise} \end{cases} \end{array} \right]$$

Obtaining a suitable account of removal on epistemic states is less straightforward. It will, of course, depend on the particular type of removal which we regard as appropriate, although results in chapter 6 indicate that it would have to be some form of reasonable withdrawal (see definition 6.3.13). For now, we restrict ourselves to a generalisation of AGM contraction and severe withdrawal to epistemic states.⁹ An AGM contraction \approx on epistemic states is required to satisfy the following postulates:

(**E** \approx 1) $K(\Phi \approx \alpha) = Cn(K(\Phi \approx \alpha))$

(**E** \approx 2) $K(\Phi \approx \alpha) \subseteq K(\Phi)$

(**E** \approx 3) If $\alpha \notin K(\Phi)$ then $K(\Phi \approx \alpha) = K$

(**E** \approx 4) If $\not\models \alpha$ then $\alpha \notin K(\Phi \approx \alpha)$

(**E** \approx 5) If $\Phi = \Psi$ and $\alpha \equiv \beta$ then $K(\Phi \approx \alpha) = K(\Psi \approx \beta)$

(**E** \approx 6) If $\alpha \in K(\Phi)$ then $K(\Phi \approx \alpha) + \alpha = K(\Phi)$

(**E** \approx 7) $K(\Phi \approx \alpha) \cap K(\Phi \approx \beta) \subseteq K(\Phi \approx (\alpha \wedge \beta))$

⁹Not too much should be read into this restriction. It is based on a purely practical consideration; the current representation of an epistemic state Φ as a belief set $K(\Phi)$ and a $K(\Phi)$ -faithful *total* preorder. A representation using $K(\Phi)$ -faithful modular weak partial orders, for example, would have resulted in a restriction to AGM contraction and systematic withdrawal.

(E \approx 8) If $\beta \notin K(\Phi \approx (\alpha \wedge \beta))$ then $K(\Phi \approx (\alpha \wedge \beta)) \subseteq K(\Phi \approx \beta)$

Definition 7.3.9 A removal on epistemic states is an *AGM contraction* iff it satisfies (E \approx 1)–(E \approx 8). \square

For the generalisation of severe withdrawal, we also need the following postulates.

(E \approx 6) If $\models \alpha$ then $K(\Phi \approx \alpha) = K(\Phi)$

(E \approx 8) If $\not\models \alpha$ then $K \approx \phi \subseteq K \approx (\alpha \wedge \beta)$

Definition 7.3.10 A removal on epistemic states is a *severe withdrawal* iff it satisfies (E \approx 1)–(E \approx 5), (E \approx 6), (E \approx 7) and (E \approx 8). \square

The following results are then easily obtained.

Theorem 7.3.11 1. Let \approx be any removal such that $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def \sim from \preceq), for every $\Phi \in \mathcal{E}$. Then \approx is an AGM contraction. Conversely, suppose that \approx is an AGM contraction. For every $\Phi \in \mathcal{E}$, $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def \approx from \preceq).

2. Let \approx be any removal such that $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def ∇_{\preceq} from \preceq), for every $\Phi \in \mathcal{E}$. Then \approx is a severe withdrawal. Conversely, suppose that \approx is a severe withdrawal. For every $\Phi \in \mathcal{E}$, $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def ∇_{\preceq} from \preceq).

Proof 1. Follows from theorem 3.2.6.

2. Follows from definition 6.3.1 and theorem 6.3.2. \square

It is also easy to verify that, on the level of belief sets, the roles of the Levi identity (Def $*$ from \sim) and the Harper identity (Def $-$ from $*$) remain unchanged.

Corollary 7.3.12 Let \approx and \approx be removals, and $*$ a revision such that, for every $\Phi \in \mathcal{E}$,

- $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def \sim from \preceq)
- $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def \sim from ∇_{\preceq})

- $K(\Phi * \alpha)$ can be defined in terms of \preceq_Φ using (Def * from \preceq)

Then, for every $\alpha \in L$,

1. $K(\Phi * \alpha) = (K(\Phi) \approx \neg\alpha) + \alpha = (K(\Phi) \ddot{\approx} \neg\alpha) + \alpha$, and
2. $K(\Phi \approx \alpha) = K(\Phi * \neg\alpha) \cap K(\Phi)$.

Proof Follows from proposition 3.2.8 and theorem 6.3.10. \square

An adherence to Levi's commensurability thesis then seems to suggest the lifting of the Levi identity to the level of epistemic states in the following manner.

$$(\mathbf{Def} * \mathbf{from} \approx) \quad \Phi * \alpha = (\Phi \approx \neg\alpha) \oplus \alpha$$

The next two results show that, where AGM contraction and revision on epistemic states are concerned, this seems to be right choice.

Proposition 7.3.13 *Let \approx and $\ddot{\approx}$ be removals such that, for every $\Phi \in \mathcal{E}$,*

- $K(\Phi \approx \alpha)$ can be defined in terms of \preceq_Φ using (Def \sim from \preceq)
- $K(\Phi \ddot{\approx} \alpha)$ can be defined in terms of \preceq_Φ using (Def \sim from ∇_{\preceq}).

Let $$ and $\ddot{*}$ be the revisions defined in terms of (Def $*$ from \approx) using \approx and $\ddot{\approx}$ respectively.*

1. For every $\alpha \in L$, $K(\Phi * \alpha) = K(\Phi \ddot{*} \alpha) = (K(\Phi) \approx \neg\alpha) + \alpha = (K(\Phi) \ddot{\approx} \neg\alpha) + \alpha$.
2. Both $*$ and $\ddot{*}$ satisfy (E*1)–(E*8).

Proof 1. Follows easily from theorem 6.3.10 and the definition of \oplus .

2. Follows easily from part (1) and proposition 3.2.8

\square

Given the connection between withdrawal and revision on the level of belief sets, the following postulates for withdrawal on epistemic states are obvious analogues of the semantic DP-postulates for revision.

(DPR \approx 1) If $u \Vdash \neg\alpha$ and $v \Vdash \neg\alpha$ then $u \preceq_\Phi v$ iff $u \preceq_{\Phi \approx \alpha} v$

(DPR \approx 2) If $u \models \alpha$ and $v \models \alpha$ then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi \approx \alpha} v$

(DPR \approx 3) If $u \models \neg \alpha$ and $v \models \alpha$ then $u \prec_{\Phi} v$ only if $u \prec_{\Phi \approx \alpha} v$

(DPR \approx 4) If $u \models \neg \alpha$ and $v \models \alpha$ then $u \preceq_{\Phi} v$ only if $u \preceq_{\Phi \approx \alpha} v$

For AGM contraction, (Def \ast from \approx) provides the expected link between these postulates and the DP-postulates for revision.

Proposition 7.3.14 *Let \approx be any AGM contraction and let \ast be the revision defined in terms of \approx using (Def \ast from \approx).*

1. *If \approx satisfies (DPR \approx 1) then \ast satisfies (DPR \ast 1).*
2. *If \approx satisfies (DPR \approx 2) then \ast satisfies (DPR \ast 2).*
3. *If \approx satisfies (DPR \approx 3) then \ast satisfies (DPR \ast 3).*
4. *If \approx satisfies (DPR \approx 4) then \ast satisfies (DPR \ast 4).*

Proof Follows from theorem 7.3.11, proposition 7.3.13, and proposition 3.2.8. \square

Interestingly enough, though, we do not get a similar result when defining revision in terms of severe withdrawal. As the next example and the proposition following it show, the revision obtained from a severe withdrawal on epistemic states satisfies (DP \ast 1), (DP \ast 3) and (DP \ast 4), but not (DP \ast 2).

Example 7.3.15 Let L be the propositional language generated by the atoms p and q , with the valuation semantics (V, \models) , where $V = \{00, 01, 10, 11\}$. Let \approx be any severe withdrawal satisfying (DPR \approx 1)–(DPR \approx 4) such that the following holds for the epistemic states Φ and Ψ , where $\Psi = \Phi \approx \neg(p \wedge q)$:

$$K(\Phi) = Cn(\neg p \wedge \neg q) \text{ and } u \preceq_{\Phi} v \text{ iff } \begin{cases} v \in V \text{ if } u = 00, \\ v \in \{01, 10, 11\} \text{ if } u \in \{01, 10, 11\}, \end{cases}$$

$$K(\Psi) = Cn(\top) \text{ and } \preceq_{\Psi} = V \times V, \text{ and}$$

$$K(\Psi \oplus p \wedge q) = Cn(p \wedge q) \text{ and } u \preceq_{\Psi \oplus (p \wedge q)} v \text{ iff } \begin{cases} v \in V \text{ if } u = 11, \\ v \in \{00, 01, 10\} \text{ otherwise} \end{cases}$$

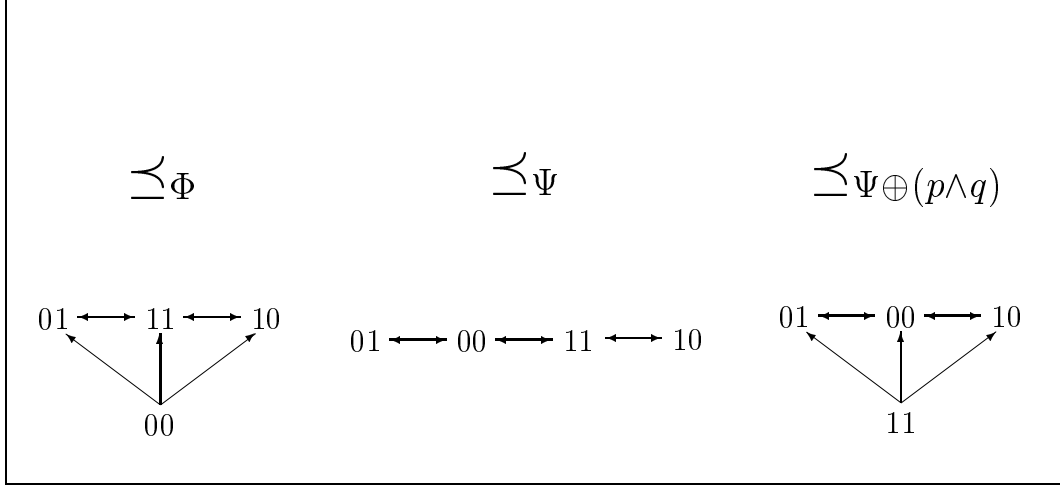


Figure 7.1: Graphical representations of the faithful total preorders \preceq_Φ , \preceq_Ψ , and $\preceq_{\Psi \oplus (p \wedge q)}$ used in example 7.3.15. Two interpretations u and v are in a faithful total preorder iff (u, v) is in the reflexive transitive closure of the relation determined by the arrows.

Figure 7.1 contains graphical representations of \preceq_Φ , \preceq_Ψ , and $\preceq_{\Psi \oplus (p \wedge q)}$. It is easily verified that such a severe withdrawal \approx exists, but that the revision \ast defined in terms of \approx using (Def \ast from \approx) does not satisfy (DPR \ast 2). In particular, it follows that $00, 10 \in M(\neg(p \wedge q))$, but it is not the case that $10 \preceq_\Phi 00$ iff $10 \preceq_{\Phi \ast (p \wedge q)} 00$. \square

Proposition 7.3.16 *Let \approx be any severe withdrawal and let \ast be the revision defined in terms of \approx using (Def \ast from \approx).*

1. *If \approx satisfies (DP \approx 1) then \ast satisfies (DP \ast 1).*
2. *If \approx satisfies (DP \approx 3) then \ast satisfies (DP \ast 3).*
3. *If \approx satisfies (DP \approx 4) then \ast satisfies (DP \ast 4).*

Proof These results follow easily by observing that, for every $\alpha \in L$, the total preorders $\preceq_{\Phi \approx \neg \alpha}$ and $\preceq_{\Phi \ast \alpha}$ are identical when restricted to elements of $V \setminus M(K(\Phi \approx \neg \alpha))$. \square

The results on iterated withdrawal suggest another reason for dropping (DPR*2) as a reasonable property of iterated revision; its incompatibility with Levi's commensurability thesis when applied to severe withdrawal on epistemic states. Semantically speaking, this incompatibility can be traced back to the fact that, unlike AGM contraction, a severe $\neg\alpha$ -withdrawal of an epistemic state Φ may result in the inclusion of models of $\neg\alpha$ to the models of $K(\Phi)$. This results in the loss of information pertaining to the relative ordering of such elements of $M(\neg\alpha)$ in \preceq_Φ , rendering a subsequent expansion unable to preserve the ordering of all models of $\neg\alpha$ in $K(\Phi * \alpha)$. It can therefore be shown that the application of (Def * from \approx) to any form of reasonable withdrawal on epistemic states, except for AGM contraction, will result in a revision which does not satisfy (DP*2).

7.4 Iterated L-revision

Lehmann [1995] considers iterated belief revision in the context of finite sequences of revisions. He extends the notion of a revision $*$ on epistemic states to a revision by a finite sequence of wffs. We use the Greek letter σ to denote such a finite sequence. $\Phi * \sigma$ then refers to the iterated revision of Φ by the wffs in σ , and if σ is the empty sequence, $\Phi * \sigma$ is just the epistemic state Φ . Concatenation of sequences is denoted by \bullet , and a wff α is identified with a sequence of length one. So the sequence $\alpha \bullet \sigma$ consists of the wff α followed by the wffs in σ , and $\sigma \bullet \alpha$ consists of the wffs in σ followed by the wff α .

Considering only sequences of *satisfiable* wffs, Lehmann proposes the following postulates for iterated revision.

$$(L*1) \quad K(\Phi) = Cn(K(\Phi))$$

$$(L*2) \quad \alpha \in K(\Phi * \alpha)$$

$$(L*3) \quad K(\Phi * \alpha) \subseteq K(\Phi) + \alpha$$

$$(L*4) \quad \text{If } \alpha \in K(\Phi) \text{ then } K(\Phi * \sigma) = K(\Phi * (\alpha \bullet \sigma))$$

$$(L*5) \quad \text{If } \alpha \models \beta \text{ then } K(\Phi * (\beta \bullet \alpha \bullet \sigma)) = K(\Phi * (\alpha \bullet \sigma))$$

$$(L*6) \quad K(\Phi) \neq Cn(\perp)$$

$$(L*7) \quad K(\Phi * (\neg\alpha \bullet \alpha)) \subseteq K(\Phi) + \alpha$$

$$(L*8) \quad \text{If } \neg\beta \notin K(\Phi * \alpha) \text{ then } K(\Phi * (\alpha \bullet \beta \bullet \sigma)) = K(\Phi * (\alpha \bullet \alpha \wedge \beta \bullet \sigma))$$

Definition 7.4.1 A revision on epistemic states is an *L-revision* iff it satisfies (L*1) to (L*8). \square

It follows easily from Lehmann's results that every L-revision also satisfies (E*1) to (E*8), (DP*1), (DP*3) and (DP*4). In fact, (L*1), (L*2), (L*3) and (L*6) correspond exactly to (E*1), (E*2), (E*3) and (E*6) respectively.¹⁰ (L*4) states that superfluous revisions are useless and should have no effect on subsequent revisions. While this may be a reasonable constraint under certain circumstances, it is a very strong restriction to impose on all rational iterated revisions. The main reason for this is that (L*4) is at odds with the notion of *corroborating evidence*; the idea that one's belief in an assertion is strengthened by repeated observations confirming that it holds.

Example 7.4.2 Suppose that an agent obtains evidence that α is the case, followed by evidence that β is the case. If subsequent evidence obtained makes it clear that exactly one of α or β holds, it seems difficult to decide between α and β . If one is inclined to trust more recent evidence, it is perhaps reasonable to entertain the option that it is β that holds. On the other hand, suppose that the agent obtains evidence that α is the case, followed by evidence that β is the case, which, in turn is followed by confirmation that α is the case. If subsequent evidence *now* makes it clear that exactly one of α or β holds, it seems reasonable to believe that α is the case, mainly because our initial belief in α was corroborated by confirming evidence that α holds. But such a conclusion is prohibited by (L*4). \square

(L*5) is a strengthening of (DP*1). It requires of an agent, when obtaining more specific information following β , not just to discard the influence of β in obtaining the resulting belief set, but also in all subsequent revisions. (L*7) is a weakened version of (DP*2). Given the rest of Lehmann's postulates, it is equivalent to the following postulate:

$$(L*9) \quad K((\Phi * \neg\alpha) * \alpha) \subseteq K(\Phi * \alpha)$$

Proposition 7.4.3 *Let $*$ be a revision satisfying (L*1)–(L*6) and (L*8). Then $*$ satisfies (L*7) iff it satisfies (L*9).*

¹⁰Since we only revise with sequences of satisfiable wffs, (L*6) is indeed equivalent to (E*6).

Proof Suppose that $*$ satisfies (L*7). Then Lehmann's Lemma 3 [1995,p. 1537] shows that $*$ satisfies (E*4). If $\neg\alpha \in K(\Phi)$ then $K((\Phi * \neg\alpha) * \alpha) = K(\Phi * \alpha)$ by (L*4). And if $\neg\alpha \notin K(\Phi)$ then

$$\begin{aligned} & K((\Phi * \neg\alpha) * \alpha) \\ \subseteq & K(\Phi) + \alpha \text{ by (L*7)} \\ \subseteq & K(\Phi * \alpha) \text{ by (E*4).} \end{aligned}$$

The converse follows immediately from (L*3). \square

In the presence of (L*5), (L*8) is a strengthening of (K*9), the postulate for AGM revision on belief sets which follows from (K*7) and (K*8).¹¹ This can be seen from Lehmann's result that (L*8) is equivalent to the following postulate whenever (L*5) holds.

(L*10) If $\neg\beta \notin K(\Phi * \alpha)$ then $K(\Phi * (\alpha \bullet \beta \bullet \sigma)) = K(\Phi * (\alpha \wedge \beta \bullet \sigma))$

As mentioned above, any L-revision also satisfies (DP*3). In fact, Lehmann shows that such a revision satisfies the following strengthened version of (DP*3).

(L*11) If $\beta \in K(\Phi * \alpha)$ then $K(\Phi * (\alpha \bullet \sigma)) = K(\Phi * (\beta \bullet \alpha \bullet \sigma))$

Since every L-revision also satisfies (E*1)–(E*8), it follows by theorem 7.3.1 and proposition 7.3.2 that we can associate with every epistemic state Φ , a unique $K(\Phi)$ -faithful total preorder \preceq_Φ such that $K(\Phi * \alpha)$ can be defined in terms of \preceq_Φ using (Def * from \preceq), for every $\alpha \in L$. Observe, though, that it is not possible to *represent* every epistemic state Φ as an ordered pair of the form $(K(\Phi), \preceq_\Phi)$. This becomes clear once we realise that every epistemic state is associated with a unique finite sequence of wffs, since there are only a finite number of such ordered pairs, but an infinite number of epistemic states.

We conclude this discussion with a brief note concerning a representation theorem proved by Lehmann. He provides a method, involving the *widening ranked models*, of constructing precisely the L-revisions. It involves the use of an implausibility ranking over sets of valuations. Lehmann warns that it is just a technical tool, though, and that it should not be seen as a description of the epistemic states of an agent.

¹¹Lehmann [1995,p. 1537] claims that (L*8) does not represent any strengthening of the postulates (E*1)–(E*8), but this is clearly incorrect. In section 7.5.1 we show that Papinis's P_{\triangleright} -revision satisfies (E*1)–(E*8), but does not satisfy (L*8), and in section 8.4.2 we give another example of a revision on epistemic states which satisfies (E*1)–(E*8), but does not satisfy (L*8).

7.5 Observation-based revision

Papini [1998, 1999] has recently proposed that iterated revision be viewed in the context of sequences of observations made by an agent. The basic idea is that the history of the agent's observations should be taken into account in some or other way. She considers the two particular constructions which lie on opposite sides of the spectrum when assessing the reliability of observations. The remainder of this section is devoted to a description of these revision operations.

7.5.1 P_{\triangleright} -revision

The intuition associated with AGM revision on belief sets contains the assumption that the wff α with which to revise a belief set K , takes precedence over the information currently contained in K . A generalisation of revision to epistemic states can accommodate this assumption in a wide variety of ways. From an information-theoretic point of view, though, it is fair to say that any such a generalisation needs to reflect the requirement that no content bit of $\neg\alpha$ may become more entrenched (or more credible) relative to the content bits of α ; a requirement that is captured by (DP*3) and (DP*4). The strongest expression of this requirement is the insistence that an α -revision should result in an epistemic state where the content bits of α are all more entrenched than the content bits of $\neg\alpha$. In model-theoretic terms, it means that an α -revision of the epistemic state Φ should result in an epistemic state in which the total preorder $\preceq_{\Phi * \alpha}$ places the models of α strictly below the countermodels of $\neg\alpha$. It can be formulated as follows:

(PR*) If $u \in M(\alpha)$ and $v \in M(\neg\alpha)$, then $u \prec_{\Phi * \alpha} v$

This is the idea underlying one of Papini's approaches to iterated revision. In such an approach, the more recent observations of the agent are to be taken more seriously. She provides the following semantic definition of revision on the level of epistemic states:¹²

$$(\text{Def } *) \left[\begin{array}{l} K(\Phi *_{\triangleright} \alpha) = Th(\text{Min}_{\preceq_{\Phi_{\triangleright}}}(\alpha)) \\ u \preceq_{\Phi *_{\triangleright} \alpha} v \text{ iff } \begin{cases} u \preceq_{\Phi} v \text{ if } u, v \in M(\alpha) \text{ or } u, v \in M(\neg\alpha), \\ u \in M(\alpha), \text{ otherwise} \end{cases} \end{array} \right]$$

¹²Papini's construction uses polynomials on the natural numbers, but it is easily seen that her definition corresponds to the one we give here.

Definition 7.5.1 The revision on epistemic states defined in terms of (Def \ast_{\triangleright}) is referred to as P_{\triangleright} -revision and denoted by \ast_{\triangleright} . \square

Papini shows that \ast_{\triangleright} is a DP-revision (i.e., it satisfies (E \ast 1)–(E \ast 8) and (DP \ast 1)–(DP \ast 4)). It turns out that \ast_{\triangleright} can be characterised precisely by adding the following postulate to (DP \ast 1) and (DP \ast 2):

(P \ast) If $\alpha \not\models \neg\beta$ then $K((\Phi \ast \alpha) \ast \beta) = K(\Phi \ast (\alpha \wedge \beta))$

Theorem 7.5.2 *The P_{\triangleright} -revision \ast_{\triangleright} is the only revision satisfying (E \ast 1)–(E \ast 8), (DP \ast 1), (DP \ast 2), and (P \ast).*

Proof From theorem 7.3.1 it follows easily that \ast_{\triangleright} satisfies (E \ast 1)–(E \ast 8), and from theorem 7.3.4 it clearly follows that \ast_{\triangleright} satisfies (DP \ast 1) and (DP \ast 2). Furthermore, it is easily verified that \ast_{\triangleright} is the only revision satisfying (E \ast 1)–(E \ast 8), (DP \ast 1), (DP \ast 2), and (PR \ast). It thus suffices to show that for every revision satisfying (E \ast 1)–(E \ast 8) and (DP \ast 1), the postulates (P \ast) and (PR \ast) are equivalent.

Pick any revision \ast satisfying (E \ast 1)–(E \ast 8) and (DP \ast 1), and suppose that \ast does not satisfy (PR \ast). Then there is an $\alpha \in L$, a $u \in M(\alpha)$ and a $v \notin M(\alpha)$ such that $v \preceq_{\Phi \ast \alpha} u$, for some epistemic state Φ . Now let β be such that $M(\beta) = \{u, v\}$. Then $\text{Min}_{\preceq_{\Phi}}(\alpha \wedge \beta) = M(\alpha \wedge \beta) = \{u\}$, and so $K(\Phi \ast \alpha \wedge \beta) = \text{Cn}(\alpha \wedge \beta)$. On the other hand, either $\text{Min}_{\preceq_{\Phi \ast \alpha}}(\beta) = \{v\}$ or $\text{Min}_{\preceq_{\Phi \ast \alpha}}(\beta) = \{u, v\}$. But then $K((\Phi \ast \alpha) \ast \beta) \neq K(\Phi \ast (\alpha \wedge \beta))$ by theorem 7.3.1, and so \ast does not satisfy (P \ast). Conversely, suppose that \ast satisfies (PR \ast). Now pick any $\alpha, \beta \in L$ such that $\alpha \not\models \neg\beta$. By (PR \ast) it follows that $\text{Min}_{\preceq_{\Phi \ast \alpha}}(\beta) \subseteq M(\alpha)$, and so $\text{Min}_{\preceq_{\Phi \ast \alpha}}(\beta) = \text{Min}_{\preceq_{\Phi \ast \alpha}}(\alpha \wedge \beta)$. Furthermore, theorem 7.3.4 guarantees that \ast satisfies (DPR \ast 1), and so $\text{Min}_{\preceq_{\Phi}}(\alpha \wedge \beta) = \text{Min}_{\preceq_{\Phi \ast \alpha}}(\alpha \wedge \beta) = \text{Min}_{\preceq_{\Phi \ast \alpha}}(\beta)$, from which it follows by theorem 7.3.1 that $K((\Phi \ast \alpha) \ast \beta) = K(\Phi \ast (\alpha \wedge \beta))$. \square

Observe that (P \ast) requires of iterated revision and simultaneous revision to yield identical results whenever α and β are compatible. In other words, an α -revision followed by a β -revision should be the same as an $\alpha \wedge \beta$ -revision. This can be seen as a strengthening of the postulate (K \ast 9), which was discussed in section 7.2. Such a property seems too strong for a general account of revision, although its reformulation in the context of nonmonotonic reasoning (see section 4.5) is one of the implicit assumptions made

about most nonmonotonic consequence relations in the literature. Future research in nonmonotonic reasoning will hopefully take this into account.

A consequence of $(P*)$ which is perhaps unexpected, is the following property.¹³

(Weak Symmetry) If $\alpha \not\models \neg\beta$ then $K((\Phi * \alpha) * \beta) = K((\Phi * \beta) * \alpha)$

Weak Symmetry suggests that it does not matter which of the observations α and β are made first, as long as α and β are compatible. At a first glance this seems at odds with Papini's intuition that more recent observations are deemed as more accurate, but closer inspection reveals this not to be the case. In fact, although the most recent observation is seen as more accurate than the previous one, both these observations are accorded higher prominence than any of the preceding observations. And as long as they are compatible, we should therefore expect the order in which these two observations were made, to be of no consequence; at least on the level of belief sets.

The intuition that Papini attaches to her construction seems to be in line with Lehmann's L-revision, and one would therefore expect it to satisfy all of Lehmann's postulates. However, as we shall see below, this turns out not to be the case. Since Papini's construction is a perfectly reasonable way of performing iterated revision, it would seem that some of Lehmann's postulates are a bit too restrictive.

Since $(L*1)$ – $(L*3)$ correspond exactly with $(E*1)$ – $(E*3)$, the former are satisfied by P_{\triangleright} -revision. P_{\triangleright} -revision also satisfies $(L*7)$, since the latter is a weakened version of $(DP*2)$. These, however, are the only of Lehmann's eight postulates that P_{\triangleright} -revision satisfies. Papini allows unsatisfiable belief sets, which violates $(L*6)$. For the remaining three of Lehmann's postulates, the following examples show that P_{\triangleright} -revision does not satisfy them.

Example 7.5.3 Let L be generated by the atoms p and q , with $V = \{00, 01, 10, 11\}$. Let Φ be an epistemic state such that $K(\Phi) = Cn(p \wedge q)$ and \preceq_{Φ} is defined as follows:

$$u \preceq_{\Phi} v \text{ iff } \begin{cases} v \in V \text{ if } u = 11, \\ v \in \{00, 01, 10\} \text{ if } u \in \{01, 10\}, \\ v = 00 \text{ if } u = 00. \end{cases}$$

¹³That $(P*)$ implies Weak Symmetry follows immediately by noting that $\alpha \models \neg\beta$ and $\beta \models \neg\alpha$ are equivalent.

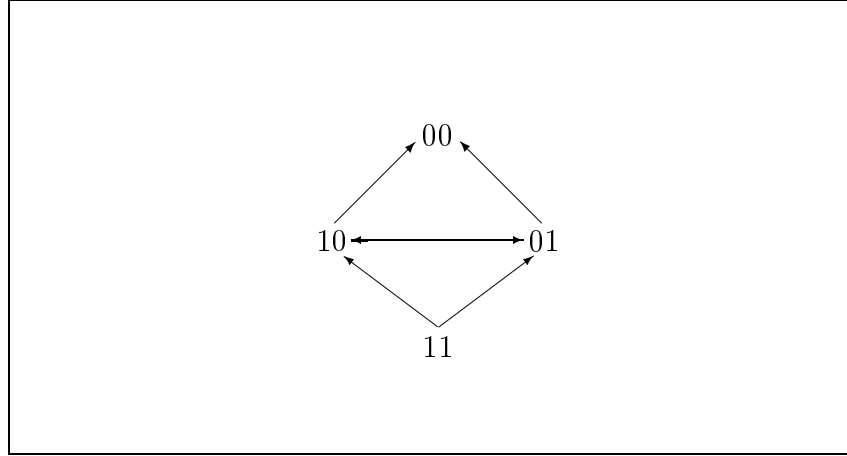


Figure 7.2: A graphical representation of the $K(\Phi)$ -faithful total preorder \preceq_Φ used in example 7.5.3. For two interpretations u and v , $u \preceq_\Phi v$ iff (u, v) is in the reflexive transitive closure of the relation determined by the arrows.

Figure 7.2 contains a graphical representation of \preceq_Φ . Now let \ast_\triangleright be the revision defined using (Def \ast_\triangleright). It is readily verified that $K((\Phi \ast_\triangleright p) \ast_\triangleright \neg(p \leftrightarrow q)) = Cn(p \wedge \neg q)$, but that $K(\Phi \ast_\triangleright \neg(p \leftrightarrow q)) = Cn(\neg(p \leftrightarrow q))$. And since $p \in K(\Phi)$, this is a violation of (L \ast 4). \square

Example 7.5.4 Let L be generated by the atoms p and q , with $V = \{00, 01, 10, 11\}$. Let Φ be an epistemic state such that $K(\Phi) = Cn(p)$ and \preceq_Φ is defined as follows:

$$u \preceq_\Phi v \text{ iff } \begin{cases} v \in V \text{ if } u \in \{10, 11\}, \\ v \in \{00, 01\} \text{ if } u \in \{00\}, \\ v = 01 \text{ if } u = 01. \end{cases}$$

Figure 7.3 contains a graphical representation of \preceq_Φ . Now let \ast_\triangleright be the revision defined using (Def \ast_\triangleright). It is readily verified that $K(((\Phi \ast_\triangleright p \vee q) \ast_\triangleright p) \ast_\triangleright \neg p) = Cn(\neg p \wedge q)$, but that $K((\Phi \ast_\triangleright p) \ast_\triangleright \neg p) = Cn(\neg p \wedge \neg q)$. And since $p \models p \vee q$, it is a violation of (L \ast 5). \square

Example 7.5.5 Let L be generated by the atoms p and q , with $V = \{00, 01, 10, 11\}$, let Φ be an epistemic state such that $K(\Phi) = Cn(\top)$ and $\preceq_\Phi = V \times V$, and let \ast_\triangleright be

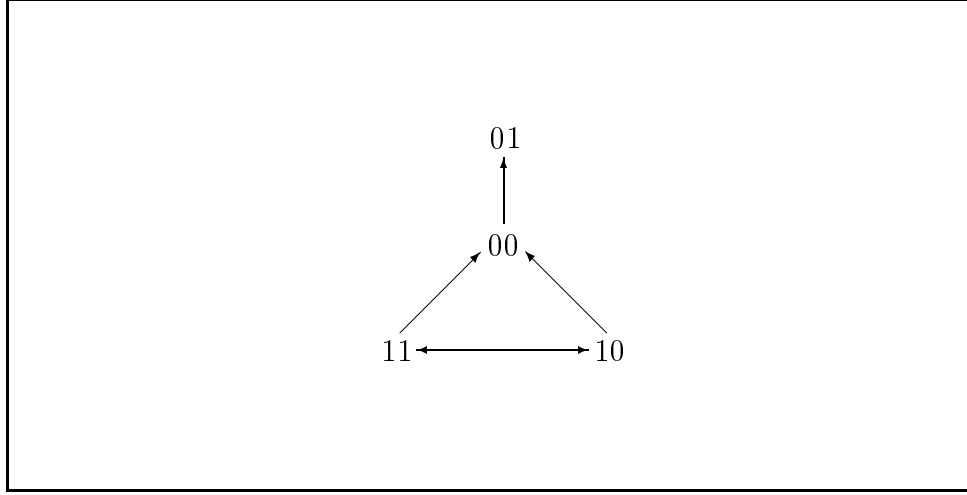


Figure 7.3: A graphical representation of the $K(\Phi)$ -faithful total preorder \preceq_Φ used in example 7.5.4. For two interpretations u and v , $u \preceq_\Phi v$ iff (u, v) is in the reflexive transitive closure of the relation determined by the arrows.

the revision defined in (Def \ast_\triangleright). It is readily verified that

$$K(((\Phi \ast_\triangleright p) \ast_\triangleright q) \ast_\triangleright \neg(p \leftrightarrow q)) = Cn(\neg p \wedge q)$$

but that

$$K(((\Phi \ast_\triangleright p) \ast_\triangleright p \wedge q) \ast_\triangleright \neg(p \leftrightarrow q)) = Cn(p \wedge \neg q).$$

And since $\neg q \notin K(\Phi \ast_\triangleright p) = Cn(p)$, this is a violation of (L \ast 8). \square

7.5.2 P_{\triangleleft} -revision

Papini also presents an operation that can be seen as dual to P_{\triangleright} -revision. Instead of letting the most recent observations carry the most weight, the situation is now reversed, with the most recent observations considered to be the least reliable. This revision operation is defined as follows:

$$(\text{Def } \ast_{\triangleleft}) \left[\begin{array}{l} u \preceq_{\Phi \ast_{\triangleleft} \alpha} v \text{ iff } \left\{ \begin{array}{l} u \preceq_\Phi v \text{ if } u \in M(\alpha) \text{ or } v \notin M(\alpha), \\ u \prec_\Phi v, \text{ otherwise} \end{array} \right. \\ K(\Phi \ast_{\triangleleft} \alpha) = \left\{ \begin{array}{l} Th(\text{Min}_{\preceq_{\Phi \ast_{\triangleleft} \alpha}}(\top)) \text{ if } K(\Phi) \neq Cn(\perp), \\ Cn(\perp) \text{ otherwise} \end{array} \right. \end{array} \right]$$

Definition 7.5.6 The revision \ast_{\triangleleft} defined in terms of $(\text{Def } \ast_{\triangleleft})$ is referred to as P_{\triangleleft} -revision. \square

Semantically speaking, a P_{\triangleleft} -revision of the epistemic state Φ by α has the following effect on \preceq_{Φ} . The relative ordering of valuations on different levels of \preceq_{Φ} are maintained, but each level is split into two by placing the models of α strictly below the models of $\neg\alpha$. The resulting belief set is then obtained from the minimal models in this new ordering, provided that the original belief set was satisfiable.

The intuition associated with P_{\triangleleft} -revision differs markedly from that normally associated with AGM-style revision, and it is not surprising that P_{\triangleleft} -revision does not satisfy all of (E*1)–(E*8). Papini shows that it satisfies (E*4), (E*5), (E*7), and the following weakened version of (E*3):

(WE*3) If $\neg\alpha \notin K(\Phi)$ then $K(\Phi \ast \alpha) \subseteq K(\Phi) + \alpha$

Furthermore, although it does not satisfy (E*2), (E*6), or (E*8), she shows that P_{\triangleleft} -revision satisfies the following dual versions of (E*2), (E*6) and the following weakened version of (E*8):

(DE*2) $K(\Phi) \subseteq K(\Phi \ast \alpha)$

(DE*6') If $K(\Phi \ast \alpha) = Cn(\perp)$ then $K(\Phi) = Cn(\perp)$

(WE*8') If $\neg\beta \notin K(\Phi \ast \alpha)$ and $\beta \in K(\Phi \ast \alpha \wedge \beta)$, then $K(\Phi \ast \alpha) + \beta \subseteq K(\Phi \ast \alpha \wedge \beta)$

That P_{\triangleleft} -revision does not satisfy (E*2) is to be expected, since it regards the most recent observation as the least reliable of all observations made thus far. The postulates (DE*2'), (DE*6') and (WE*8') are all in line with this view. (DE*2') requires all the wffs in Φ to be retained after a revision of Φ , and (DE*6') states that an α -revision of Φ will result in the unsatisfiable belief set only if Φ contained the unsatisfiable belief set to begin with. (WE*6') differs from (E*8) only in adding to the antecedent of (E*8) the requirement that $\beta \in K(\Phi \ast \alpha \wedge \beta)$.

We conclude this section by showing that the results above can be sharpened somewhat. Firstly, it is easily shown that P_{\triangleleft} -revision satisfies both (E*1) and (E*3). Moreover, we can improve on (DE*6') and (WE*8'). We show that the converse of (DE*6') also holds, and that the requirement added to the antecedent of (E*8) can be replaced with one that is, in our view, more natural.

(DE*6) $K(\Phi * \alpha) = Cn(\perp)$ iff $K(\Phi) = Cn(\perp)$

(WE*8) If $\neg\beta \notin K(\Phi * \alpha)$ and $\neg\alpha \notin K(\Phi)$, then $K(\Phi * \alpha) + \beta \subseteq K(\Phi * \alpha \wedge \beta)$

Proposition 7.5.7 P_{\triangleleft} -revision satisfies (E*1), (E*3), (DE*6) and (WE*8).

Proof (E*1) follows immediately from the semantic definition of P_{\triangleleft} -revision. For (E*3), we only need to consider the case where $\neg\alpha \in K(\Phi)$ because Papini has shown that P_{\triangleleft} -revision satisfies (WE*3), and the result then follows immediately. For (DE*6), we only need to show the right-to-left direction, and this follows immediately from (Def * \triangleleft). For (WE*8), suppose that $\neg\beta \notin K(\Phi * \alpha)$ and $\neg\alpha \notin K(\Phi)$. So $K(\Phi) \neq Cn(\perp)$. Since $\neg\alpha \notin K(\Phi)$ it follows from (Def * \triangleleft) that $K(\Phi * \alpha) = Th(Min_{\preceq_{\Phi}}(\alpha))$ and thus that $K(\Phi * \alpha) + \beta = Th(Min_{\preceq_{\Phi}}(\alpha) \cap M(\beta))$. Furthermore, it follows from $\neg\beta \notin K(\Phi * \alpha)$ that $Min_{\preceq_{\Phi}}(\alpha) \cap M(\beta) \neq \emptyset$. So $\neg(\alpha \wedge \beta) \notin K(\Phi)$ and it thus follows from (Def * \triangleleft) that $K(\Phi * \alpha \wedge \beta) = Th(Min_{\preceq_{\Phi}}(\alpha \wedge \beta)) = Th(Min_{\preceq_{\Phi}}(\alpha) \cap M(\beta))$. Therefore $K(\Phi * \alpha) + \beta = K(\Phi * \alpha \wedge \beta)$. \square

7.6 Merging epistemic states

While both of Papini's constructions may formally be viewed as revision operations, P_{\triangleleft} -revision does not quite conform to the intuition associated with revision. The reason for this is twofold. Firstly, revision has thus far referred to operations in which the wff with which to revise is *fully* accepted into the resulting belief set, and P_{\triangleleft} -revision thus represents a significant departure from this assumption. Secondly, the informal description of P_{\triangleleft} -revision, coupled with properties such as (DE*2) and (DE*6), suggests that it may also be seen as an operation in which a wff is being “revised” by an epistemic state, and not the other way around. The problem with the latter view, of course, is the asymmetry built into a revision on epistemic states; its first argument is an epistemic state, while its second argument is an element of L . To obtain the required symmetry, it is necessary to generalise the notion of revision. Instead of revising an epistemic state by a wff, we consider the process of revising an epistemic state by another epistemic state. In fact, since we wish to include cases where the second epistemic state is not regarded as more reliable than the first, it is more appropriate to refer to the *merging* of epistemic states, an area of research which has already received some

attention by Borgida and Imielinski [1984], Baral et al. [1991, 1992], Subrahmanian [1994], Liberatore and Schaerf [1998], Konieczny and Pino-Pérez [1998].

Formally then, a merge operation \otimes is a function from $\mathcal{E} \times \mathcal{E}$ to \mathcal{E} (where \mathcal{E} is the set of all epistemic states).

It is not our intention to provide a detailed discussion of merging here. At this stage, we merely wish to argue that merging is an area of research which needs to be investigated more thoroughly, and to put forward some basic properties with which every merge operation should comply. We also consider some specific merge operations, one of which bolsters the claim that revision on epistemic states is indeed a special case of merging.

7.6.1 Basic properties of merge operations

Intuitively, the merging of epistemic states is intended to be a coherent fusion of the information contained in both. There is no assumption that one of the epistemic states is deemed to be more reliable than the other. Instead, merging is intended to cover the whole spectrum; from the case where the first epistemic state takes absolute priority over the second one, to the case where the second epistemic state has complete precedence over the first one. Our point of departure in this investigation is the assumption that every epistemic state Φ has associated with it a belief set $K(\Phi)$ and a $K(\Phi)$ -faithful total preorder \preceq_Φ . The information contained in two epistemic states Φ and Ψ to be merged, does not just refer to the beliefs contained in $K(\Phi)$ and $K(\Psi)$, but also to the information contained in the orderings \preceq_Φ and \preceq_Ψ . Observe that the idea is still one of a minimal model semantics. Given the fact that $\preceq_{\Phi \otimes \Psi}$ has to be a $K(\Phi \otimes \Psi)$ -faithful total preorder, this assumption is built into the definition of an epistemic state.

With these guidelines in mind, we propose the following general properties for merging:

$$(\otimes 1) \quad K(\Phi) \cap K(\Psi) \subseteq K(\Phi \otimes \Psi)$$

$$(\otimes 2) \quad K(\Phi \otimes \Psi) \subseteq Cn(K(\Phi) \cup K(\Psi))$$

$$(\otimes 3) \quad \text{If } K(\Phi) \neq Cn(\perp) \text{ and } K(\Psi) \neq Cn(\perp) \text{ then } K(\Phi \otimes \Psi) \neq Cn(\perp)$$

$$(\otimes 4) \quad \text{If } K(\Phi) = K(\Omega) \text{ and } K(\Psi) = K(\Upsilon) \text{ then } K(\Phi \otimes \Psi) = K(\Phi \otimes \Upsilon)$$

These four properties involve the belief set obtained from a merge operation. $(\otimes 1)$ provides a lower bound on the resulting belief set. It states that the new belief set has to contain those beliefs associated with both epistemic states to be merged. $(\otimes 2)$ on the other hand, provides an upper bound for the resulting belief set. It may not contain any belief which does not occur in at least one of the two epistemic states to be merged. $(\otimes 3)$ requires that the resulting belief set be unsatisfiable only if at least one of the belief sets associated with the two epistemic states to be merged are unsatisfiable. And $(\otimes 4)$ is an expression of the principle of the Irrelevance of Syntax, applied to the belief sets associated with epistemic states.

The next two properties are concerned with the faithful total preorder resulting from a merge operation.

$(\otimes 5)$ If $u \preceq_{\Phi} v$ and $u \preceq_{\Psi} v$ then $u \preceq_{\Phi \otimes \Psi} v$

$(\otimes 6)$ If $u \preceq_{\Phi \otimes \Psi} v$ then $u \preceq_{\Phi} v$ or $u \preceq_{\Psi} v$

Both $(\otimes 5)$ and $(\otimes 6)$ are motivated by the intuition that the merging of two epistemic states Φ and Ψ depends, in the first place, on the information contained in Φ and Ψ . $(\otimes 5)$ states that information contained in both Φ and Ψ should also occur in $\Phi \otimes \Psi$. $(\otimes 6)$ is almost the converse of $(\otimes 5)$. It asserts that information contained in $\Phi \otimes \Psi$ must have been obtained from either Φ or Ψ .

7.6.2 Constructing merge operations

In this section we take a brief look at the construction of some merge operations. The first two we have in mind represent the two extremes on the spectrum of merging. They involve the cases where one of the two epistemic states to be merged takes complete precedence over the other, and can be defined as follows:

(Def \otimes_{\leftarrow}) $\Phi \otimes_{\leftarrow} \Psi = \Phi$

(Def \otimes_{\rightarrow}) $\Phi \otimes_{\rightarrow} \Psi = \Psi$

It is easily verified that the merge operations defined using $(\text{Def } \otimes_{\rightarrow})$ and $(\text{Def } \otimes_{\leftarrow})$ both satisfy $(\otimes 1)$ – $(\otimes 6)$.

The next two merge operations to be presented can also be seen as opposites. In this case though, it involves a preference for one epistemic state over the other which

is of a less extreme kind. For their definition, we need to broaden the definition of the minimal models of a wff to apply to sets of interpretations.

Definition 7.6.1 For a binary relation \preceq on V , and $W \subseteq V$, we let $Min_{\preceq}(W) = \{v \mid v \text{ is } \preceq\text{-minimal in } W\}$. \square

So $Min_{\preceq}(W)$ is the set of \preceq -minimal elements in W . The two merge operations are defined as follows:

$$\begin{aligned}
 (\text{Def } \otimes_{\triangleleft}) \quad & \left[\begin{array}{l} K(\Phi \otimes_{\triangleleft} \Psi) = Min_{\preceq_{\Psi}}(M(K(\Phi))) \\ u \preceq_{\Phi \otimes_{\triangleleft} \Psi} v \text{ iff } u \prec_{\Phi} v \text{ or } (u \equiv_{\preceq_{\Phi}} v \text{ and } u \preceq_{\Psi} v) \end{array} \right] \\
 (\text{Def } \otimes_{\triangleright}) \quad & \left[\begin{array}{l} K(\Phi \otimes_{\triangleright} \Psi) = Min_{\preceq_{\Phi}}(M(K(\Psi))) \\ u \preceq_{\Phi \otimes_{\triangleright} \Psi} v \text{ iff } u \prec_{\Psi} v \text{ or } (u \equiv_{\preceq_{\Psi}} v \text{ and } u \preceq_{\Phi} v) \end{array} \right]
 \end{aligned}$$

These two merge operations can perhaps best be described as lexicographic orderings of the faithful total preorders associated with the epistemic states; \otimes_{\triangleleft} ensures that Ψ orders \preceq_{Φ} lexicographically, while \otimes_{\triangleright} ensures that Φ orders \preceq_{Ψ} lexicographically. Again, both these merge operations satisfy $(\otimes 1)$ – $(\otimes 6)$.

Proposition 7.6.2 *The merge operations \otimes_{\triangleleft} and \otimes_{\triangleright} defined using (Def \otimes_{\triangleleft}) and (Def \otimes_{\triangleright}) respectively, both satisfy $(\otimes 1)$ – $(\otimes 6)$.*

Proof For $(\otimes 1)$ and $(\otimes 2)$, observe that $M(K(\Phi)) \cap M(K(\Psi)) \subseteq Min_{\preceq_{\Psi}}(M(K(\Phi))) \subseteq M(K(\Phi))$ and that $M(K(\Phi)) \cap M(K(\Psi)) \subseteq Min_{\preceq_{\Phi}}(M(K(\Psi))) \subseteq M(K(\Psi))$. For $(\otimes 3)$, note that if $K(\Phi \otimes_{\triangleleft} \Psi) = Cn(\perp)$ then $M(K(\Phi)) = \emptyset$, and if $K(\Phi \otimes_{\triangleright} \Psi) = Cn(\perp)$ then $M(K(\Psi)) = \emptyset$. $(\otimes 4)$ is trivial. For $(\otimes 5)$, suppose that $u \preceq_{\Phi} v$ and $u \preceq_{\Psi} v$. If $u \prec_{\Phi} v$ then $u \preceq_{\Phi \otimes_{\triangleleft} \Psi} v$, and if $u \equiv_{\preceq_{\Phi}} v$ then $u \preceq_{\Phi \otimes_{\triangleleft} \Psi} v$ since $u \preceq_{\Psi} v$. The case for \otimes_{\triangleright} is similar. For $(\otimes 6)$, suppose that $u \preceq_{\Phi \otimes_{\triangleleft} \Psi} v$. Then it has to be the case that $u \preceq_{\Phi} v$ and so $(\otimes 6)$ holds for \otimes_{\triangleleft} . The case for \otimes_{\triangleright} is similar. \square

The merge operation \otimes_{\triangleright} defined using (Def \otimes_{\triangleright}) corresponds to a proposal of Nayak [Nayak, 1994b, Nayak et al., 1996]. His FPO (fixed point ordering) revision operation is a generalisation of AGM revision based on modified versions of the EE-orderings (see 2.3), but it is clear from his semantic description [Nayak, 1994b] that it is, essentially, the same construction as \otimes_{\triangleright} . Furthermore, Papini's P_{\triangleright} -revision can be seen as a

special case of \otimes_{\triangleright} , while her P_{\triangleleft} -revision can be seen as a special case of the merge operation \otimes_{\triangleleft} defined using (Def \otimes_{\triangleleft}). It is simply a matter of associating with every wff α the unique $Cn(\alpha)$ -faithful total preorder in which the countermodels of α are all on the same level. That is, every wff α is associated with the epistemic state Ψ_{α} where $K(\Psi_{\alpha}) = Cn(\alpha)$ and $u \preceq_{\Psi_{\alpha}} v$ iff $u \in M(\alpha)$ or $v \in M(\neg\alpha)$. It then follows immediately that $\Phi \otimes_{\triangleright} \Psi_{\alpha} = \Phi *_{\triangleright} \alpha$ and $\Phi \otimes_{\triangleleft} \Psi_{\alpha} = \Phi *_{\triangleleft} \alpha$ for every $\Phi \in \mathcal{E}$ and every $\alpha \in L$.

Finally, we propose a class of merge operations which regard the two epistemic states to be merged as equally important; at least on the level of belief sets. Revesz [1993] uses the term “arbitration” for merge operations conforming to this intuition. Information-theoretically, our proposal draws a distinction between two cases. Firstly, if the two epistemic states Φ and Ψ to be merged are compatible on the level of belief sets, the belief set resulting from an arbitration of Φ and Ψ are obtained by pooling the content bits of $K(\Phi)$ and $K(\Psi)$. Secondly, if Φ and Ψ are incompatible on the level of belief sets, the belief set resulting from an arbitration of Φ and Ψ is built up using those infatoms that are content bits of $K(\Phi)$ as well as $K(\Psi)$.

$$(\text{Def } K(\hat{\otimes})) \quad K(\Phi \hat{\otimes} \Psi) = \begin{cases} K(\Phi) \cap K(\Psi) & \text{if } K(\Phi) \cup K(\Psi) \models \perp, \\ Cn(K(\Phi) \cup K(\Psi)) & \text{otherwise} \end{cases}$$

Definition 7.6.3 A merge operation $\hat{\otimes}$ on epistemic states is an *arbitration* iff the belief set $K(\Phi \hat{\otimes} \Psi)$ associated with the arbitration of two epistemic states can be defined using (Def $K(\hat{\otimes})$). \square

Arbitration, as defined above, is only concerned with belief sets, and therefore it does not satisfy $(\otimes 5)$ or $(\otimes 6)$. However, it does satisfy the remaining properties for merging.

Proposition 7.6.4 *Every arbitration satisfies $(\otimes 1)$ to $(\otimes 4)$.*

Proof Pick any arbitration $\hat{\otimes}$. $(\otimes 1)$, $(\otimes 2)$ and $(\otimes 4)$ are trivial. For $(\otimes 3)$, observe that if $K(\Phi \hat{\otimes} \Psi) = Cn(\perp)$ then $K(\Phi) \cup K(\Psi) \models \perp$ and thus $K(\Phi \hat{\otimes} \Psi) = K(\Phi) \cap K(\Psi)$. And $K(\Phi) \cap K(\Psi) = Cn(\perp)$ iff $K(\Phi) = K(\Psi) = Cn(\perp)$. \square

Liberatore and Schaerf [1998] propose a class of merge operations which is similar in spirit to definition 7.6.3. They provide the following eight postulates for arbitration operations.

$$(\text{LS}\hat{\otimes}1) \quad K(\Phi \hat{\otimes} \Psi) = K(\Psi \hat{\otimes} \Phi)$$

$$(\mathbf{LS}\hat{\otimes}2) \quad K(\Phi\hat{\otimes}\Psi) \subseteq Cn(K(\Phi) \cup K(\Psi))$$

$$(\mathbf{LS}\hat{\otimes}3) \quad \text{If } K(\Phi) \cup K(\Psi) \not\models \perp \text{ then } Cn(K(\Phi) \cup K(\Psi)) \subseteq K(\Phi\hat{\otimes}\Psi)$$

$$(\mathbf{LS}\hat{\otimes}4) \quad K(\Phi\hat{\otimes}\Psi) = Cn(\perp) \text{ iff } K(\Phi) = K(\Psi) = Cn(\perp)$$

$$(\mathbf{LS}\hat{\otimes}5) \quad \text{If } K(\Phi) = K(\Omega) \text{ and } K(\Psi) = K(\Upsilon) \text{ then } K(\Phi\hat{\otimes}\Psi) = K(\Omega\hat{\otimes}\Upsilon)$$

$$(\mathbf{LS}\hat{\otimes}6) \quad \text{If } K(\Psi) = K(\Omega) \cap K(\Upsilon) \text{ then } K(\Phi\hat{\otimes}\Psi) = \begin{cases} K(\Phi\hat{\otimes}\Omega) \text{ or} \\ K(\Phi\hat{\otimes}\Upsilon) \text{ or} \\ K(\Phi\hat{\otimes}\Omega) \cap K(\Phi\hat{\otimes}\Upsilon) \end{cases}$$

$$(\mathbf{LS}\hat{\otimes}7) \quad K(\Phi) \cap K(\Psi) \subseteq K(\Phi\hat{\otimes}\Psi)$$

$$(\mathbf{LS}\hat{\otimes}8) \quad \text{If } K(\Phi) \neq Cn(\perp) \text{ then } K(\Phi) \cup K(\Phi\hat{\otimes}\Psi) \not\models \perp$$

We show that an arbitration, in the sense of definition 7.6.3, satisfies all eight of these postulates.

Proposition 7.6.5 *Every arbitration $\hat{\otimes}$ satisfies $(\mathbf{LS}\hat{\otimes}1)$ to $(\mathbf{LS}\hat{\otimes}8)$.*

Proof $(\mathbf{LS}\hat{\otimes}1)$ – $(\mathbf{LS}\hat{\otimes}5)$ and $(\mathbf{LS}\hat{\otimes}7)$ are trivial. Now pick any arbitration $\hat{\otimes}$. For $(\mathbf{LS}\hat{\otimes}6)$, pick any $\Psi, \Omega, \Upsilon \in \mathcal{E}$ such that $K(\Psi) = K(\Omega) \cap K(\Upsilon)$. We need to consider four cases. First we consider the case where both $K(\Phi) \cup K(\Omega) \not\models \perp$ and $K(\Phi) \cup K(\Upsilon) \not\models \perp$. Then $K(\Phi\hat{\otimes}\Omega) = Cn(K(\Phi) \cup K(\Omega))$ and $K(\Phi\hat{\otimes}\Upsilon) = Cn(K(\Phi) \cup K(\Upsilon))$. So

$$\begin{aligned} & K(\Phi\hat{\otimes}\Omega) \cap K(\Phi\hat{\otimes}\Upsilon) \\ &= Cn(K(\Phi) \cup K(\Omega)) \cap Cn(K(\Phi) \cup K(\Upsilon)) \\ &= Cn(K(\Phi) \cup (K(\Omega) \cap K(\Upsilon))) \\ &= Cn(K(\Phi) \cup K(\Psi)) \\ &= K(\Phi\hat{\otimes}\Psi) \text{ since } K(\Phi) \cup K(\Psi) \not\models \perp. \end{aligned}$$

Next we consider the case where both $K(\Phi) \cup K(\Omega) \models \perp$ and $K(\Phi) \cup K(\Upsilon) \models \perp$. Then $M(K(\Phi)) \cap M(K(\Omega)) = M(K(\Phi)) \cap M(K(\Upsilon)) = \emptyset$, and so

$$\begin{aligned} & M(K(\Phi)) \cap M(K(\Psi)) \\ &= M(K(\Phi)) \cap M(K(\Omega) \cap K(\Upsilon)) \\ &= M(K(\Phi)) \cap (M(K(\Omega)) \cup M(K(\Upsilon))) \\ &= (M(K(\Phi)) \cap M(K(\Omega))) \cup (M(K(\Phi)) \cap M(K(\Upsilon))) \\ &= \emptyset \end{aligned}$$

Therefore $K(\Phi) \cup K(\Psi) \models \perp$. Furthermore, $K(\Phi \hat{\otimes} \Omega) = K(\Phi) \cap K(\Omega)$ and $K(\Phi \hat{\otimes} \Upsilon) = K(\Phi) \cap K(\Upsilon)$, and so

$$\begin{aligned}
& K(\Phi \hat{\otimes} \Omega) \cap K(\Phi \hat{\otimes} \Upsilon) \\
&= (K(\Phi) \cap K(\Omega)) \cap (K(\Phi) \cap K(\Upsilon)) \\
&= K(\Phi) \cap K(\Omega) \cap K(\Upsilon) \\
&= K(\Phi) \cap K(\Psi) \\
&= K(\Phi \hat{\otimes} \Psi) \text{ since } K(\Phi) \cup K(\Psi) \models \perp.
\end{aligned}$$

Finally, we consider the case where $K(\Phi) \cup K(\Omega) \models \perp$ and $K(\Phi) \cup K(\Upsilon) \not\models \perp$. (The remaining case, where $K(\Phi) \cup K(\Omega) \not\models \perp$ and $K(\Phi) \cup K(\Upsilon) \models \perp$, is similar.) Then $K(\Phi) \cup K(\Psi) \not\models \perp$ and so

$$\begin{aligned}
& K(\Phi \hat{\otimes} \Psi) \\
&= Cn(K(\Phi) \cup K(\Psi)) \\
&= Cn(K(\Phi) \cup (K(\Omega) \cap K(\Upsilon))) \\
&= Cn\left((K(\Phi) \cup K(\Omega)) \cap (K(\Phi) \cup K(\Upsilon))\right) \\
&= Cn(K(\Phi) \cup K(\Upsilon)) \text{ since } K(\Phi) \cup K(\Omega) \models \perp \\
&= K(\Phi \hat{\otimes} \Upsilon) \text{ since } K(\Phi) \cup K(\Upsilon) \not\models \perp.
\end{aligned}$$

For (LS $\hat{\otimes}$ 8), suppose that $K(\Phi) \neq Cn(\perp)$. If $K(\Phi) \cup K(\Psi) \not\models \perp$ then $K(\Phi \hat{\otimes} \Psi) = Cn(K(\Phi) \cup K(\Psi)) \neq Cn(\perp)$ and so $K(\Phi \hat{\otimes} \Psi) \cup K(\Phi) \not\models \perp$. And if $K(\Phi) \cup K(\Psi) \models \perp$ then $K(\Phi \hat{\otimes} \Psi) = K(\Phi) \cap K(\Psi)$ and since $K(\Phi) \neq Cn(\perp)$, it follows that $K(\Phi \hat{\otimes} \Psi) \cup K(\Phi) \not\models \perp$. \square

Finally, observe that there are some similarities between the properties for merge operations that we have proposed, and the postulates of Liberatore and Schaerf. In particular, $(\otimes 1)$ and (LS $\hat{\otimes}$ 7) are identical, $(\otimes 2)$ and (LS $\hat{\otimes}$ 2) are identical, $(\otimes 4)$ and (LS $\hat{\otimes}$ 5) are identical, and $(\otimes 3)$ corresponds to the one direction of (LS $\hat{\otimes}$ 4). The remaining postulates of Liberatore and Schaerf seem to be specifically concerned with arbitration, and are thus not suitable as properties for the more general notion of a merge operation. On the other hand, $(\otimes 5)$ and $(\otimes 6)$ are concerned with the faithful total orders associated with epistemic states, and have no counterparts among the postulates of Liberatore and Schaerf, which are only concerned with the belief sets associated with epistemic states.

7.7 Conclusion

Questions concerning iterated belief change can be traced back to a violation of the principle of Categorical Matching in the AGM approach to belief change. The latter requires an epistemic state to perform belief change operations, but delivers just a belief set. The AGM postulates can thus be seen as constraints placed on just one part of the epistemic state of an agent. This realisation has prompted various authors to extend the AGM postulates in order to place constraints on the other parts of an epistemic state as well. While the work of Spohn [1988, 1991] has been instrumental in this regard, the account provided by Darwiche and Pearl [1994, 1997] is arguably the most influential. Although some of the postulates they provide seem too strong, their decision to associate with every epistemic state a unique faithful total preorder has proved to play a central role in the understanding of their constraints on epistemic states pertaining to theory revision.

A semantic consideration of epistemic states also promises to have a significant impact on the investigation of the merging operations of section 7.6. Much work still needs to be done in this area, but the work of Nayak [1994b], Nayak et al. [1996], Liberatore and Schaerf [1998] and Konieczny and Pino-Pérez [1998] have opened fruitful areas of investigation.

Chapter 8

Infobase change

*It is undesirable to believe a proposition when
there is no ground whatsoever for believing it true.*

Bertrand Russell

We have seen in section 7.2 that frameworks for belief change which operate on the level of belief sets are not rich enough in structure to provide a proper treatment of change operations. In particular, from the work of Darwiche and Pearl [1994, 1997], it has emerged that belief change ought to be described on the level of epistemic states. While the proposal of Darwiche and Pearl is an important contribution to the enterprise of belief change on an abstract level, it does not address the equally important question of what it is that prompts an agent to adopt a particular epistemic state in a given situation. In this chapter we investigate an approach to find a solution to this problem using structures that we refer to as *infobases*.

The assumption underlying infobase change is that an agent obtains information (in the form of wffs of L) which is to be stored in an infobase; a *finite* sequence of wffs consisting of information obtained independently from different sources. Infobases thus have more structure than finite sets of wffs.¹ From this description it might seem as if infobase change is a slightly generalised instance of *base change*, the proposal to

¹This chapter is an expanded version of the paper by Meyer et al. [1999a]. In that paper we took an infobase to be a finite *set* of wffs, but acknowledged at the same time that such a representation is not entirely satisfactory.

replace change operations on belief sets with change operations on arbitrary sets of wffs (known as belief bases). While it is indeed possible to classify infobase change as such, the phrase “base change” has become so synonymous with the particular kind of base change championed by Fuhrmann and Hansson in particular, that it is perhaps more appropriate to regard infobase change as an altogether different kind of belief change. Section 8.1 contains a brief discussion of base change. It is not intended as a comprehensive introduction to the field, but is included primarily for purposes of comparison with infobase change.

8.1 Base change

The realisation that belief sets do not have a rich enough structure to serve as appropriate models for epistemic states (see section 7.2) has led some researchers to regard AGM theory change as an elegant idealisation of a more general theory of belief change in which belief sets are replaced by arbitrary sets of wffs (known as belief bases).² The intuition is that some of our beliefs have no independent standing, but arise only as beliefs derived from our more basic beliefs. And if our reason for believing such a derived belief disappears, then so should the belief. Martins and Shapiro [1988] refer to this process as disbelief propagation. It is also known as *reason maintenance* [Doyle, 1979], and is the principle underlying Fuhrmann’s [1991] filtering condition, which we encounter in section 8.2.3.

A belief base B is taken to consist of such basic beliefs, with B being *associated with* the belief set K (and K being the belief set *associated with* a belief base B) iff $Cn(B) = K$. The classic example in the base change literature (perhaps analogous to the Tweety example in nonmonotonic reasoning) is Hansson’s hamburger example.

Example 8.1.1 [Hansson, 1989] “On a public holiday you are standing in the street in a town that has two hamburger restaurants. Let us consider the subset of your belief set that represents your beliefs about whether or not each of these two restaurant is open.

When you meet me, eating a hamburger, you draw the conclusion that at least one of the restaurant is open ($a \vee b$). Further, seeing from a distance that one of the

²Although the original AGM postulates are not exclusively concerned with belief sets, the major results in Alchourrón et al. [1985] only hold for belief sets.

two restaurants has its lights on, you believe that this particular restaurant is open (a). This situation can be represented by the set of beliefs $\{a, a \vee b\}$. When you have reached the restaurant however, you find a sign saying that it is closed all day. The lights are only turned on for the purposes of cleaning. You now have to include the negation of a , i.e. $\neg a$, into your belief set. The revision of $\{a, a \vee b\}$ to include $\neg a$ should still contain $a \vee b$, since you still have reason to believe that one of the two restaurants is open.

In contrast, suppose you had not met me or anyone else eating a hamburger. Then your only clue would have been the lights from the restaurant. The original belief system in this case can be represented by the set $\{a\}$. After finding out that the restaurant is closed, the resulting set should not contain $a \vee b$, since in this case you have no reason to believe that one of the restaurants is open.” \square

The difference in the treatment of the belief bases $\{a\}$ and $\{a, a \vee b\}$ is attributable to the fact that $a \vee b$ is an explicit belief with independent standing in $\{a, a \vee b\}$, while it is a mere derived belief of the belief base $\{a\}$. The two belief bases should therefore treat an a -contraction differently even though $Cn(a) = Cn(a, a \vee b)$.

One of the basic principles of base change is that it is sensitive to syntax. What is usually *not* made explicit, though, is that such an assertion can be interpreted in many ways. In the context of belief change, this sensitivity to syntax usually refers to the following two properties:

1. Belief bases offer a finer-grained approach than belief sets in the sense that two different belief bases may both be associated with the same belief set.
2. Contraction is interpreted on the symbol level and not on the knowledge level (see page 3). In particular, this means that a base contraction \sim is expected to satisfy the property of Inclusion, which requires that $B \sim \alpha \subseteq B$, and not merely that $Cn(B \sim \alpha) \subseteq Cn(B)$.

Observe that there are other ways for base change to be sensitive to syntax as well. To name just two, a change effected by two logically equivalent wffs may be treated differently, or belief bases containing different but logically equivalent wffs may be treated differently.

Even though base change is more sensitive to syntax than theory change, it is not intended to be totally oblivious to knowledge level matters. For example, a base change

operation \sim is expected to satisfy the property that $\alpha \notin Cn(B \sim \alpha)$ whenever $\not\models \alpha$, which involves the belief set associated with the base $B \sim \alpha$ as well.

Descriptions of base change usually subscribe to some form of Levi's commensurability thesis (see page 7), and contraction is thus defined explicitly, while revision is defined in terms of some analogue of the Levi Identity (the identity (Def \sim from *)). Both Fuhrmann [1991] and Hansson [1989, 1992a, 1993b] define versions of base contraction which can be viewed as generalisations of theory contraction in which the contraction of belief sets is a special case. Accordingly, their methods for constructing these base contraction operations are appropriate generalisations of methods for constructing (basic AGM) theory contractions. Fuhrmann generalises the entailment sets used to construct safe contractions (see section 2.4), while Hansson generalises the remainders used to construct partial meet contractions (see section 2.2).

Base contractions are operations on belief bases, but it has been pointed out by Nebel [1989] and Fuhrmann [1991], amongst others, that there is a theory contraction – *associated with* every base contraction \sim , which can be obtained as follows: $Cn(B) - \alpha = Cn(B \sim \alpha)$.³ In this way it is possible to provide a knowledge level analysis of base contraction, and to make (indirect) comparisons between base contraction and theory contraction.

One of the first observations to be made in this regard concerns the controversial Recovery postulate for theory contraction. Given the symbol level interpretation of base contraction, a simple example suffices to show that Recovery does not hold for the associated theory contractions.

Example 8.1.2 Let $B = \{p\}$ and let \sim be a base contraction. Given the restrictions that $B \sim \alpha \subseteq B$ and that $\alpha \notin Cn(B \sim \alpha)$ if $\not\models \alpha$, it has to be the case that $B \sim p \vee q = \emptyset$. Therefore $Cn(B) \neq Cn(B \sim p \vee q) + p \vee q$ even though $p \vee q \in Cn(B)$; a violation of Recovery. \square

With the emphasis on the syntactic structure of a belief base, it has been remarked by Gärdenfors and Rott [1995, p. 87] that a semantic characterisation of base change seems to be out of the question. It is possible, though, to obtain an indirect semantic characterisation of bases change, by focussing on the theory change operations associated

³This construction only makes sense for a fixed belief base B , though, since the same belief set may be associated with different belief bases, which may violate the assumed functionality of theory contraction.

with base change operations. Hansson [1996] provides postulates and representation results for the theory contractions associated with some base contraction operations.

A different (though not completely unrelated) view of base change is that more emphasis should be placed on knowledge level matters, and that a belief base should be thought of as providing more structure to its associated belief set. The idea is that the added structure of a belief base can be used, in one way or another, to pick an appropriate associated theory change operation, from which the base change operation can then be constructed. This is the view encountered in Nebel's [1989, 1990, 1991, 1992] description of base contraction. Nebel himself describes his own work as a knowledge level analysis of base change. In concentrating on knowledge level matters, his construction violates one of the cornerstones of base contraction; the property of Inclusion, which requires of a base contraction \sim to satisfy $B \sim \alpha \subseteq B$. This violation has resulted in these operations being labelled as pseudo-contraction by Hansson [1993a, 1999].

In conclusion, observe that if one is interested in moving towards a realistic representation of the epistemic states of agents, it seems reasonable to insist that such a representation be finite. Such a move is sometimes held up as a reason for preferring base change to theory change. But to do so, is to disregard the distinction between an arbitrary finite representation of a particular belief set, and a set of finite wffs occurring in a belief base because of their independent standing. For example, recall from section 3.2.1 that Katsuno and Mendelzon use single propositional wffs to represent belief sets. But since the particular wff representing a belief set is unimportant, their work should be classified as research about theory change, and not about base change.

8.2 Constructing infobase change

Infobase change is similar in spirit to the knowledge level approach to base change favoured by Nebel [1989]. The basic idea is to use the assumption of independence of the wffs in an infobase IB to construct the structures necessary for performing *theory* change. Both the current infobase and the obtained theory change operations are then used in the process of determining how to modify the existing infobase when confronted with new information, resulting in an operation which produces a new infobase from the current one.

An infobase will be represented as a list of wffs enclosed by square brackets. For

example, the infobase IB containing the three wffs p , q and p , in that order, will be denoted as $[p, q, p]$. Although infobases are sensitive to the order in which wffs occur, as well as to their syntactical form, we shall see that these superficial qualities can be done away with by employing the notion of element-equivalence.

Definition 8.2.1 Two infobases IB and IC are *element-equivalent*, written as $IB \approx IC$, iff for every β occurring in IB such that $\not\models \beta$, there is a unique logically equivalent wff γ occurring in IC , and for every γ occurring in IC such that $\not\models \gamma$, there is a unique logically equivalent wff β occurring in IB . \square

The intuition is that element-equivalent infobases contain exactly the same information.

For any finite sequence σ of wffs, we let $|\sigma|$ denote the *number* of wffs occurring in σ , and we use the symbol \bullet to denote concatenation by a single wff. Thus, if $\sigma = [p, q]$, then $|\sigma| = 2$, $[p, q] \bullet p$ denotes the sequence $[p, q, p]$, and $|\sigma \bullet p| = 3$. The converse of concatenation (removing the last wff from a finite sequence σ) will be denoted by $\bar{\sigma}$. In other words, if $\sigma = [p, q, r]$ then $\bar{\sigma} = [p, q]$. Furthermore, in our discussion of the construction of infobase change operations it will frequently be necessary to refer to the (finite) *set* of wffs occurring in a finite sequence of wffs σ . We denote this set by $S(\sigma)$. Thus, for any finite sequence σ of wffs, $S(\sigma) = \{\beta \mid \beta \text{ occurs in } \sigma\}$.

Definition 8.2.2 An infobase IB is *associated* with a belief set K (and K is *associated* with IB) iff $Cn(S(IB)) = K$. \square

For any two finite sequences σ and τ of wffs, τ is a *subsequence* of σ iff for every wff in τ there is a unique occurrence of the same wff in σ . τ is an *ordered subsequence* of σ iff τ is a subsequence of σ and the wffs in τ occur in the same order in both τ and σ .

In our description of infobase change, we subscribe to Levi's commensurability thesis, by viewing infobase contraction as more primitive than infobase revision, and preferring to define infobase revision in terms of infobase contraction by means of an infobase change analogue of the Levi Identity (see definition 2.1.1). Formally, we consider infobase change operations (which include contraction and revision operations) as functions from $\mathcal{IB} \times L$ to \mathcal{IB} , where \mathcal{IB} is the set of all infobases. We shall also frequently assume the existence of a fixed infobase IB , and consider infobase IB -change operations as functions from L to \mathcal{IB} .

8.2.1 Infobase contraction

To construct an infobase contraction, we first use the structure of the infobase IB to obtain an $S(IB)$ -faithful total preorder (see definition 3.2.5). The theory contraction obtained from the $S(IB)$ -faithful total preorder is taken to be the theory contraction associated with the infobase contraction that we aim to construct.

Definition 8.2.3 For every infobase IB , a theory contraction $-$ is *associated with* an infobase IB -contraction \ominus iff $Cn(IB) - \alpha = Cn(IB \ominus \alpha)$ for every $\alpha \in L$. \square

Using the intuition associated with an infobase, we order the interpretations in U according to the number of wffs of IB they satisfy; the more they satisfy, the “better” they are deemed to be, and the lower down in the ordering they will be.

Definition 8.2.4 For $u \in U$, we define u_{IB} , the *IB-number of u* , as the number of wffs β in IB such that $u \models \beta$ and $u \in M(\beta)$. \square

This ordering is used to obtain an appropriate $S(IB)$ -faithful total preorder in terms of IB as follows:

(Def \preceq from IB) $u \preceq v$ iff $u_{IB} \leq v_{IB}$

Definition 8.2.5 We refer to the faithful total preorder \preceq_{IB} defined in terms of an infobase IB using (Def \preceq from IB) as the *IB-induced* faithful total preorder. \square

The construction of the IB -induced faithful total preorders is perhaps best justified from an information-theoretic point of view (see section 3.1). Suppose that the infobase IB represents the information that an agent has obtained from its sources. Since the wffs in IB are assumed to have been obtained independently, every occurrence of an infatom i as a content bit of one of these wffs, corroborates the claim that i forms part of the content bits of the belief set $Cn(S(IB))$ of the agent. From definition 3.1.3 on 34 it can be verified that a wff β is satisfied by an interpretation u iff the infatom i_u associated with u is a content bit of $\neg\beta$. So, with \preceq_{IB} seen as an ordering on infatoms, it follows that being *higher* up in \preceq_{IB} corresponds to *more* occurrences of an infatom as the content bit of some wffs in IB , which is in line with the view of a faithful total preorder as an ordering of entrenchment or credibility on infatoms (see 3.2, page 44). Note also that since logically valid wffs have no content bits, their presence in an infobase is superfluous since they do not contribute towards the entrenchment or

credibility of any of the infatoms. This explains why the definition of IB -numbers disregards the logically valid wffs in infobases.

The IB -induced faithful total preorder is used to construct a theory contraction as follows:

(Def $-_{IB}$ from IB) $Cn(S(IB)) -_{IB} \alpha = Th(M(S(IB)) \cup Min_{\preceq_{IB}}(\neg\alpha))$

Definition 8.2.6 The theory contraction $-_{IB}$ defined in terms of an infobase IB using (Def $-_{IB}$ from IB) is referred to as the *IB -induced theory contraction*. \square

It is easy to verify that the IB -induced faithful total preorder is indeed an $S(IB)$ -faithful total preorder and by theorem 3.2.6 it thus follows that the IB -induced theory contraction is an AGM theory contraction. Associating the IB -induced theory contraction with the infobase IB -contraction allows us to determine which wffs in IB should be retained and which cannot be retained, after a contraction of IB .

Definition 8.2.7 The set of α -discarded wffs (of an infobase IB) is defined as $IB^{-\alpha} = \{\beta \in S(IB) \mid \beta \notin Cn(S(IB)) -_{IB} \alpha\}$. We refer to $S(IB) \setminus IB^{-\alpha}$ as the set of α -retained wffs (of IB). \square

Clearly the α -retained wffs are precisely the wffs in IB that should be retained when contracting IB by α . Unlike the dominant approaches to base contraction discussed in section 8.1, however, we don't simply expunge the α -discarded wffs, but instead opt to replace them with appropriately weakened wffs. (It is only when the weakened version of such a wff is logically valid that we can think of the wff as being completely discarded.) The strategy is to retain as much of the information contained in a wff as possible, even if not all the information in the wff can be retained. This is in line with the intuition that infobases consist of independently obtained wffs. Of course, these weakened wffs cannot be chosen in an arbitrary fashion. Since the IB -induced theory contraction $-_{IB}$ has already been identified as the theory contraction to be associated with the infobase IB -contraction, the weakened wffs, together with the α -retained wffs, have to generate the belief set $Cn(S(IB)) -_{IB} \alpha$.

In deciding on an appropriate method for the weakening of the α -discarded wffs, it is necessary to strike the right balance between a *coherentist* approach, emphasising knowledge level matters, and a *foundationalist* approach, emphasising the independence of the wffs occurring in IB (see page 2). The following example serves to make these matters concrete.

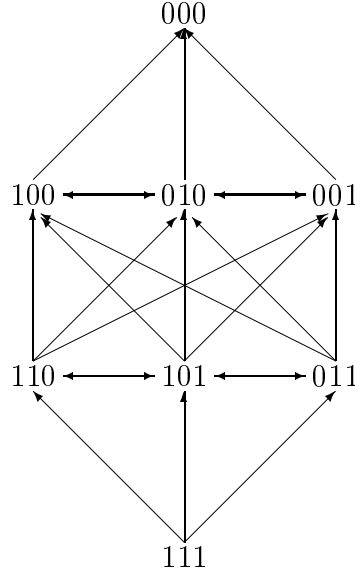


Figure 8.1: A graphical representation of the IB -induced faithful total preorder \preceq_{IB} , with $IB = [p, q, r]$. For every $u, v \in U$, $u \preceq_{IB} v$ iff (u, v) is in the reflexive transitive closure of the relation determined by the arrows.

Example 8.2.8 Let L be the finitely generated propositional language generated by the three atoms p , q , and r , with a valuation semantics (V, \models) , where

$$V = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$

Consider the infobase $IB = [p, q, r]$. Figure 8.1 gives a graphical representation of the IB -induced faithful total preorder \preceq_{IB} . Because p , q and r each represents independently obtained information, a $(p \wedge q)$ -contraction of IB should have no effect on r . That is, when contracting IB by $p \wedge q$, the resulting infobase should contain weakened versions of the two $(p \wedge q)$ -discarded wffs p and q , and should contain the $(p \wedge q)$ -retained wff r itself. But what should the weakened versions of p and q look like?

An application of the coherentist approach on a local level suggests that, in order to minimise the loss of information, one should add only the minimal models of $\neg(p \wedge q)$ to the models of both p and q , and let the corresponding wffs be the appropriate weakened versions. Since $\text{Min}_{\preceq_{IB}}(\neg(p \wedge q)) = \{101, 011\}$, the weakened version of p would be logically equivalent to $p \vee (q \wedge r)$ and the weakened version of q would be logically equivalent to $q \vee (p \wedge r)$.

On the other hand, the foundationalist approach, which stresses the independence of the wffs in IB , suggests that the presence of r should have no effect on the weakened

versions of p and q . In this view, the wff $p \vee q$ (or any wff logically equivalent to it) would be a suitable choice for the weakened versions of both p and q . \square

There does not seem to be a definite answer to the question of which one of these two approaches to infobase change is the “correct” one. They should rather be seen as opposites on a whole spectrum of possibilities. The coherentist approach can be described as the case where *all* the wffs in IB play a role in determining the weakened versions of the α -discarded wffs, while the foundationalist approach ensures that only the set of α -discarded wffs themselves is involved in the construction of their weakened versions. Given these two opposites, it also seems perfectly reasonable to allow for *any* set of wffs in between (i.e., containing the α -discarded wffs and included in $S(IB)$) to be involved in the construction of the weakened versions of the α -discarded wffs.

Definition 8.2.9 Given an infobase IB and a wff α , a set R is said to be (IB, α) -*relevant* iff $IB^{-\alpha} \subseteq R \subseteq S(IB)$. \square

Our goal is to ensure that, in the process of obtaining the weakened versions of the α -discarded wffs, the effect of the wffs *not* in the (IB, α) -relevant set R are neutralised. To do so, we should not just add the \preceq_{IB} -minimal models of $\neg\alpha$, but also any other models of $\neg\alpha$ that behave exactly like the \preceq_{IB} -minimal models with respect to the wffs in R , but that might differ from the \preceq_{IB} -minimal models on the truth value of the wffs in $S(IB) \setminus R$.

Definition 8.2.10 For $X \subseteq L$ and $u, v \in U$, u is X -*equivalent* to v , written $u \equiv_X v$, iff for every $\chi \in X$, $u \in M(\chi)$ iff $v \in M(\chi)$. \square

Observe that, for the $(IB, p \wedge q)$ -relevant set $R = \{p, q\}$ in example 8.2.8, it follows that 100 and 010 are R -equivalent to the minimal models 101 and 011 respectively, and adding them to the models of p (and q) as well, results in weakened versions of p and q that are logically equivalent to $p \vee q$, which is in line with the foundationalist intuition described above.

In general, we obtain the weakened version of every α -discarded wff β as follows. We need some appropriate set of interpretations that can be added to the models of β to obtain the set of models of its weakened version. Once we have decided on an (IB, α) -relevant set R , we use the set of minimal models of $\neg\alpha$ as our starting point and then try to expand it so that only elements in R have any influence, thus neutralising

the possible influence of any of remaining wffs in IB . This is accomplished by including all the models of $\neg\alpha$ that are R -equivalent to some minimal model of $\neg\alpha$.

Definition 8.2.11 Let R be any (IB, α) -relevant set. For every $u \in \text{Min}_{\preceq_{IB}}(\neg\alpha)$, we let $N_u^R(\neg\alpha) = \{v \in M(\neg\alpha) \mid v \equiv_R u\}$, and we let

$$N_{IB}^R(\neg\alpha) = \bigcup_{u \in \text{Min}_{\preceq_{IB}}(\neg\alpha)} N_u^R(\neg\alpha).$$

We refer to $N_{IB}^R(\neg\alpha)$ as the (R, α) -neutralised models of IB . □

We take the (R, α) -neutralised models as the set of interpretations to be added to the models of each α -discarded wff. We can think of the (R, α) -neutralised models as a set of interpretations in which the influence of the wffs not in R has been removed, but in which the wffs in R have the same impact as on the minimal models of $\neg\alpha$.

To summarise, we intend to obtain the infobase resulting from an α -contraction of the infobase IB by weakening the α -discarded wffs in the manner described above, and keeping the α -retained wffs as they are. It turns out that there is an elegant way to provide a uniform description of this process. In doing so, we describe infobase contraction as a process in which *all* the wffs in the current infobase are replaced with weaker versions, but where the “weaker” version of every α -retained wff turns out to be logically equivalent to the wff itself.

Definition 8.2.12 Let R be any (IB, α) -relevant set. For every $\beta \in S(IB)$, we let

$$N_\beta^R(\neg\alpha) = \bigcup_{u \in \text{Min}_{\preceq_{IB}}(\neg\alpha) \setminus M(\beta)} N_u^R(\neg\alpha).$$

We refer to $N_\beta^R(\neg\alpha)$ as the (R, α, β) -neutralised models of IB . □

The next proposition shows that an α -retained wff β has no (R, α, β) -neutralised models, and that, for an α -discarded wff β , adding the (R, α, β) -neutralised models to the models of β has the same effect as adding the (R, α) -neutralised models.

Proposition 8.2.13 Let R be any (IB, α) -relevant set.

1. If $\beta \in S(IB) \setminus IB^{-\alpha}$ then $N_\beta^R(\neg\alpha) = \emptyset$.
2. If $\beta \in IB^{-\alpha}$ then $M(\beta) \cup N_\beta^R(\neg\alpha) = M(\beta) \cup N_{IB}^R(\neg\alpha)$.

Proof 1. Suppose that $\beta \in S(IB) \setminus IB^{-\alpha}$. Then $\beta \in Cn(S(IB)) -_{IB} \alpha$ and thus $Min_{\preceq_{IB}}(\neg\alpha) \subseteq M(\beta)$. And therefore

$$N_{\beta}^R(\neg\alpha) = \bigcup_{u \in Min_{\preceq_{IB}}(\neg\alpha) \setminus M(\beta)} N_u^R(\neg\alpha) = \emptyset.$$

2. Suppose that $\beta \in IB^{-\alpha}$. The left-to-right inclusion is immediate. For the right-to-left inclusion we have to show that

$$\bigcup_{u \in Min_{\preceq_{IB}}(\neg\alpha) \cap M(\beta)} N_u^R(\neg\alpha) \subseteq M(\beta).$$

So pick any $u \in Min_{\preceq_{IB}}(\neg\alpha) \cap M(\beta)$ and $v \in N_u^R(\neg\alpha)$. Then $v \equiv_R u$ and since $\beta \in R$, it follows that $v \in M(\beta)$. □

Proposition 8.2.13 allows us to describe an α -contraction of an infobase IB by adding to the models of a wff β in IB , the set $N_{\beta}^R(\neg\alpha)$, and replacing β with an axiomatisation of this set of interpretations. Of course, such a description only makes sense if these sets of interpretations can be axiomatised by single wffs. While this is immediate for the finitely generated propositional logics, the next result shows that it also holds in the more general case.

Definition 8.2.14 Let R be any (IB, α) -relevant set, and for $\beta \in S(IB)$, let IB_{β}^{α} be the set containing every ordered subsequence C of IB such that $|C| = u_{IB}$ for some $u \in (Min_{\preceq_{IB}}(\neg\alpha) \cap M(S(C))) \setminus M(\beta)$ (where u_{IB} is the IB -number of u). We define the α -weakened version of β , with respect to R , as

$$w_{(IB, \alpha)}^R(\beta) = \beta \vee \left(\bigvee_{C \in IB_{\beta}^{\alpha}} \left(\left(\bigwedge (S(C) \setminus (S(IB) \setminus R)) \right) \wedge \left(\bigwedge \neg (R \setminus S(C)) \right) \wedge \neg\alpha \right) \right)$$

□

Proposition 8.2.15 Let R be an (IB, α) -relevant set. For every $\alpha \in L$ and every $\beta \in S(IB)$, $M(w_{(IB, \alpha)}^R(\beta)) = M(\beta) \cup N_{\beta}^R(\neg\alpha)$.

Proof Define IB_{β}^{α} as in definition 8.2.14. If $IB_{\beta}^{\alpha} = \emptyset$ then it follows easily that $Min_{\preceq_{IB}}(\neg\alpha) \setminus M(\beta) = \emptyset$, which means that $Min_{\preceq_{IB}}(\neg\alpha) \subseteq M(\beta)$ and therefore

that $N_\beta^R(\neg\alpha) = \emptyset$. So we only need to consider the case where $IB_\beta^\alpha \neq \emptyset$. Then every $u \in \text{Min}_{\preceq_{IB}}(\neg\alpha) \setminus M(\beta)$ is a model of $S(C)$ for some $C \in IB_\beta^\alpha$. Pick any $C \in IB_\beta^\alpha$ and any $u \in (\text{Min}_{\preceq_{IB}}(\neg\alpha) \cap M(S(C))) \setminus M(\beta)$. Observe that every model of $S(C) \cup \{\neg\alpha\}$ is a \preceq_{IB} -minimal element of $M(\neg\alpha)$, which ensures that every element of $(R \setminus S(C)) \setminus \text{Cn}(\top)$ is false in all the models of $S(C) \cup \{\neg\alpha\}$. We record this result formally.

$$\forall \gamma \in (R \setminus S(C)) \setminus \text{Cn}(\top), \forall v \in M(S(C) \cup \{\neg\alpha\}), v \notin M(\gamma) \quad (8.1)$$

We show that $M((S(C) \setminus (S(IB) \setminus R)) \cup \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \cup \{\neg\alpha\}) = N_u^R(\neg\alpha)$. From (8.1) it follows that $u \notin M(\gamma)$ for every $\gamma \in (R \setminus S(C)) \setminus \text{Cn}(\top)$ and therefore that

$$u \in M((S(C) \setminus (S(IB) \setminus R)) \cup \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \cup \{\neg\alpha\}).$$

Now pick any $v \in M((S(C) \setminus (S(IB) \setminus R)) \cup \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \cup \{\neg\alpha\})$ and any $\rho \in R$. We only consider the case where $\rho \not\equiv \top$. If $\rho \in S(C)$ then clearly $u \in M(\rho)$ iff $v \in M(\rho)$, so suppose $\rho \notin S(C)$. Then by (8.1) again, $u \notin M(\rho)$. Furthermore, since $v \in M(\neg((R \setminus S(C)) \setminus \text{Cn}(\top)))$, it follows that $v \notin M(\rho)$ and thus that $u \in M(\rho)$ iff $v \in M(\rho)$. Finally, it is clear that $v \in M(\neg\alpha)$. We have thus shown that $v \in N_u^R(\neg\alpha)$. Conversely, pick any $v \in N_u^R(\neg\alpha)$. Clearly $v \in M(\neg\alpha)$, and since $u \in M((S(C) \setminus (S(IB) \setminus R)) \cup \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \cup \{\neg\alpha\})$, so is v .

It is clear that $M((S(C) \setminus (S(IB) \setminus R)) \cup \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \cup \{\neg\alpha\})$ is axiomatised by the wff

$$(\neg\alpha)_C^R = \left(\bigwedge (S(C) \setminus (S(IB) \setminus R)) \right) \wedge \left(\bigwedge \neg((R \setminus S(C)) \setminus \text{Cn}(\top)) \right) \wedge \neg\alpha$$

and it thus follows that $M((\neg\alpha)_C^R) = N_u^R(\neg\alpha)$. So we have shown that if $IB_\beta^\alpha \neq \emptyset$, then

$$\forall C \in IB_\beta^\alpha, \exists u \in (\text{Min}_{\preceq_{IB}}(\neg\alpha) \cap M(S(C))) \setminus M(\beta) \text{ and} \quad (8.2)$$

$$\forall C \in IB_\beta^\alpha, \forall u \in (\text{Min}_{\preceq_{IB}}(\neg\alpha) \cap M(S(C))) \setminus M(\beta),$$

$$M((\neg\alpha)_C^R) = N_u^R(\neg\alpha). \quad (8.3)$$

We now show that $N_\beta^R(\alpha) = M\left(\bigvee_{C \in IB_\beta^\alpha} (\neg\alpha)_C^R\right)$, from which the required result follows. Pick a $v \in N_\beta^R(\alpha)$. There is a $u \in \text{Min}_{\preceq_{IB}}(\neg\alpha) \setminus M(\beta)$ such that $v \in N_u^R(\neg\alpha)$, and by (8.3) it follows that for some $C \in IB_\beta^\alpha$, $v \in N_u^R(\neg\alpha) = M((\neg\alpha)_C^R)$. So clearly $v \in M\left(\bigvee_{C \in IB_\beta^\alpha} (\neg\alpha)_C^R\right)$. Conversely, pick any $v \in M\left(\bigvee_{C \in IB_\beta^\alpha} (\neg\alpha)_C^R\right)$. Then v is a

model of $(\neg\alpha)_C^R$ for some $C \in IB_\beta^\alpha$. By (8.2) there is a $u \in (Min_{\preceq_{IB}}(\neg\alpha) \cap M(S(C))) \setminus M(\beta)$, and by (8.3), $N_u^R(\neg\alpha) = M((\neg\alpha)_C^R)$. So $v \in N_u^R(\neg\alpha)$ and thus $v \in N_\beta^R(\neg\alpha)$. \square

We are now almost in a position to define *basic* infobase contraction.

Definition 8.2.16 A function $rs : \mathcal{IB} \times \wp L \rightarrow \wp \wp L$ is a *relevance selection function* iff

1. $IB^{-\alpha} \subseteq rs(IB, \alpha) \subseteq IB$,
2. if $\alpha \equiv \beta$ then $rs(IB, \alpha) = rs(IB, \beta)$, and
3. if $IB \approx IC$ (that is, IB and IC are element-equivalent) then $rs(IB, \alpha) = rs(IC, \alpha)$.

\square

Intuitively, a relevance selection function indicates which of the wffs in IB should play a role in determining the weakened versions during a contraction. Observe that $rs(IB, \alpha)$ is (IB, α) -relevant.

- Definition 8.2.17**
1. An infobase change operation \ominus is a *basic infobase contraction* iff there is a relevance selection function rs such that, for every $IB \in \mathcal{IB}$ and every $\alpha \in L$, $IB \ominus \alpha$ is obtained by replacing every wff β in IB with $w_{(IB, \alpha)}^{rs(IB, \alpha)}(\beta)$, the α -weakened version of β with respect to $rs(IB, \alpha)$.
 2. For every $IB \in \mathcal{IB}$, an infobase IB -change operation \ominus_{IB} is a *basic infobase IB -contraction* iff it can be obtained from an infobase contraction \ominus by fixing the infobase IB . That is, iff $IB \ominus_{IB} \alpha = IB \ominus \alpha$ for every $\alpha \in L$.

\square

We conclude this section with an example illustrating the partial construction of some basic infobase contractions.

Example 8.2.18 Let $IB = [p, q]$. Figure 8.2 contains a graphical representation of the IB -induced faithful total preorder \preceq_{IB} . Then

$$\begin{aligned}
 Cn(S(IB)) -_{IB} p &= Cn(q), IB^{-p} = \{p\} \\
 IB_p^p &= \{[q]\}, IB_q^p = \emptyset, \\
 Cn(S(IB)) -_{IB} (p \wedge q) &= Cn(p \vee q) \\
 IB^{-(p \wedge q)} &= \{p, q\}, IB_p^{p \wedge q} = \{[q]\}, \text{ and } IB_q^{p \wedge q} = \{[p]\}.
 \end{aligned}$$

Now observe that $w_{(IB,p)}^{IB^{-p}}(p) = p \vee (\top \wedge \neg p \wedge \neg p) \equiv \top$ and that $w_{(IB,p)}^{IB^{-p}}(q) = q \vee \perp \equiv q$. Furthermore, since $S(IB) = B^{-p \wedge q}$, note that

$$w_{(IB,p \wedge q)}^{S(IB)}(p) = w_{(IB,p \wedge q)}^{IB^{-p \wedge q}}(p) = p \vee (q \wedge \neg p \wedge \neg(p \wedge q)) \text{ and}$$

$$w_{(IB,p \wedge q)}^{S(IB)}(q) = w_{(IB,p \wedge q)}^{IB^{-p \wedge q}}(q) = q \vee (p \wedge \neg q \wedge \neg(p \wedge q)).$$

It can be verified that both $w_{(IB,p \wedge q)}^{S(IB)}(p)$ and $w_{(IB,p \wedge q)}^{S(IB)}(q)$ are logically equivalent to $p \vee q$. There is thus at least one infobase contraction \ominus such that

$$IB \ominus p = [w_{(IB,p)}^{IB^{-p}}(p), w_{(IB,p)}^{IB^{-p}}(q)] \approx [\top, q]$$

and

$$IB \ominus (p \wedge q) = [w_{(IB,p \wedge q)}^{IB^{-p \wedge q}}(p), w_{(IB,p \wedge q)}^{IB^{-p \wedge q}}(q)] \approx [p \vee q, p \vee q].$$

Furthermore, observe that $w_{(IB,p)}^{S(IB)}(p) = p \vee (q \wedge \neg p \wedge \neg p) \equiv p \vee q$ and that $w_{(IB,p)}^{S(IB)}(q) = q \vee \perp \equiv q$. So there is least one infobase contraction \ominus' such that

$$IB \ominus' p = [w_{(IB,p)}^{S(IB)}(p), w_{(IB,p)}^{S(IB)}(q)] \approx [p \vee q, q]$$

and

$$IB \ominus' (p \wedge q) = [w_{(IB,p \wedge q)}^{S(IB)}(p), w_{(IB,p \wedge q)}^{S(IB)}(q)] \approx [p \vee q, p \vee q].$$

□

8.2.2 Properties of basic infobase contraction

In the discussion of infobase contraction thus far, it has been implied that the α -weakened versions of the α -discarded wffs are appropriate choices for weakened versions of these wffs, and that the IB -induced theory contraction is the theory contraction associated with every basic infobase IB -contraction. The first point has already been dealt with in the previous section. For the second point, we first present a preliminary result, indicating that for every (IB, α) -relevant set R , the models of the α -retained wffs that are also (R, α) -neutralised models, are precisely the \preceq_{IB} -minimal models of $\neg\alpha$.

Lemma 8.2.19 *If R is an (IB, α) -relevant set, then*

$$N_{IB}^R(\neg\alpha) \cap M(S(IB) \setminus IB^{-\alpha}) = \text{Min}_{\preceq_{IB}}(\neg\alpha).$$

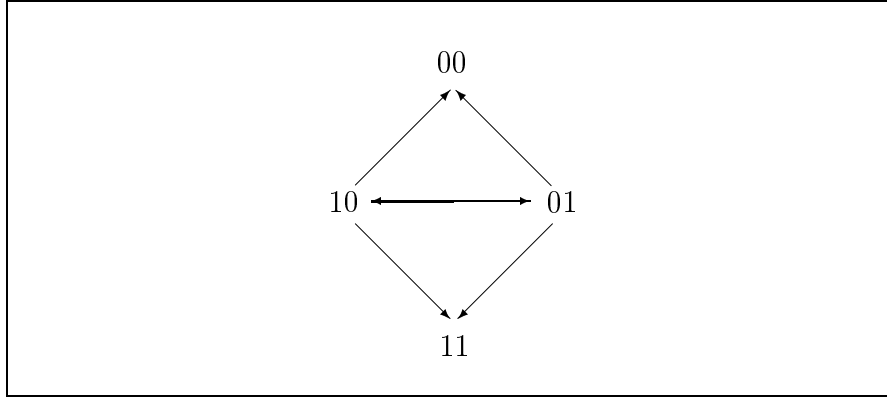


Figure 8.2: A graphical representation of the IB -induced faithful total preorder \preceq_{IB} , with $IB = [p, q]$. For every $u, v \in U$, $u \preceq_{IB} v$ iff (u, v) is in the reflexive transitive closure of the relation determined by the arrows.

Proof By definition, $S(IB) \setminus IB^{-\alpha} \subseteq Cn(S(IB)) -_{IB} \alpha$ and thus

$$M(S(IB)) \cup Min_{\preceq_{IB}}(\neg\alpha) \subseteq M(S(IB) \setminus IB^{-\alpha}).$$

Furthermore, $Min_{\preceq_{IB}}(\neg\alpha) \subseteq N_{IB}^R(\neg\alpha)$, and so $Min_{\preceq_{IB}}(\neg\alpha) \subseteq N_{IB}^R(\neg\alpha) \cap M(S(IB) \setminus IB^{-\alpha})$. Conversely, pick any $v \in N_{IB}^R(\neg\alpha) \cap M(S(IB) \setminus IB^{-\alpha})$. That is, v satisfies all the α -retained wffs, v is a model of $\neg\alpha$ and there is a \preceq_{IB} -minimal model u of $\neg\alpha$ that satisfies exactly the same wffs in R as v does (which includes the α -discarded wffs). Because $u \in Min_{\preceq_{IB}}(\neg\alpha)$, it follows from the definition of $-_{IB}$ and $IB^{-\alpha}$ that u also satisfies all the wffs in $S(IB) \setminus IB^{-\alpha}$. So u and v satisfy exactly the same wffs occurring in IB , which means that $v \in Min_{\preceq_{IB}}(\neg\alpha)$. \square

The result above is used to prove that the IB -induced contraction $-_{IB}$ is the theory contraction associated with every basic infobase IB -contraction.

Proposition 8.2.20 *Let \ominus be any basic infobase contraction. Then*

$$Cn(S(IB)) -_{IB} \alpha = Cn(S(IB \ominus \alpha)).$$

Proof Let rs be the relevance selection function used to define \ominus . By propositions 8.2.13 and 8.2.15,

$$M(S(IB \ominus \alpha)) = \left[\bigcap_{\beta \in IB^{-\alpha}} \left(M(\beta) \cup N_{IB}^{rs(IB, \alpha)}(\neg\alpha) \right) \right] \cap M(S(IB) \setminus IB^{-\alpha})$$

$$\begin{aligned}
&= \left[\left(\bigcap_{\beta \in IB^{-\alpha}} M(\beta) \right) \cup N_{IB}^{rs(IB, \alpha)}(\neg \alpha) \right] \cap M(S(IB) \setminus IB^{-\alpha}) \\
&= \left(M(IB^{-\alpha}) \cup N_{IB}^{rs(IB, \alpha)}(\neg \alpha) \right) \cap M(S(IB) \setminus IB^{-\alpha}) \\
&= M(S(IB)) \cup \left(N_{IB}^{rs(IB, \alpha)}(\neg \alpha) \cap M(S(IB) \setminus IB^{-\alpha}) \right) \\
&= M(S(IB)) \cup \text{Min}_{\preceq_{IB}}(\neg \alpha) \text{ by lemma 8.2.19,}
\end{aligned}$$

and thus $Cn(S(IB)) -_{IB} \alpha = Cn(S(IB \ominus \alpha))$. \square

Since one of the basic tenets of infobase change is that knowledge level issues matter, one would not expect syntax to play too big a role in the construction of infobase change operations. We show that the syntactic form of the wffs in an infobase, as well as form of the wff with which to contract, are irrelevant.

Proposition 8.2.21 *Let \ominus be a basic infobase contraction, and suppose that $IB \approx IC$ and $\beta \equiv \gamma$. Then $IB \ominus \beta \approx IC \ominus \gamma$.*

Proof Let rs be the relevance selection function used to obtain \ominus . Since IB and IC are element-equivalent, $u_{IB} = u_{IC}$ for every $u \in U$, and so the IB -induced faithful total preorder is exactly the same as the IC -induced faithful preorder. By the properties of a relevance selection function, it then follows that $N_{IB}^{rs(IB, \beta)}(\neg \beta) = N_{IC}^{rs(IC, \gamma)}(\neg \gamma)$. So, by propositions 8.2.13 and 8.2.15, $w_{(IB, \beta)}^{IB}(\beta') \equiv w_{(IC, \gamma)}^{IC}(\gamma')$ for every β' in IB and every γ' in IC such that $\beta' \equiv \gamma'$, from which the required result follows. \square

8.2.3 Infobase contraction and reason maintenance

In section 8.1 it was pointed out that base change came about as an attempt to perform reason maintenance, the process in which the removal of a basic belief forces the removal of the consequences of the basic belief as well, unless the latter wffs can be derived from other basic beliefs. In the context of infobase change, the wffs in an infobase IB are viewed as such basic beliefs of the belief set associated with IB . Reason maintenance would thus ensure that the contraction of IB by a wff α in IB results in the removal of all the wffs that are dependent on α for being in $Cn(S(IB))$. Fuhrmann [1991] has given a precise meaning to the idea of a wff being dependent on α (for being in $Cn(S(IB))$).⁴

⁴Fuhrmann works with belief bases and not infobases, and our definition of IB -dependence is thus a slight generalisation of the notion he defines.

Definition 8.2.22 A wff $\beta \in L$ is *IB-dependent on α* iff $\alpha \in S(IB)$ and $\beta \in Cn(S(IB))$, but $\beta \notin Cn(S(IB) \setminus \{\alpha\})$. \square

The next result shows that basic infobase contraction incorporates reason maintenance.

Proposition 8.2.23 *Let \ominus be a basic infobase contraction. If β is IB-dependent on α then $\beta \notin Cn(S(IB \ominus \alpha))$.*

Proof Since $\beta \in Cn(S(IB))$, but $\beta \notin Cn(S(IB) \setminus \{\alpha\})$, there has to be a model u of $S(IB) \setminus \{\alpha\}$ in which both α and β are false. So $u \in M(\neg\alpha)$ and $u \notin M(S(IB))$. Now, there is only one wff in IB , namely α , that is false in u (although IB may contain multiple instances of α). So any interpretation v for which $v_{IB} > u_{IB}$, has to be a model of $S(IB)$ and hence of α . Therefore $u \in Min_{\preceq_{IB}}(\neg\alpha)$, and because $u \notin M(\beta)$, it follows that $\beta \notin Cn(S(IB)) -_{IB} \alpha$. So $\beta \notin Cn(S(IB \ominus \alpha))$ by proposition 8.2.20. \square

Of course, the contraction of IB by a wff α in IB is not the only way to remove α from the infobase IB . In the light of this, it seems reasonable to inquire whether the wffs that are *IB-dependent on α* will also be discarded if α is discarded during the contraction of IB by some wff other than α itself. That is, if α is in IB and $\alpha \notin Cn(S(IB \ominus \gamma))$, will it be the case that $\beta \notin Cn(S(IB \ominus \gamma))$ for every β that is *IB-dependent on α* ? This property is known as Fuhrmann's [1991] *filtering condition*. It is easy to see that basic infobase contraction can violate the filtering condition. For example, it is readily verified that for *any* basic infobase contraction, the contraction of the infobase $IB = [p \wedge q]$ by p results in an infobase in which $p \wedge q$ is replaced by the wff $w_{(IB,p)}^{S(IB)}(p \wedge q)$ which is logically equivalent to $p \rightarrow q$. And since $w_{(IB,p)}^{S(IB)}(p \wedge q)$ is clearly *IB-dependent on $p \wedge q$* , the filtering condition is violated. But such a violation is to be expected. Given the intuition associated with infobases, the filtering condition is clearly too strong a requirement to impose. For the filtering condition requires that for any infobase contraction \ominus , $Cn(S(IB \ominus \gamma)) = Cn(\top)$ for any singleton infobase IB , and any $\gamma \in Cn(S(IB))$ (where $\neq \gamma$), thus leaving no room for weakening the wff in IB to anything but a logically valid wff.

8.2.4 Infobase revision

Basic infobase revision is defined by an appeal to the following infobase analogue of the Levi Identity:

(Def \otimes from \ominus) $IB \otimes \alpha = (IB \ominus \neg\alpha) \bullet \alpha$

Definition 8.2.24 An infobase change operation \otimes is a *basic infobase revision* iff it can be defined in terms of a basic infobase contraction \ominus using (Def \otimes from \ominus). \square

Given this connection, it is to be expected that basic infobase revision satisfies properties that are similar to those proved in sections 8.2.2 and 8.2.3. The next corollary shows that this is indeed the case.

Definition 8.2.25 A theory revision $*$ is *associated with* an infobase IB -revision \otimes (for some infobase IB) iff $Cn(IB) * \alpha = Cn(IB \otimes \alpha)$ for every $\alpha \in L$. \square

(Def $*_{IB}$ from IB) $Cn(S(B)) *_{IB} \alpha = Th(Min_{\preceq_{IB}}(\alpha))$

Definition 8.2.26 The theory revision $*_{IB}$ defined in terms of an infobase IB using (Def $*_{IB}$ from IB) is referred to as the *IB -induced theory revision*. \square

From theorem 3.2.6 it follows that the IB -induced theory revision is an AGM theory revision.

Corollary 8.2.27 Let \ominus be a basic infobase contraction, and let \otimes be the infobase revision defined in terms of \ominus using (Def \otimes from \ominus).

1. If $IB \approx IC$ and $\alpha \equiv \beta$ then $IB \otimes \alpha \approx IC \otimes \beta$.
2. $Cn(S(IB \otimes \alpha)) = Cn(S(IB)) *_{IB} \alpha$.
3. If β is IB -dependent on α , then $\beta \notin Cn(S(B \otimes \neg\alpha))$.

Proof 1. Follows from proposition 8.2.21.

2. Follows from proposition 8.2.20, by noting that $Min_{\preceq_{IB}}(\alpha) \subseteq M(S(IB))$ if $\neg\alpha \notin Cn(S(IB))$, and by recalling that $Cn(S(IB)) *_{IB} \alpha = Th(Min_{\preceq_{IB}}(\alpha))$.

3. Follows from part (2) of this corollary, and by an argument similar to the proof of proposition 8.2.23.

\square

Part (1) of corollary 8.2.27 shows that basic infobase revision is insensitive to the syntactic form of the wffs in an infobase, as well as to the syntactic form of the wff with which to revise, part (2) shows that the theory revision associated with a basic infobase revision is the IB -induced revision function, and part (3) shows that basic infobase revision can be said to perform reason maintenance.

It is also possible to provide a result for infobase change which is reminiscent of the Harper Identity (the identity (Def – from $*$)).

Proposition 8.2.28 *Let \otimes be a basic infobase revision, and let \ominus be an infobase change operation such that $IB \ominus \alpha \approx \overline{IB \otimes \neg \alpha}$. Then \ominus is a basic infobase contraction.*

Proof Follows from the fact that there is a basic infobase contraction \ominus' such that $IB \otimes \neg \alpha = (IB \ominus' \neg \neg \alpha) \bullet \neg \alpha$ and that $\alpha \equiv \neg \neg \alpha$. \square

To conclude this section, we provide an example to show that infobase change is able to accommodate Hansson's hamburger example (example 8.1.1) in an appropriate fashion.

Example 8.2.29 Let L be the propositional language generated by the two atoms p and q with a valuation semantics (V, \models) , where $V = \{00, 01, 10, 11\}$. We let p denote the assertion that the restaurant whose lights are on is open, and we let q denote the assertion that the second restaurant is open. Now, let $IB = [p, p \vee q]$ and let $IC = [p]$. Since $IB^{\neg \neg p} = \{p\}$, it follows from propositions 8.2.13 and 8.2.15 that for every basic infobase revision \otimes , there is a β in $IB \otimes \neg p$ such that $\beta \equiv p \vee q$. Furthermore, since $IC^{\neg \neg p} = IC$, it follows that for every basic infobase revision \otimes , $IC \otimes \neg p \approx [\top, \neg p] \approx [\neg p]$. As our intuition suggests, revising IB by $\neg p$ yields an infobase containing $p \vee q$ (or something locally equivalent to it). In contrast, a revision of IC by $\neg p$ does not contain such a wff. Nor, for that matter, does $p \vee q$ follow logically from the infobase resulting from a $\neg p$ -revision of IC . \square

8.3 Related approaches

Infobase change relies heavily on the IB -induced faithful total preorders, which are obtained by counting the number of wffs in an infobase IB . As such, its roots can be found in the work of Dalal [1988], Borgida [1985], Satoh [1988], Weber [1986], Winslett [1988], all of whom use the idea of distinguishing between interpretations based on the number of propositional atoms that they satisfy (at least in the propositional

case). However, these approaches do not distinguish between different infobases (or belief bases) generating the same belief set, and are thus more properly classified as instances of theory change for the same reasons that Katsuno and Mendelzon's work is seen as research on theory change, rather than research on base change (see section 8.1 page 243).

As discussed in section 8.1, base contraction is usually associated with the requirement that the belief base resulting from a base contraction ought to be a subset of the original belief base. Two notable exceptions to this are the base contraction operations of Nebel [1989, 1990, 1991, 1992] and Nayak [1994a], which allow wffs into the resulting belief base that were not in the original belief base. In this section we compare these two approaches with infobase change.

8.3.1 Nebel's approach

Nebel's base change operations in [Nebel, 1990, 1991, 1992] make use of an *epistemic relevance* ordering on the wffs in the belief set generated by the base, which is taken to denote relative epistemic importance. This is a generalisation of the case considered in [Nebel, 1989], which can be seen as the special case where all wffs in the base have equal epistemic weight. Since the latter is closer to infobase change, we shall mainly concern ourselves with the work in [Nebel, 1989].

Nebel's construction of base contraction functions uses the maximal subsets of a set X that do not entail α . It can thus be seen as a generalisation of the construction of the partial meet functions (see section 2.2). For every $X \subseteq L$, let $X \downarrow \alpha$, the set of *remainders of X after removing α* , be defined as

$$X \downarrow \alpha = \{Y \subseteq X \mid Y \not\models \alpha \text{ and for every } Z \subseteq L \text{ such that } Y \subseteq Z \subset X, Z \models \alpha\}.$$

Nebel defines the base contraction $\hat{\sim}$, in a somewhat opaque fashion, as

$$B \hat{\sim} \alpha = \begin{cases} \bigvee_{C \in (B \downarrow \alpha)} C \wedge (B \vee \{\neg \alpha\}) & \text{if } B \not\models \alpha, \\ B, & \text{otherwise.} \end{cases}$$

This construction is justified by a closer look at the theory contraction associated with $\hat{\sim}$. He defines a B -faithful weak partial order \leq as: $x \leq y$ iff $(Th(x) \cap B) \supseteq (Th(y) \cap B)$,

and then obtains a $Cn(B)$ -contraction $\hat{-}$ from \leq as follows:⁵

$$Cn(B)\hat{-}\alpha = Th(M(B) \cup Min_{\leq}(-\alpha)).$$

He then proceeds to show that $\hat{-}$ is the $Cn(B)$ -contraction associated with $\hat{\sim}$ (i.e. $Cn(B)\hat{-}\alpha = Cn(B\hat{\sim}\alpha)$), and that $\hat{-}$ satisfies (K-1) to (K-7), but does not, in general, satisfy (K-8).

A comparison of Nebel's $Cn(B)$ -contraction $\hat{-}$ (which is obtained from \leq) with the IB -induced contraction (where B is a belief base and IB an infobase) shows that the intuitions employed in both cases are very similar. But whereas \leq is defined in terms of the satisfaction of maximal *subsets* of B , the IB -induced faithful total preorder relies on the satisfaction of the maximum *number* of wffs in IB . While this difference allows for Nebel's $\hat{-}$ to be defined for infinite bases as well, it ensures that $\hat{-}$ does not always satisfy (K-8), while the IB -induced contraction does. Below we provide an example in which it seems desirable for a base contraction operation to satisfy (K-8), at least under the assumption of the independence of the wffs in a belief base B .

Example 8.3.1 Let $B = \{p \vee q, \neg p \vee q, p\}$ and let \sim be a base contraction in which the wffs in B are regarded as being independently obtained. A contraction with $p \wedge q$ would force us to remove at least one of p and q from $Cn(B)$, and since $p \in B$ but $q \notin B$, it seems reasonable to require that if one of the two is retained, it should be p and not q . So, regardless of whether p is being retained, q should not be an element of $Cn(B \sim (p \wedge q))$. Furthermore, since $p \vee q$ is explicitly contained in B , a contraction of B by $p \wedge q$ should not remove $p \vee q$, and we should thus have $p \vee q \in Cn(B \sim (p \wedge q))$. Finally, although the presence of both $p \vee q$ and $\neg p \vee q$ in B suggests that p and q are independent (since $p \vee q$ is logically equivalent to $\neg p \rightarrow q$, and $\neg p \vee q$ to $p \rightarrow q$), this is, to some extent, offset by the presence in B of both p and $\neg p \vee q$. The inconclusive evidence regarding the independence of p and q , coupled with the fact that p itself is in B , then suggests that p should be an element of $Cn(B \sim q)$. It is easy to see that the failure of the intuition expressed above would amount to a violation of (K-8). By taking α as p and β as q , it is easily seen that Nebel's $Cn(B)$ -contraction function $\hat{-}$ violates (K-8) ($p \vee q \in Cn(B)\hat{-}(p \wedge q)$, but $p \vee q \notin Cn(B)\hat{-}q$). \square

Nebel also considers a modification of $\hat{-}$ that satisfies (K-8) (which allows him to set $B\hat{\sim}\alpha$ equal to some element of $B \downarrow \alpha$) but it presupposes a linear order on the wffs

⁵Nebel's construction of the theory contraction function $\hat{-}$ is phrased in terms of partial meet functions, but it is easily seen that it can also be phrased semantically, as we have done.

in B , which is a very strong restriction indeed. The restriction is relaxed to a total preorder in [Nebel, 1990, 1991, 1992], but then (K-8) does not hold in the general case.

We have thus far considered the $Cn(B)$ -contraction $\hat{-}$ in detail, but have said very little about $\hat{\sim}$ itself. From some comments made in his conclusion, it seems that Nebel regards the set of wffs $B\hat{\sim}\alpha$ merely as a convenient finite representation from which the belief set $B\hat{-}\alpha$ can be generated, and nothing more. He writes: "...iterated contractions were ignored because they present serious problems.", and "Choosing the 'right' form of the premises seems to be one of the central tasks before any kind of belief revision can be applied". The latter statement seems to suggest that $B\hat{\sim}\alpha$ cannot be seen as a base with the wffs contained in it being epistemologically more important than the wffs in $Cn(B\hat{\sim}\alpha)$, a view that is also supported by his proposal for a base revision $\hat{*}$. He defines $B\hat{*}\alpha$ as $(B\hat{\sim}\neg\alpha) \wedge \{\alpha\}$, which means that the newly obtained basic belief α occurs in $B\hat{*}\alpha$ as a conjunct of every wff in $(B\hat{\sim}\neg\alpha)$. And there certainly is no intuition of a weakening of the wffs contained in B , as with infobase change. For example, if $B = \{p, q, r\}$, it can be verified that $B\hat{\sim}(p \wedge q \wedge r)$ contains 24 elements and is element-equivalent to the infobase $[p \vee q, p \vee r, q \vee r, p \vee q \vee r]$. In contrast, consider the infobase contraction \ominus obtained from the relevance selection function sr where $sr(IB, \alpha) = IB^{-\alpha}$ for every $IB \in \mathcal{IB}$, and every $\alpha \in L$. It can be verified that, for the infobase $IB = [p, q, r]$, $IB \ominus (p \wedge q \wedge r)$ contains three logically non-equivalent wffs (weakened versions of each of the wffs in IB) and is element-equivalent to the infobase $[p \vee (q \wedge r), q \vee (p \wedge r), r \vee (p \wedge q)]$.

8.3.2 Nayak's approach

In some ways, Nayak's [1994a] approach to base change is more general than infobase change since it accommodates infinite bases. (On the other hand, of course, infobases have a richer structure than finite sets of wffs.) He takes Fuhrmann's [1991] generalised safe contraction as a starting point. When contracting a base B by α (a base contraction which we denote by $\tilde{-}$) he first finds the set $E(\alpha)$ of minimal subsets of B that entail α . The idea is to construct a *reject set* $R(\alpha)$ (wffs of B that will be discarded), consisting of wffs from every element of $E(\alpha)$. To ensure that the $Cn(B)$ -contraction associated with $\tilde{-}$ satisfies (K-1) to (K-5), he assumes a *choice function* C from $\wp B$ to $\wp B$ that picks the "most rejectable" elements of any subset of B . Up to this

point the construction corresponds roughly to Fuhrmann's base contraction. However, Fuhrmann's version of the choice function does not have to conform to the stringent restrictions that Nayak places on C . Furthermore, Nayak does not take the set $R_0(\alpha)$, which consists of the most rejectable elements of all members of $E(\alpha)$, to be the reject set, as Fuhrmann does. Instead, he uses C to choose a particular subset of $R_0(\alpha)$, which also happens to be an element of $B \downarrow \alpha$, as the reject set $R(\alpha)$. $B \tilde{\sim} \alpha$ is then defined as the wffs in B that are not rejected, together with weakened versions of the rejected wffs. To be precise, $B \tilde{\sim} \alpha = B \setminus R(\alpha) \cup \{\beta \rightarrow \alpha \mid \beta \in R(\alpha)\}$. Nayak proves that the $Cn(B)$ -contraction function $\tilde{-}$ associated with $\tilde{\sim}$ satisfies all eight AGM contraction postulates. The addition of the weakened versions of wffs in the reject set ensures that $\tilde{-}$ satisfies (K-6), but it is currently unclear whether it plays a role in the satisfaction of (K-7) and (K-8) as well.

The strict conditions imposed on C , together with the insistence that the reject set $R(\alpha)$ be an element of $B \downarrow \alpha$, are akin to placing a linear order on B . This means that Nayak's base contraction function $\tilde{\sim}$ is closely related to Nebel's modified version of the base contraction function $\hat{\sim}$, for which $B \hat{\sim} \alpha$ is an element of $B \downarrow \alpha$. It is thus difficult to draw a direct comparison between $\tilde{\sim}$ and infobase contraction, mainly because the construction of $\tilde{\sim}$ needs so much more extra-logical information. A feature that Nayak's base contraction does have in common with infobase contraction concerns the wffs contained in the resulting base (or infobase) after a contraction has taken place. Both retain a number of wffs and replace the wffs that are removed with weakened versions. Currently, the closest we can come to a comparison is to give an example showing that any reasonable modification to $\tilde{\sim}$ which caters for situations in which less extra-logical information is available will probably not always give the desired results, at least not when the wffs in a base are assumed to be independent. This does not, of course, suggest that infobase contraction will always be preferable to such modified versions of Nayak's approach. It merely serves to indicate that, given the assumption of the independence of wffs, there are cases in which infobase contraction is preferable to any modification that retains the spirit of Nayak's original approach.

Example 8.3.2 Let $B = \{p, q\}$. The requirement that the reject set be a subset of B seems to form an integral part of Nayak's approach, which means that the reject set $R(p \wedge q)$ has to be one of \emptyset , $\{p\}$ or $\{q\}$, irrespective of any restrictions on the choice function C . The only candidates for $B \tilde{\sim} (p \wedge q)$ are thus $\{(p \wedge q) \rightarrow p, (p \wedge q) \rightarrow q\}$, $\{p, (p \wedge q) \rightarrow q\}$ and $\{q, (p \wedge q) \rightarrow p\}$. Now, if p and q have equal weight then the

desired result when contracting B with $p \wedge q$ is $\{p \vee q\}$ (or some set containing elements that are logically equivalent to $p \vee q$), a set of wffs which Nayak's approach is not able to produce. In contrast, it was shown in example 8.2.18 that there is a basic infobase contraction \ominus for which $[p, q] \ominus (p \wedge q) \approx [p \vee q, p \vee q]$. In fact, it can be shown that *every* basic infobase contraction yields the same result. \square

8.4 Iterated infobase change

Although an infobase IB induces the unique theory contraction $-_{IB}$, infobases do not contain enough information to determine a basic infobase contraction or revision. To do that, we also need a relevance selection function rs . Once rs is fixed, though, we are dealing with a specific basic infobase contraction and revision, which allows for the possibility of iterated infobase change. In this section we investigate whether iterated infobase change measures up to the postulates supplied by Darwiche and Pearl (see section 7.3) and Lehmann (see section 7.4). To do so, we have to work on the level of epistemic states. Recall from section 7.3 that every epistemic state Φ is assumed to have associated with it a belief set $K(\Phi)$ and a $K(\Phi)$ -faithful total preorder \preceq_Φ . To bring infobase change into this framework, we assume that it is possible to extract a unique infobase IB_Φ from every epistemic state Φ . This implies that $K(\Phi) = Cn(S(IB_\Phi))$ and that \preceq_Φ is identical to the IB_Φ -induced faithful total preorder \preceq_{IB_Φ} . Furthermore, since Darwiche and Pearl operate under the assumption of a finitely generated propositional language L with a valuation semantics (V, \models) , we shall do the same for the rest of this section.

Recall from our discussion of DP-revision in section 7.3 that in order to simplify matters, we decided to *equate* every epistemic state Φ with the ordered pair $(K(\Phi), \preceq_\Phi)$. With the incorporation of infobases into epistemic states, it is no longer possible to adhere to this simplification. The reason is that infobases contain more information than such ordered pairs. That is, while every infobase IB is uniquely associated with the ordered pair $(Cn(S(IB)), \preceq_{IB})$, this ordered pair may be associated with different infobases. For example, letting $IB = [p, q]$ and $IC = [p \wedge q, p \vee q]$, it is easy to check that $Cn(S(IB)) = Cn(S(IC))$, and that \preceq_{IB} and \preceq_{IC} are identical. Furthermore, the fact that we only deal with finitely generated propositional logics makes it easy to see that every ordered pair of this kind can be obtained from some infobase.

Lemma 8.4.1 *For every ordered pair of the form (K, \preceq) where K is a belief set and*

\preceq is a K -faithful total preorder, there is an infobase IB such \preceq and \preceq_{IB} are identical, and $K = Cn(S(IB))$.

Proof Pick any ordered pair of the form (K, \preceq) where K is a belief set and \preceq is a K -faithful total preorder. Since L is a finitely generated propositional language, V contains a finite number of interpretations. The total preorder \preceq thus partitions V into a finite number of subsets (blocks). Let us assume that there are n such blocks. We assign each of them a unique index from 1 to n according to their relative positions in \preceq , leaving us with the n indexed blocks P_1, \dots, P_n . That is, for $1 \leq i, j \leq n$, $i < j$ iff for every $u \in P_i$ and every $v \in P_j$, $u \prec v$. Now, for any $W \subseteq V$, let α_W be some wff that axiomatises W . (Since L is finitely generated, such a wff always exists.) For $1 \leq i \leq n$, let $\beta_i \equiv \alpha_W$ where $W = \bigcup_{1 \leq j \leq i} P_j$. We define an infobase IB as follows: if $\perp \in K$, then IB contains exactly one instance of each of the wffs in $\{\perp\} \cup \bigcup_{1 \leq i \leq n} \{\beta_i\}$, otherwise IB contains exactly one instance of each of the wffs in $\bigcup_{1 \leq i \leq n} \{\beta_i\}$. It is easily verified that \preceq and \preceq_{IB} are identical, and that $Cn(S(IB)) = K$. \square

More importantly, perhaps, is the fact that the extra information contained in infobases plays an important role in the process of infobase change, as the next example shows.

Example 8.4.2 Let \ominus be the basic infobase contraction obtained from the relevance selection function rs , where $rs(IB, \alpha) = IB^{-\alpha}$, for every $IB \in \mathcal{IB}$ and every $\alpha \in L$, and let \otimes be the basic infobase revision defined in terms of \ominus using (Def \otimes from \ominus). Now, let $IB = [p, q]$ and let $IC = [p \wedge q, p, q, p \vee q, p \rightarrow q, q \rightarrow p]$. Clearly $Cn(S(IB)) = Cn(S(IC))$ and it is also easy to see that \preceq_{IB} and \preceq_{IC} are identical, and are represented graphically in figure 8.2. Yet, it can be verified that $IB \otimes (p \wedge \neg q) \approx [p, \top, p \wedge \neg q]$ and that $IC \otimes (p \wedge \neg q) \approx [p, p, p \vee q, p \vee q, \top, q \rightarrow p, p \wedge \neg q]$. So $IB \otimes (p \wedge \neg q)$ and $IC \otimes (p \wedge \neg q)$ induce different faithful total preorders, as can be seen in figure 8.3. \square

Having established that epistemic states need to have a richer structure than ordered pairs of the form $(K(\Phi), \preceq_\Phi)$, we now turn to the definition of revision on epistemic state in terms of basic infobase revision.

$$(\text{Def } * \text{ from } \otimes) \quad \left[\begin{array}{l} K(\Phi * \alpha) = Cn(IB_\Phi \otimes \alpha) \\ \preceq_{\Phi * \alpha} = \preceq_{(IB_\Phi \otimes \alpha)} \end{array} \right]$$

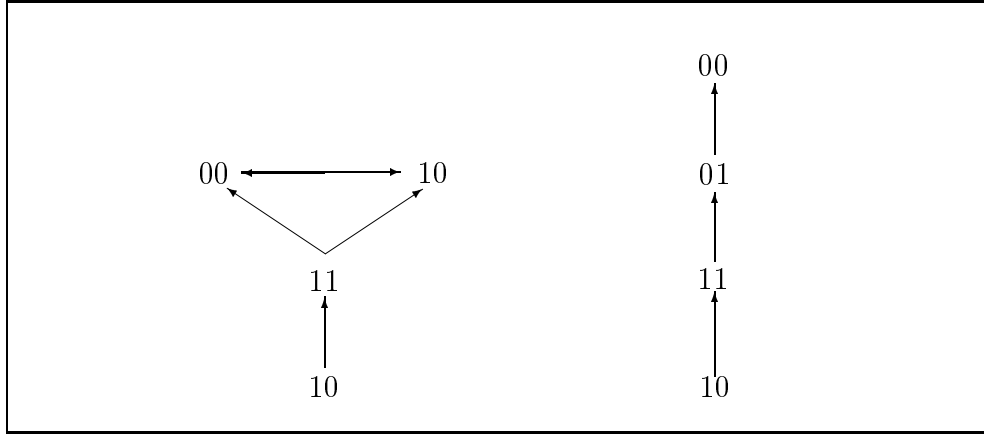


Figure 8.3: A graphical representation of the total preorders used in example 8.4.2. On the left is the $(IB \otimes (p \wedge \neg q))$ -induced faithful total preorder and on the right the $(IC \otimes (p \wedge \neg q))$ -induced faithful total preorder. As usual, the applicable preorder is the reflexive transitive closure of the relation determined by the arrows.

Definition 8.4.3 We refer to the revision on epistemic states defined in terms of a basic infobase revision \otimes using $(\text{Def} \ast \text{from } \otimes)$ as the \otimes -associated revision on epistemic states. \square

It is easily verified that the revisions on epistemic states associated with basic infobase revisions all satisfy $(E\ast 1)$ to $(E\ast 8)$.

Proposition 8.4.4 Let \otimes be a basic infobase revision, and let \ast be the \otimes -associated revision on epistemic states. Then \ast satisfies $(E\ast 1)$ to $(E\ast 8)$.

Proof Follows from theorem 7.3.1 and part (2) of corollary 8.2.27. \square

8.4.1 DP-revision

When placed in the framework for iterated belief change proposed by Darwiche and Pearl, basic infobase revision yields favourable results. The revisions on epistemic states associated with basic infobase revisions satisfy all but the first one of the four DP-postulates. The satisfaction of these three DP-postulates rely on the following two simple results.

Lemma 8.4.5 *Let \otimes be a basic infobase revision and let rs be the relevance selection function from which \otimes is obtained.*

1. *If $v \in M(\neg\alpha)$ then, for every β in IB , $v \in M(\beta)$ iff $v \in M\left(w_{(IB, \neg\alpha)}^{rs(IB, \neg\alpha)}(\beta)\right)$.*
2. *For every β in IB , if $v \in M(\beta)$ then $v \in M\left(w_{(IB, \neg\alpha)}^{rs(IB, \neg\alpha)}(\beta)\right)$.*

Proof By proposition 8.2.15, $M\left(w_{(IB, \neg\alpha)}^{rs(IB, \neg\alpha)}(\beta)\right) = M(\beta) \cup N_\beta^{rs(IB, \neg\alpha)}(\neg\neg\alpha)$ for every β in IB .

1. Follows from the fact that $N_\beta^{rs(IB, \neg\alpha)}(\neg\neg\alpha) \subseteq M(\alpha)$ for every β in IB .
2. Follows from the fact that $M(\beta) \subseteq M\left(w_{(IB, \neg\alpha)}^{rs(IB, \neg\alpha)}(\beta)\right)$.

□

Proposition 8.4.6 *Let \otimes be a basic infobase revision, and let \ast be the \otimes -associated revision on epistemic states. Then \ast satisfies (DP2)–(DP4), but does not necessarily satisfy (DP1).*

Proof To show that \ast does not necessarily satisfy (DP1), let L be generated by the atoms p and q , with a valuation semantics (V, \models) where $V = \{00, 01, 10, 11\}$. Let $IB_\Phi = [p \leftrightarrow q, p \vee \neg q, \neg p \vee \neg q, \neg q]$ and let \otimes be the basic infobase revision obtained from the relevance selection function rs for which $rs(IB, \alpha) = IB^{-\alpha}$ for every $IB \in \mathcal{IB}$ and every $\alpha \in L$. It can be verified that

$$IB_\Phi \otimes (p \vee q) \approx [p \vee \neg q, \neg p \vee \neg q, \neg q, p \vee q],$$

$$K((\Phi \ast (p \vee q)) \ast q) = Cn(S((IB_\Phi \otimes (p \vee q)) \otimes q)) = Cn(q), \text{ and}$$

$$K(\Phi \ast q) = Cn(S(IB_\Phi \otimes q)) = Cn(p \wedge q).$$

So $q \models p \vee q$, but $K((\Phi \ast (p \vee q)) \ast q) \neq K(\Phi \ast q)$, which is a violation of (DP1).

For (DP2)–(DP4), it suffices, by theorem 7.3.4, to show that \ast satisfies (DPR2)–(DPR4). Let rs be the relevance selection function from which \otimes is obtained and pick any epistemic state Φ .

For (DPR2), observe that since $IB_\Phi \otimes \alpha$ is obtained by replacing every wff β in IB_Φ with $w_{(IB_\Phi, \neg\alpha)}^{rs(IB_\Phi, \neg\alpha)}(\beta)$ and then adding α , it follows from part (1) of lemma 8.4.5 that

$u_{IB_\phi} = u_{IB_\phi \otimes \alpha}$ (where u_{IB_ϕ} and $u_{IB_\phi \otimes \alpha}$ are the IB_ϕ -number and the $(IB_\phi \otimes \alpha)$ -number of u), for every $u \in M(-\alpha)$. And since \preceq_Φ is the IB_ϕ -induced faithful total preorder, and $\preceq_{\Phi * \alpha}$ is the $(IB_\phi \otimes \alpha)$ -induced faithful total preorder, it then follows that $u \preceq_\Phi v$ iff $u \preceq_{\Phi * \alpha} v$ for every $u, v \in M(-\alpha)$. So (DPR2) is satisfied.

For (DPR3) and (DPR4), note that part (2) of lemma 8.4.5 ensures that $u_{IB_\phi} \leq u_{IB_\phi \otimes \alpha}$. Combined with part (1) of lemma 8.4.5, it then follows for every $u \in M(\alpha)$ and every $v \in M(-\alpha)$, that if $u_{IB_\phi} > v_{IB_\phi}$ then $u_{IB_\phi \otimes \alpha} > v_{IB_\phi \otimes \alpha}$. So, for every $u \in M(\alpha)$ and every $v \in M(-\alpha)$, if $u \prec_\Phi v$ then $u \prec_{\Phi * \alpha} v$, which means that (DPR3) holds. Similarly, from parts (1) and (2) of lemma 8.4.5 it follows for every $u \in M(\alpha)$ and every $v \in M(-\alpha)$, that if $u_{IB_\phi} \geq v_{IB_\phi}$ then $u_{IB_\phi \otimes \alpha} \geq v_{IB_\phi \otimes \alpha}$. So, for every $u \in M(\alpha)$ and every $v \in M(-\alpha)$, if $u \preceq_\Phi v$ then $u \preceq_{\Phi * \alpha} v$; that is, (DPR4) holds. \square

It is our contention that the violation of (DP1) by basic infobase revision is an indication that this postulate is perhaps too restrictive to accommodate a wide range of rational forms of revision. Below we give a realistic example in support of this claim.⁶

Example 8.4.7 I have a circuit containing two components; an adder and a multiplier. I have made three independent observations about these components.

1. The adder is working.
2. The multiplier is working.
3. If the adder doesn't work then the multiplier also doesn't work.

Another observation now indicates that at least one of the two components is not working. In trying to incorporate this new information, we have to discard (or weaken) at least one of the first two observations. Moreover, we cannot retain both observations (2) and (3), for they imply observation (1). So it seems reasonable to retain the belief that the adder is working and the belief that a broken adder implies a broken multiplier. Together with the new information that at least one of the components is broken, it then follows that it is the multiplier that is broken.

This line of reasoning can be formalised by using a propositional language generated by the two atoms a (indicating that the adder is working) and m (indicating that the multiplier is working) with a valuation semantics (V, \models) , where $V = \{00, 01, 10, 11\}$.⁷

⁶This example was inspired by a similar one proposed by Darwiche and Pearl [1997, p. 12].

⁷We adopt the convention of letting the first digit denote the truth value of a and the second digit the truth value of m .

My initial infobase then looks like this: $IB = [a, m, \neg a \rightarrow \neg m]$. Figure 8.4 contains a graphical representation of the IB -induced faithful total preorder \preceq_{IB} . It is easily verified that for any basic infobase revision \otimes , $Cn(S(IB \otimes \neg(a \wedge m))) = Cn(a \wedge \neg m)$, which means that m should be discarded and that a and $\neg a \rightarrow \neg m$ should be retained. But what should the weakened version of the discarded wff m look like?

One reasonable option is to discard it completely, or, what amounts to the same thing, to weaken it so that it becomes logically valid. Formally, this can be accomplished as follows. Let rs be a relevance selection function such that $rs(IB, a \wedge m) = IB^{-(a \wedge m)} = \{m\}$. Since $IB^{-(a \wedge m)}$ is $(IB, \alpha \wedge m)$ -relevant, there is such an rs . Now consider the basic infobase contraction \ominus which is obtained using rs . It can be verified that $IB \ominus \neg\neg(a \wedge m) \approx IB \ominus (a \wedge m) \approx [a, \top, \neg a \rightarrow \neg m]$ and therefore $IB \otimes \neg(a \wedge m) \approx [a, \top, \neg a \rightarrow \neg m, \neg(a \wedge m)]$, where \otimes is the basic infobase revision defined in terms of \ominus using (Def \otimes from \ominus). Figure 8.4 contains a graphical representation of the $(IB \otimes \neg(a \wedge m))$ -induced faithful total preorder.

To see that the revision \ast defined in terms of \otimes using (Def \ast from \otimes) violates (DP1), note that an inspection of figure 8.4 shows that $Cn(S(IB \otimes \neg a)) = Cn(\neg a)$, but that $Cn(S((IB \otimes \neg(a \wedge m)) \otimes \neg a)) = Cn(\neg a \wedge \neg m)$. So $K((\Phi \ast \neg(a \wedge m)) \ast \neg a) \neq K(\Phi \ast \neg a)$ even though $\neg a \models \neg(a \wedge m)$ where Φ is an epistemic state such that $IB_\Phi = IB$. And this constitutes a violation of (DP1). \square

There is a particular form of basic infobase revision which does satisfy (DP1), though. It corresponds to what we have referred to as the coherentist approach to infobase change on page 246 in section 8.2.1.

Definition 8.4.8 A *coherentist* basic infobase revision \otimes is a basic infobase revision such that $rs(IB, \alpha) = IB$ for every $\alpha \in L$, for the relevance selection function rs from which \otimes is obtained. \square

To show that a coherentist basic infobase revision satisfies (DP1) we need the following two lemmas.

Lemma 8.4.9 For every $u \in \text{Min}_{\preceq_{IB}}(\alpha)$, $N_u^{IB}(\alpha) \subseteq \text{Min}_{\preceq_{IB}}(\alpha)$.

Proof Pick any $u \in \text{Min}_{\preceq_{IB}}(\alpha)$ and any $v \in N_u^{IB}(\alpha)$. By definition, $v \in M(\alpha)$, and u and v satisfy exactly the same wffs in IB . So the IB -numbers of u and v are the same, and therefore $v \in \text{Min}_{\preceq_{IB}}(\alpha)$. \square

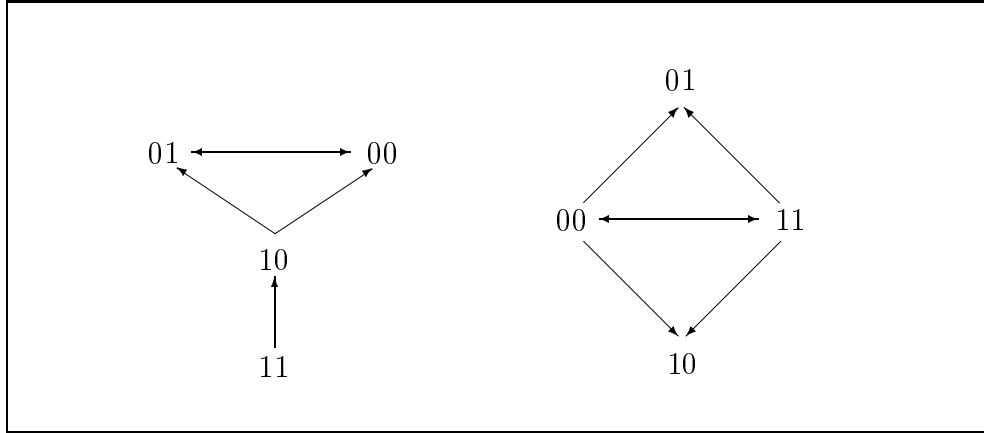


Figure 8.4: A graphical representation of the total preorders used in example 8.4.7. On the left is the IB -induced faithful total preorder and on the right the $(IB \otimes \neg(a \wedge m))$ -induced faithful total preorder. As usual, the applicable preorder is the reflexive transitive closure of the relation determined by the arrows.

Lemma 8.4.10 *If $v \in M(\alpha) \setminus \text{Min}_{\preceq_{IB}}(\alpha)$ then, for every β in IB , $v \in M(\beta)$ iff $v \in M(w_{(IB, \neg\alpha)}^{IB}(\beta))$.*

Proof Pick any $v \in M(\alpha) \setminus \text{Min}_{\preceq_{IB}}(\alpha)$ and any β in IB . By proposition 8.2.15, $M(\beta) \subseteq M(w_{(IB, \neg\alpha)}^{IB}(\beta))$, and so $v \in M(\beta)$ implies $v \in M(w_{(IB, \neg\alpha)}^{IB}(\beta))$. Conversely, suppose that $v \in M(w_{(IB, \neg\alpha)}^{IB}(\beta))$. By lemma 8.4.9, $v \notin N_{\beta}^{IB}(\alpha)$, and it therefore follows from proposition 8.2.15 that $v \in M(\beta)$. \square

Proposition 8.4.11 *Let \otimes be the coherentist basic infobase revision and let \ast be the revision on epistemic states defined in terms of \otimes using (Def \ast from \otimes). Then \ast satisfies (DP1).*

Proof By theorem 7.3.4, it suffices to show that \ast satisfies (DPR1). Let Φ be any epistemic state. So \preceq_{Φ} is the IB_{Φ} -induced faithful total preorder. We have to show that $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi \ast \alpha} v$ for every $u, v \in M(\alpha)$.

Recall from definitions 8.2.14 and 8.2.24 that $IB_{\Phi} \otimes \alpha$ is obtained by replacing every wff β in IB_{Φ} with $w_{(IB_{\Phi}, \neg\alpha)}^{IB_{\Phi}}(\beta)$ and then adding α . From lemma 8.4.10 it follows that the $(IB_{\Phi} \otimes \alpha)$ -number of u is one more than the IB_{Φ} -number of u , for every $u \in M(\alpha) \setminus \text{Min}_{\preceq_{\Phi}}(\alpha)$. So $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi \ast \alpha} v$ for every $u, v \in M(\alpha) \setminus \text{Min}_{\preceq_{\Phi}}(\alpha)$.

Next, observe that the IB_Φ -number of every $u \in \text{Min}_{\preceq_\Phi}(\alpha)$ is greater than the IB_Φ -number of every $v \in M(\alpha) \setminus \text{Min}_{\preceq_\Phi}(\alpha)$. Moreover, by part (2) of corollary 8.2.27 it follows that $M(S(IB_\Phi \otimes \alpha)) = \text{Min}_{\preceq_\Phi}(\alpha)$. So the $(IB_\Phi \otimes \alpha)$ -number of every $u \in \text{Min}_{\preceq_\Phi}(\alpha)$ is greater than the $(IB_\Phi \otimes \alpha)$ -number of every $v \in M(\alpha) \setminus \text{Min}_{\preceq_\Phi}(\alpha)$. Therefore $u \preceq_\Phi v$ iff $u \preceq_{\Phi * \alpha} v$ for every $u \in \text{Min}_{\preceq_\Phi}(\alpha)$ and every $v \in M(\alpha) \setminus \text{Min}_{\preceq_\Phi}(\alpha)$, and $u \preceq_\Phi v$ iff $u \preceq_{\Phi * \alpha} v$ for every $v \in \text{Min}_{\preceq_\Phi}(\alpha)$ and every $u \in M(\alpha) \setminus \text{Min}_{\preceq_\Phi}(\alpha)$.

Finally, observe that elements of $\text{Min}_{\preceq_\Phi}(\alpha)$ all have the same IB_Φ -number, and since $M(S(IB_\Phi \otimes \alpha)) = \text{Min}_{\preceq_\Phi}(\alpha)$, the elements of $\text{Min}_{\preceq_\Phi}(\alpha)$ all have the same $(IB_\Phi \otimes \alpha)$ -number as well. So $u \preceq_\Phi v$ iff $u \preceq_{\Phi * \alpha} v$ for every $u, v \in \text{Min}_{\preceq_\Phi}(\alpha)$, which means we are done. \square

8.4.2 L-revision

We turn now to Lehmann's framework for iterated revision which was discussed in section 7.4.⁸ Since his postulates (L*1), (L*2), (L*3) and (L*6) correspond exactly to (E*1), (E*2), (E*3) and (E*6) respectively, it follows from proposition 8.4.4 that the revision $*$ on epistemic states obtained in terms of a basic infobase revision using (Def $*$ from \otimes) satisfy these four postulates of Lehmann. Furthermore, since (L*7) is a weakened version of (DP2) (see section 7.4, page 223), it follows from proposition 8.4.6 that $*$ also satisfies (L*7). It does not necessarily satisfy (L*4), (L*5) and (L*8), though, as the following example shows.

Example 8.4.12 Let \otimes be the basic infobase revision obtained from the relevance selection function rs for which $rs(IB, \alpha) = IB^{-\alpha}$ for every $IB \in \mathcal{IB}$ and every $\alpha \in L$.

1. Let $IB = [p \wedge \neg q, p \vee q]$. Clearly $IB \otimes p \approx [p \wedge \neg q, p \vee q, p]$. It can be verified that $Cn(S((IB \otimes p) \otimes q)) = Cn(p \wedge q)$, but that $Cn(S(IB \otimes q)) = Cn(q)$. Taking p as α and q as the sequence of wffs σ , this is a violation of (L*4).
2. Let $IB = [p \leftrightarrow q, p \vee \neg q, \neg p \vee \neg q, \neg q]$. It can be verified that

$$\begin{aligned}
 IB \otimes q &\approx [p \leftrightarrow q, p \vee \neg q, p \vee \neg q, q], \\
 IB \otimes p \vee q &\approx [p \vee \neg q, \neg p \vee \neg q, \neg q, p \vee q], \\
 (IB \otimes p \vee q) \otimes q &\approx [p \vee q, q], \\
 Cn(S(((IB \otimes p \vee q) \otimes q) \otimes \neg q)) &= Cn(p \wedge \neg q), \text{ and} \\
 Cn(S((IB \otimes q) \otimes \neg q)) &= Cn(\neg p \wedge \neg q).
 \end{aligned}$$

⁸Recall that Lehmann concerns himself only with revisions by satisfiable wffs.

Taking $p \vee q$ as β , q as α , and $\neg q$ as the sequence of wffs σ , this constitutes a violation of (L*5).

3. Let $IB = [p \vee q, p \vee \neg q]$. Clearly

$$\begin{aligned} IB \otimes p &\approx [p \vee q, p \vee \neg q, p], \\ (IB \otimes p) \otimes q &= [p \vee q, p \vee \neg q, p, q], \text{ and} \\ (IB \otimes p) \otimes p \wedge q &= [p \vee q, p \vee \neg q, p, p \wedge q]. \end{aligned}$$

It can be verified that

$$\begin{aligned} Cn(S(((IB \otimes p) \otimes q) \otimes \neg p)) &= Cn(\neg p \wedge q), \text{ and} \\ Cn(S(((IB \otimes p) \otimes p \wedge q) \otimes \neg p)) &= Cn(\neg p). \end{aligned}$$

With p as α , q as β , and $\neg p$ as the sequence of wffs σ , it follows that (L*8) is violated.

□

An examination of this example suggests that, unlike the DP-postulates, (L*4), (L*5) and (L*8) are fundamentally incompatible with basic infobase revision.

8.5 Future research

This chapter has laid the foundation for a theory of infobase change, but it is clear that much still needs to be done. Infobase change, as we have currently defined it, assumes that the wffs contained in the infobase IB have equal epistemic weight. But there may be good reasons for regarding some wffs in IB as epistemologically more important than others, as the following example, which is part of an example by Hansson [1992b], attests to.

Example 8.5.1 “A geography student sees one of his fellow students pick up a book in the library. The title of the book is *The University at Niamey*. He asks, ‘Where is Niamey?’, and receives the answer, ‘It is a Nigerian city’.

Next day, in an oral examination, the professor asks our student, ‘What do you know about Niamey?’—‘It is a university town in Nigeria’—‘It most certainly isn’t’...the student believes what the professor says, and adjust his beliefs accordingly.” □

We use the propositional language generated by the atoms p and q to represent the situation above, where p denotes the assertion that there is university in Niamey, and q denotes the assertion that Niamey is a town in Nigeria. So the infobase IB is $[p, q]$ and the student performs a $p \wedge q$ -contraction of IB . It is easy to verify that every basic infobase contraction of IB by $p \wedge q$ yields an infobase that is element-equivalent to $[p \vee q, p \vee q]$ (see example 8.2.18). But, as Hansson [1992b] argues, it is reasonable to assume that the result of the above contraction should be element-equivalent to $[p]$. This is because of the extra-logical assumption that information obtained in library books is more reliable than information obtained from fellow students, which allows us to retain p rather than q .

One way in which to represent such extra-logical information is in terms of orderings of *epistemic relevance* on IB . Nebel [1990, 1991, 1992] requires of epistemic relevance orderings to be total preorders on a base B . When applied to infobase change, the aim would be to use an epistemic relevance ordering on an infobase IB to obtain a suitable $S(IB)$ -faithful total preorder. An appropriate infobase change operation would then be constructed in a manner analogous to the way it is currently being constructed.

One of the main differences between infobase change and many approaches to base change is illustrated by example 8.2.18, where a wff that is not contained in the infobase $IB = [p, q]$ finds its way into the resulting infobase $IB \ominus p \wedge q$. And while this seems to be the correct solution in many respects, it is not quite in tune with the intuition that the wffs in an infobase represent independently obtained beliefs. For it seems counterintuitive to regard a wff that is merely entailed by the wffs in IB as an independently obtained belief contained in $IB \ominus p \wedge q$. It is with this kind of example in mind that Rott [1992a] writes as follows (In the quotation H represents the base $\{p, q\}$):

“...Even after conceding that one of p and q may be false, we should still cling to the belief that the other one is true. But $H' = \{p \vee q\}$ is no base which can be constructed naturally from H —it certainly does not record any explicit belief. We are faced with a deep-seated dilemma...”

Rott ultimately decides against the inclusion of such wffs, arguing that bases should only contain explicit beliefs.⁹ We conclude this section by arguing that a priority or-

⁹Hansson [1996] mentions the use of disjunctively closed bases (in which the disjunction $\alpha \vee \beta$ of every $\alpha, \beta \in B$ is also in B) as a possible solution to problems of this kind. Unfortunately this ensures

dering, similar in spirit to the epistemic relevance orderings, may provide an acceptable solution. The idea is to split the infobase obtained from an infobase contraction into two partitions; one containing the explicit beliefs and the other containing the *introspective beliefs*. After an infobase contraction of the infobase IB by the wff α , the explicit beliefs consists of the α -retained wffs of IB , while the introspective beliefs are appropriately weakened versions of the α -discarded wff β of IB . Wffs that were, at some stage, obtained directly from independent sources thus constitute the explicit beliefs, while wffs such as the ones logically equivalent to $p \vee q$ in example 8.2.18 are regarded as beliefs obtained by introspection during the contraction process, and are thus to be seen as carrying less epistemic weight than the explicit beliefs.

that bases can't be finite. And in any case, Hansson does not regard it as an acceptable solution, warning that it should be seen as an interesting special case, rather than a required property of bases.

Chapter 9

Conclusion

It can go on and on, or someone must write “The End” to it.

Gerald R. Ford, 38th US President

One of the most important issues in the area of knowledge representation is to find appropriate representations of the epistemic states of agents equipped with the ability to reason intelligently. In this dissertation we have concentrated on semantic representations of the part of an epistemic state pertaining to belief change. We chose the AGM approach to theory change as our starting point, primarily because of its importance in the study of belief change. Historically, AGM theory change has become synonymous with a presentation in terms of postulates, as outlined in section 2.1, plus the following four basic construction methods:

1. The method of partial meet contraction, which uses remainders [Alchourrón et al., 1985].
2. The method of safe contraction, which makes use of entailment sets [Alchourrón and Makinson, 1985, Rott, 1992b].
3. A construction method involving the EE-orderings of Gärdenfors and Makinson [1988] and Gärdenfors [1988].
4. A semantic method of construction, in terms of systems of spheres [Grove, 1988].

The semantic method we have chosen to focus on is a slight variation on Grove's systems of spheres. It involves a set of total preorders on the set of interpretations of the logic language under consideration, which we have chosen to refer to as the faithful total preorders [see Katsuno and Mendelzon, 1991, Peppas and Williams, 1995].¹ While the representation theorems involving these construction methods allow us to move from any one construction method to any one of the others (at least in principle), it is, in our view, difficult to escape the conclusion that the semantic methods are, in an important sense, more fundamental than the others.

The use of faithful total preorders is model-theoretic in nature, and has been used as such in our technical results. But it can also be given an information-theoretic flavour in terms of the infatoms introduced in section 3.1. The basic idea is that the bits of information making up the belief set of an agent are ordered according to their entrenchment or credibility, and that any changes in beliefs are ultimately made with this ordering in mind. It is our contention that such an information-theoretic view of belief change provides an appropriate setting for further studies in belief change.

Not long after the inception of the research area known as nonmonotonic reasoning, researchers started to point out connections between this field and the enterprise of belief change. The link is provided by theory revision operations on the one hand, and nonmonotonic consequence relations on the other. The basic idea is that results obtained from an α -revision can be seen as the plausible (but nonmonotonic) consequences resulting from the adoption of the evidence α , and vice versa. What is most interesting from our point of view, is that a slight variation on expectation based nonmonotonic reasoning [Gärdenfors and Makinson, 1994] can be constructed from the faithful total preorders, thus leading to the claim that the processes involved in theory revision and nonmonotonic reasoning are identical. While such results indicate a formal connection between theory revision and nonmonotonic reasoning, it has been argued that one should not attempt to extend this link to the epistemological level as well [Gärdenfors and Makinson, 1994]. It is our view that both these areas can be incorporated into a more general formal theory of cautious and bold reasoning, with nonmonotonic reasoning being viewed as a form of reasoning which is bolder than the type of reasoning encountered in theory revision.

¹For some reason it seems that the use of systems of spheres is more popular in philosophical circles, while artificial intelligence researchers prefer to use faithful total preorders.

It is a curious feature of many a nonmonotonic reasoning system that, while the examples used in justifying the formal construction have a dynamic quality to them, the construction itself is viewed as the description of a static process. This behaviour can only be explained by the (implicit) assumption that the adoption of two pieces of evidence in sequence, yields results that are identical to that obtained from the simultaneous adoption of the same bits of evidence. Using the connection between nonmonotonic reasoning and theory revision, we have argued that this is too strong a restriction to place on all forms of nonmonotonic reasoning. It is hoped that future research on nonmonotonic reasoning systems will take this result into account.

Orderings of entrenchment on wffs are frequently advanced as appropriate representations of the epistemic states of agents; at least with regard to belief change. We have surveyed the forms of entrenchment found in the literature, and presented a novel version of entrenchment — *refined entrenchment* — which is intended as an alternative to the EE-orderings of Gärdenfors and Makinson [1988] and Gärdenfors [1988]. The construction of refined entrenchment orderings involves the use of the faithful modular weak partial orders, instead of the faithful total preorders, thereby ensuring the elimination of some of the undesirable properties of the EE-orderings. The use of the faithful modular weak partial orders paves the way for the introduction of a more general set of faithful orderings, the faithful layered preorders, from which both the EE-orderings and the refined entrenchment orderings can be constructed. Using these results, we have argued that such orderings on interpretations (and on their information-theoretic counterparts) ought to be seen as more fundamental than the entrenchment orderings on wffs generated from them.

One of the most controversial aspects of AGM theory change is the insistence on the inclusion of the Recovery postulate (K–6). Those theory removal operations that satisfy the first five basic AGM postulates have come to be known as withdrawal operations. In recent years, there have been a number of proposals aimed at constructing rational forms of theory withdrawal that do not, in general, satisfy Recovery. Following a survey of withdrawal operations, we have introduced a new member of this family, dubbed *systematic withdrawal*. The method for constructing systematic withdrawal is semantic in nature. It involves the faithful modular weak partial orders; the preorders used in the construction of refined entrenchment. We have argued that systematic withdrawal seems to retain the advantages of other forms of withdrawal, but does not suffer from their undesirable properties. By applying the method used in the construc-

tion of systematic withdrawal to the faithful layered preorders, we have obtained a set of principled withdrawal operations which includes systematic withdrawal as well as the severe withdrawal operations of Rott and Pagnucco [1999]. From our investigation into withdrawal it seems reasonable to advance the thesis that any principled form of withdrawal will be amenable to semantical construction, in terms of some kind of ordering on interpretations (or infatoms).

Due to its violation of the principle of Categorical Matching, AGM theory change has been shown not be suitable for a satisfactory description of iterated belief change. In one of the most important recent advances in the field of belief change, Darwiche and Pearl [1994, 1997] have shown that investigations of iterated belief change ought to be conducted on the level of epistemic states. The main results about their proposed framework rely on a semantic view of epistemic states, which states that it is possible to extract from every epistemic state a unique faithful total preorder and a unique belief set. Our investigation into iterated belief change consist of a survey of the proposed frameworks of Darwiche and Pearl [1994, 1997] and Lehmann [1995], a discussion of transmutation, which can be viewed as a generalised version of iterated belief change, and a discussion about two revision operations recently proposed by Papini [1998, 1999]. Papini's revision operations and work done by Nayak [1994b], Nayak et al. [1996] and Liberatore and Schaerf [1998], coupled with the move to view revision as an operation on epistemic states, have also served as inspiration for the proposal to investigate operations involving the merging of two epistemic states. It seems difficult to conduct an investigation into merging without incorporating the semantic view.

Most of the work in this dissertation is of a declarative nature. It addresses the question of how an agent may employ the semantic structures extracted from an epistemic state to perform belief change, but ignores, for the most part, the equally important question of how an agent may arrive at a particular epistemic state. We have shown how data structures called infobases can be used to achieve the latter objective. An infobase is a finite ordered list of wffs. It is associated with a belief set — the set of wffs entailed by the wffs in the infobase, and the structure of an infobase is exploited to induce a faithful total preorder. Every infobase is thus associated with a unique belief set and a unique faithful total preorder; the two components of an epistemic state needed to perform theory change. While the basic idea of associating extra-logical information with the structure of a set of wffs is nothing new, the particular method we have employed ensures that faithful *total* preorders are obtained, and is a novel

contribution. Although infobase change views knowledge level matters as important, it is also concerned with appropriate issues on the symbol level. Unlike the process that has become known as base change, infobase change involves the weakening of wffs in an infobase, rather than the removal of some wffs.

In conclusion, we restate the three main questions with which this dissertation is concerned and indicate to what extent answers have been provided for them.

1. How should an epistemic state (or at least the part pertaining to belief change) be represented?

The work in this dissertation suggests that the answer to this question consists of a single word; “semantically”. Much of the work done here indicates that an ordered pair, consisting of a belief set and a layered preorder on a set of infatoms of the logic language under consideration, is an appropriate representation of an epistemic state, at least for belief change operations such as theory revision and theory withdrawal. And while some of the later chapters, in particular chapter 8, suggest that richer structures are needed for more realistic belief change operations, there is a clear indication that such richer structures also need to be semantic in nature.

2. How does an agent use an epistemic state to perform belief change?

The most important issue that has been resolved in connection with this question is that any belief change operation ought to produce, not just a belief set, but rather a complete new epistemic state. Furthermore, it has become clear that, while the process of identifying the belief sets associated with those epistemic states resulting from belief change operations is well laid out (notwithstanding some variations in the methods for doing so), much work still has to be done to determine the permissible ways of arriving at complete epistemic states resulting from belief change operations.

3. How does an agent arrive at a particular epistemic state?

Our main contribution in providing an answer to this question is the use of infobases. We assume that wffs in an infobase are independently obtained and then exploit the structure of the infobase to aid in the construction of an appropriate epistemic state.

The use of infobases in this fashion is just a first approximation, although it seems to have the potential for developing into a full-fledged theory.

9.1 Future research

This dissertation provides guidelines for some promising areas of future research, some of which have already been touched on in the relevant chapters. We briefly outline the most interesting of these.

It seems worthwhile to explore the connections between the information-theoretic semantics described in chapter 3 and other logic-oriented approaches such as that of Barwise and Seligman [1997]. It is also possible that there might be a link with algorithmic information theory in the sense of Shannon [1964, 1993] and Chaitin [1987].

Having accepted the importance of semantic structures for the construction of belief change operations, it is tempting to re-evaluate some of the generalisations of AGM theory change which do not admit semantic descriptions. This has, to some extent, already been accomplished with base change, resulting in the definition of infobase change. Another such area is that of multiple change; theory change operations involving sets of wffs instead of single wffs.² Some proposals for multiple contraction have been made by Fuhrmann and Hansson [1994]. One of their proposals, *package contraction*, is constructed from generalised versions of the partial meet contractions (see section 2.2). A closer look at this construction from a semantic point of view seems to point to some inconsistencies in the choice of admissible belief sets when contracting by certain sets of wffs, and also suggests a possible solution to this problem.

In recent years, considerable progress has been made in the area of iterated belief change. The framework provided by Darwiche and Pearl [1994, 1997], in particular, has provided an excellent starting point. However, much work still needs to be done in this regard. Some recent results suggest that the first two DP-postulates may be too restrictive. The challenge is thus to weaken these two postulates in an appropriate fashion. One possibility is suggested by concentrating on the semantic versions of the four DP-postulates. It involves the kind of restriction placed on the relative ordering of interpretations which is found in (DPR4). The idea is that an α -revision need not leave the relative order of the models of α unchanged, as (DPR1) requires. Instead, it

²Peppas and Sprakis [1999] have recently provided a semantic description of Lindström's [1991] proposed version of multiple revision.

only requires of model v of α that is strictly higher up in the ordering than a model u of α , to be at least as high up as u after an α -revision. From an information-theoretic point of view, this means that an α -revision may cause a content bit i of $\neg\alpha$ to become less entrenched, but i may not become strictly less entrenched than any of the content bits of $\neg\alpha$ which are currently at most as entrenched as i . A similar weakening of (DPR2) would require of a model u of $\neg\alpha$ that is strictly lower down in the ordering than a model v of $\neg\alpha$, to be at most as high up as v after an α -revision. Information-theoretically, this means that an α -revision permits a content bit i of α to become more entrenched, but i may not become strictly more entrenched than any of the content bits of α which are currently at least as entrenched as i . It remains to be seen whether these suggested properties will turn out to be appropriate postulates for iterated revision.

Further investigations into iterated belief change are also bound to have an impact on research concerning multi-agent belief change [Kfir-Dahav and Tennenholtz, 1996], and in particular, the merging of epistemic states [Borgida and Imielinski, 1984, Baral et al., 1991, 1992, Subrahmanian, 1994, Liberatore and Schaerf, 1998, Konieczny and Pino-Pérez, 1998]. Currently the major results in these areas seem to be focused on the level of the belief sets associated with epistemic states. An area that needs to be looked at is the establishment of a framework involving restrictions on the faithful total preorders associated with epistemic states.

The faithful total preorders have played a major role in many of the belief change operations described in this dissertation. As such it may be seen as a suitable point of departure for the description of appropriate semantic structures to be used in belief change. Two obvious generalisations of these orderings seem to be worthy of investigation: Firstly, the role of the faithful modular weak partial orders, both in the construction of refined entrenchment and systematic withdrawal, is an indication that one needs to move to a set of faithful preorders which includes both the faithful total preorders and the faithful modular weak partial orders. A candidate which seems to be appropriate is the set of faithful layered preorders. Encompassing both the faithful total preorders and the faithful modular weak partial orders, it retains an important characteristic shared by these two sets of preorders; the idea of layers of elements, with elements on different levels being comparable. It is, essentially, this property which ensures the satisfaction of the postulates (K-8) and (K*8) in the context of AGM theory change and the postulate known as Rational Monotonicity in the context of nonmonotonic reasoning. Secondly, while the faithful layered preorders may be sufficient for the

definition of some belief change operations such as revision and withdrawal, it has been pointed out by Rott [1991, p. 172], amongst others, that richer structures are needed for others. A proposal that immediately springs to mind is to use structures along the lines of Spohn's [1988, 1991] ordinal conditional functions. (Observe that we may view the faithful total preorders induced by infobases in this light, since these orderings are obtained from the *IB*-numbers of the interpretations.) Considerable progress has been made in this regard by Goldszmidt and Pearl [1996]. Amongst many other desirable properties, their formalism is able to deal with observations with a varying degree of firmness. They also provide a link with qualitative probabilities. It remains to be seen whether their approach can be combined with the use of faithful layered preorders, where elements on the same level (with the same ordinal assigned to them) need not be seen as comparable.

With the exception of the *IB*-induced faithful total preorders, the use of semantic structures for defining belief change has been of a declarative nature in this dissertation, with not much attention being paid to the equally important question of how to extract suitable semantic structures from the data structures at one's disposal in order to perform belief change. This question has received some attention in the nonmonotonic literature [Geffner and Pearl, 1992, Geffner, 1992, Delgrande and Schaub, 1997], and has also been addressed by Goldszmidt and Pearl [1996] in the context of belief change, but much work still needs to be done.

Finally, we come to an extremely important general aspect which has received no attention in this dissertation, and indeed, very little, in the research field of belief change; implementational considerations and the computational complexity of proposed belief change operations. While some researchers [Lehmann and Magidor, 1992, Gärdenfors and Rott, 1995, Goldszmidt and Pearl, 1996, Greiner, 1999] have presented relevant results, a more general picture has yet to emerge.

Appendix A

Proofs of some results in chapter 3

A.1 Theorems 3.2.3 and 3.3.1

Theorem 3.2.3

1. *A removal defined in terms of a semantic selection function using $(\text{Def} \sim \text{from } sm_K)$ is a basic AGM contraction. Conversely, every basic AGM contraction can be defined in terms of a semantic selection function using $(\text{Def} \sim \text{from } sm_K)$.*
2. *A revision defined in terms of a semantic selection function using $(\text{Def} * \text{from } sm_K)$ is a basic AGM revision. Conversely, every basic AGM revision can be defined in terms of a semantic selection function using $(\text{Def} * \text{from } sm_K)$.*

Proof 1. Let sm_K be a semantic selection function and let $-$ be defined in terms of sm_K using $(\text{Def} \sim \text{from } sm_K)$. We construct a selection function s_K such that $-$ is defined in terms of s_K using $(\text{Def} \sim \text{from } s_K)$. By theorem 2.2.4 it then follows that $-$ is a basic AGM contraction. Pick any $\alpha \in L$. If $\alpha \notin K$ or if $\models \alpha$ then $s_K(K \perp \alpha) = \{K\}$ and by definition $sm_K(\alpha) = \emptyset$, and thus $\cap s_K(K \perp \alpha) = Th(M(K) \cup sm_K(\alpha))$. So we suppose that $\not\models \alpha$ and $\alpha \in K$. Then $\emptyset \subset sm_K(\alpha) \subseteq M(\neg\alpha)$. By proposition 3.2.1 it follows that for every $u \in sm_K(\alpha)$, there is an $A_u \in K \perp \alpha$ such that $Th(M(K) \cup \{u\}) = A_u$. We let $s_K(K \perp \alpha)$ be the set consisting of all these A_u 's. That is,

$$s_K(K \perp \alpha) = \{A \in K \perp \alpha \mid \exists u \in sm_K(\alpha) \text{ such that } Th(M(K) \cup \{u\}) = A\}.$$

Since $sm_K(\alpha) \neq \emptyset$ it clearly follows that $s_K(K \perp \alpha) \neq \emptyset$. Furthermore, it is clear that $s_K(K \perp \alpha) \subseteq K \perp \alpha$. We still need to show that $\cap s_K(K \perp \alpha) = Th(M(K) \cup$

$sm_K(\alpha)$). By proposition 3.2.1 we have that, for every $u \in sm_K(\alpha)$, there is an $A \in s_K(K \perp \alpha)$ such that $Th(M(K) \cup \{u\}) = A$, and for every $A \in s_K(K \perp \alpha)$, there is a $u \in sm_K(\alpha)$ such that $A = Th(M(K) \cup \{u\})$, and therefore

$$\bigcap_{u \in sm_K(\alpha)} Th(M(K) \cup \{u\}) = \cap_{s_K(K \perp \alpha)}.$$

So it suffices to show that

$$\bigcap_{u \in sm_K(\alpha)} Th(M(K) \cup \{u\}) = Th(M(K) \cup sm_K(\alpha)).$$

Pick any $\beta \in \bigcap_{u \in sm_K(\alpha)} Th(M(K) \cup \{u\})$. Then $u \in M(\beta)$ for every $u \in sm_K(\alpha)$ and $v \in M(\beta)$ for every $v \in M(K)$. Therefore $M(K) \cup sm_K(\alpha) \subseteq M(\beta)$ and thus $\beta \in Th(M(K) \cup sm_K(\alpha))$. Conversely, suppose that $\beta \in Th(M(K) \cup sm_K(\alpha))$. Then $M(K) \cup sm_K(\alpha) \subseteq M(\beta)$ and therefore $M(K) \cup \{u\} \subseteq M(\beta)$ for every $u \in sm_K(\alpha)$. So $\beta \in Th(M(K) \cup \{u\})$ for every $u \in sm_K(\alpha)$, which means that $\beta \in \bigcap_{u \in sm_K(\alpha)} Th(M(K) \cup \{u\})$.

Conversely, pick any basic AGM contraction $-$. By theorem 2.2.4, there is a selection function s_K in terms of which $-$ is defined using (Def \sim from s_K). We construct a semantic selection function sm_K such that for every α , $\cap_{s_K(K \perp \alpha)} = Th(M(K) \cup sm_K(\alpha))$. Pick any $\alpha \in L$. The cases in which $\alpha \notin K$ or $\models \alpha$ have already been dealt with above, so suppose that $\not\models \alpha$ and $\alpha \in K$. Then $\emptyset \subset s_K(K \perp \alpha) \subseteq K \perp \alpha$. By proposition 3.2.1 it follows that for every $A \in s_K(K \perp \alpha)$, there is a $u_A \in M(\neg\alpha)$ such that $Th(M(K) \cup \{u_A\}) = A$. We let $sm_K(\alpha)$ be the set consisting of all these u_A 's. That is,

$$sm_K(\alpha) = \{u \in M(\neg\alpha) \mid \exists A \in s_K(K \perp \alpha) \text{ such that } Th(M(K) \cup \{u\}) = A\}.$$

Clearly $\emptyset \subset sm_K(\alpha) \subseteq M(\neg\alpha)$ and if $\alpha \equiv \beta$ then $sm_K(\alpha) = sm_K(\beta)$. To show that $\cap_{s_K(K \perp \alpha)} = Th(M(K) \cup sm_K(\alpha))$ we proceed exactly as above, which means we are done.

2. By theorem 2.1.6 and part (1) above, it suffices to show that the revision $*$, defined in terms of a semantic selection function sm_K using (Def $*$ from sm_K), can also be defined in terms of $-$ using (Def $*$ from \sim), where $-$ is the removal defined in terms of sm_K using (Def \sim from sm_K). So, ignoring the trivial cases,

observe that if $\neg\alpha \in K$ and $\not\models \neg\alpha$, then

$$\begin{aligned} & (K - \neg\alpha) + \alpha \\ &= Th(M(K) \cup sm_K(\neg\alpha)) + \alpha \\ &= Th(sm_K(\neg\alpha)) \text{ by lemma 1.3.4.} \end{aligned}$$

□

Theorem 3.3.1

1. *Every faithful total preorder defines a GE-ordering using $(Def \sqsubseteq_G \text{ from } \preceq)$. Conversely, every GE-ordering can be defined in terms of a faithful total preorder using $(Def \sqsubseteq_G \text{ from } \preceq)$.¹*
2. *Every faithful total preorder defines an EE-ordering using $(Def \sqsubseteq_E \text{ from } \preceq)$. Conversely, every EE-ordering can be defined in terms of a faithful total preorder using $(Def \sqsubseteq_E \text{ from } \preceq)$.*

Proof 1. Let \preceq be any faithful total preorder. We show that the relation \sqsubseteq_{GE} on L , defined in terms of \preceq using $(Def \sqsubseteq_G \text{ from } \preceq)$ is a GE-ordering. For (GE1), pick any $\alpha, \beta \in L$ and suppose that $\beta \not\sqsubseteq_{GE} \alpha$. That is, there is a $y \in M(\alpha)$ such that $x \not\preceq y$ (and thus $y \preceq x$) for every $x \in M(\beta)$, and so $\alpha \sqsubseteq_{GE} \beta$. For (GE2), suppose that $\alpha \sqsubseteq_{GE} \beta$ and $\beta \sqsubseteq_{GE} \gamma$, and pick any $y \in M(\gamma)$. There is an $x \in M(\beta)$ such that $x \preceq y$, and a $z \in M(\alpha)$ such that $z \preceq x$. By the transitivity of \preceq , $z \preceq y$. For (GE3), suppose that $\alpha \models \beta \vee \gamma$ and assume that $\beta \not\sqsubseteq_{GE} \alpha$ and $\gamma \not\sqsubseteq_{GE} \alpha$. So there is a $y \in M(\alpha)$ such that $y \prec x$ for every $x \in M(\beta)$, and there is a $v \in M(\alpha)$ such that $v \prec u$ for every $u \in M(\gamma)$. Clearly $y \notin M(\beta)$ and $v \notin M(\gamma)$. If $y \preceq v$ then $y \notin M(\gamma)$, which contradicts the fact that $\alpha \models \beta \vee \gamma$. Similarly, if $v \preceq y$ then $v \notin M(\beta)$, contradicting the fact that $\alpha \models \beta \vee \gamma$. For (GE4), suppose that $K \neq Cn(\perp)$ and pick an $\alpha \in L$ such that $\neg\alpha \notin K$. Then $M(K) \cap M(\alpha) \neq \emptyset$, and it thus follows that $x \preceq y$ for every $\beta \in L$ and every $y \in M(\beta)$; i.e. $\alpha \sqsubseteq_{GE} \beta$ for every $\beta \in L$. Conversely, suppose that $\alpha \sqsubseteq_{GE} \beta$ for every $\beta \in L$. In particular then, $\alpha \sqsubseteq_{GE} \top$; i.e. for every $y \in U$ there is an $x \in M(\alpha)$ such that $x \preceq y$. Because $K \neq L$, this means that $M(K) \cap M(\alpha) \neq \emptyset$,

¹This result fixes up some small inaccuracies of Grove [1988] and Boutilier [1992], and it sharpens a result of Boutilier [1994].

from which it follows that $\neg\alpha \notin K$. For (GE5), suppose that $\models \neg\alpha$. So $M(\alpha) = \emptyset$, and it thus follows vacuously that $\beta \sqsubseteq_{GE} \gamma$ for every $\beta \in L$. On the other hand, suppose that $\beta \sqsubseteq_{GE} \alpha$ for every $\beta \in L$. Then, in particular, $\perp \sqsubseteq_{GE} \alpha$, which has to mean that $M(\alpha) = \emptyset$, and thus that $\models \neg\alpha$.

For the converse, let \sqsubseteq_{GE} be a GE-ordering. We construct a faithful total preorder \preceq in terms of which \sqsubseteq_{GE} can be defined using (Def \sqsubseteq_G from \preceq). For any $\alpha \in L$, let $\Delta\alpha = \{\beta \in L \mid \alpha \sqsubseteq_{GE} \beta\}$. Grove [1988] refers to these sets as *cuts*. It is easy to see that the set of cuts is totally ordered under set inclusion. Pick any two wffs α and β , and suppose that $\Delta\alpha \not\subseteq \Delta\beta$. Then $\Delta\alpha \setminus \Delta\beta = \{\gamma \mid \alpha \sqsubseteq_{GE} \gamma \not\sqsubseteq_{GE} \beta\} \neq \emptyset$, and so it follows that $\Delta\beta \subseteq \Delta\alpha$. Now, for every $x \in U$, let

$$\sqcup x = \bigcup \{\Delta\alpha \mid \alpha \in L \text{ and } x \in M(\neg(\Delta\alpha))\}.$$

So $\sqcup x$ is the largest cut that contains none of the wffs satisfied by x . We define \preceq as follows:

$$\text{For every } x, y \in U, x \preceq y \text{ iff } \sqcup x \supseteq \sqcup y.$$

First we show that \preceq is a faithful total preorder. Cuts are totally ordered by set inclusion, so it clearly follows that \preceq is a total preorder. Now pick any $x, y \in M(K)$. By (GE4), $\sqcup x = \sqcup y = \{\alpha \mid \neg\alpha \in K\}$ and so $x \preceq y$. Furthermore, if we pick a $z \notin M(K)$, then there is at least one $\neg\alpha \in K$ such that $z \models \alpha$. So $\sqcup x \supset \sqcup z$, which means that $x \prec z$. To prove that \preceq is smooth, it suffices to show that for every $\alpha \in L$ such that $\not\models \neg\alpha$, $\text{Min}_{\preceq}(\alpha) \neq \emptyset$. The following result shows that there is an interesting connection between a cut C and the set $M(\neg C)$ of all interpretations that satisfy none of the wffs in C .

$$\text{For every cut } C \text{ and every } \alpha \in L, \alpha \in C \text{ iff } M(\neg C) \subseteq M(\neg\alpha). \quad (\text{A.1})$$

For the proof of (A.1), pick any cut C and any $\alpha \in L$ and suppose that $\alpha \in C$. Then, by definition, $M(\neg C) \subseteq M(\neg\alpha)$. Conversely, suppose $\alpha \notin C$. If $\neg C \not\models \neg\alpha$, then there is a model of $\neg C$ that satisfies α , and thus $M(\neg C) \not\subseteq M(\neg\alpha)$. So suppose that $\neg C \models \neg\alpha$. By compactness, there is a finite subset C_{Fin} of C such that $\neg C_{Fin} \models \neg\alpha$, and so $\alpha \models \bigvee C_{Fin}$. By repeated applications of (GE3), it then follows that for some $\beta \in C_{Fin}$, $\beta \sqsubseteq_{GE} \alpha$, and thus $\alpha \in C$, contradicting the supposition.

Now, pick any $\alpha \in L$ such that $\not\models \neg\alpha$, and let $C_\alpha = \bigcup \{\Delta\beta \mid \alpha \notin \Delta\beta\}$. So C_α is the largest cut not containing α . From (A.1) it follows that $M(\neg C_\alpha)$ contains

an interpretation y that satisfies α . We show that $C_\alpha = \sqcup y$. If $C_\alpha \not\subseteq \sqcup y$, then $\sqcup y \subset C_\alpha$, which means there has to be a $\beta \in C_\alpha$ that is satisfied by y , contradicting (A.1). So $C_\alpha \subseteq \sqcup y$. Conversely, since C_α is the largest cut not containing α , and since $\alpha \notin \sqcup y$, it follows that $\sqcup y \subseteq C_\alpha$. Now assume there is an $x \in M(\alpha)$, and thus $\alpha \notin \sqcup x$, such that $x \prec y$. Then $\sqcup x \supset \sqcup y = C_\alpha$, contradicting the fact that C_α is the largest cut not containing α . Therefore $y \in \text{Min}_\preceq(\alpha)$.

Finally, let \sqsubseteq be the GE-ordering defined in terms of \preceq using (Def \sqsubseteq_G from \preceq). We show that $\sqsubseteq = \sqsubseteq_{GE}$. Pick any $\alpha, \beta \in L$. If $\models \neg\beta$ then by the definition of \sqsubseteq , $\alpha \sqsubseteq \beta$ and, by (GE5), $\alpha \sqsubseteq_{GE} \beta$. Furthermore, if $\models \neg\alpha$ and $\alpha \sqsubseteq_{GE} \beta$, then by (GE5), $\models \neg\beta$, and if $\models \neg\alpha$ and $\alpha \sqsubseteq \beta$, it follows from the definition of \sqsubseteq that $\models \neg\beta$. Hence, if $\models \neg\alpha$ or $\models \neg\beta$, then $\alpha \sqsubseteq_{GE} \beta$ iff $\alpha \sqsubseteq \beta$. So we suppose that $\not\models \neg\gamma$ and $\not\models \neg\delta$. By (A.1), there is a $y \in M(\neg C_\beta)$ that satisfies β , and an $x \in M(\neg C_\alpha)$ that satisfies α . As above, it then also follows that $C_\beta = \sqcup y$, $C_\alpha = \sqcup x$, $y \in \text{Min}_\preceq(\beta)$ and that $x \in \text{Min}_\preceq(\alpha)$. If $\alpha \sqsubseteq_{GE} \beta$, then $C_\beta \subseteq C_\alpha$, and thus $\sqcup y \subseteq \sqcup x$. So, by the definition of \preceq , $x \preceq y$, and therefore $\alpha \sqsubseteq \beta$. On the other hand, if $\alpha \sqsubseteq \beta$, it means that $u \preceq v$ for every $u \in \text{Min}_\preceq(\alpha)$, and every $v \in \text{Min}_\preceq(\beta)$. So in particular, $x \preceq y$, which means that $C_\beta = \sqcup y \subseteq \sqcup x = C_\alpha$, and thus that $\alpha \sqsubseteq_{GE} \beta$.

2. Follows from part (1) and theorem 2.3.5.

□

A.2 Results used in the proof of theorem 3.2.6

This section contains the results used to prove that AGM contraction and AGM revision can be characterised in terms of faithful total preorders. First we provide a “soundness” result for AGM contraction.

Proposition A.2.1 *Every removal defined in terms of a faithful total preorder \preceq using (Def \sim from \preceq) is an AGM contraction.*

Proof For (K-1) to (K-6), it suffices, by theorem 3.2.3, to show that the function $sm_K : L \rightarrow \wp U$ obtained by setting $sm_K(\alpha) = \text{Min}_\preceq(\neg\alpha) \setminus M(K)$ is a semantic selection function. If $\alpha \equiv \beta$ then $\text{Min}_\preceq(\neg\alpha) = \text{Min}_\preceq(\neg\beta)$ and so $sm_K(\alpha) =$

$sm_K(\beta)$. Furthermore, if $\alpha \notin K$ then there is an $x \in M(K)$ such that $x \notin M(\alpha)$. So $Min_{\preceq}(\neg\alpha) \subseteq M(K)$ and thus $sm_K(\alpha) = \emptyset$. On the other hand, if $\models \alpha$ then $Min_{\preceq}(\neg\alpha) = \emptyset$ and thus $sm_K(\alpha) = \emptyset$. So suppose that $\alpha \in K$ and $\not\models \alpha$. Then $M(K) \cap Min_{\preceq}(\neg\alpha) = \emptyset$, but by smoothness, $Min_{\preceq}(\neg\alpha) \neq \emptyset$. So $\emptyset \subset sm_K(\alpha)$. Finally, $sm_K(\alpha) \subseteq M(\neg\alpha)$ since $Min_{\preceq}(\neg\alpha) \subseteq M(\neg\alpha)$.

For (K-7), suppose that $\gamma \in K - \alpha$ and $\gamma \in K - \beta$. That is,

$$\begin{aligned} M(Th(M(K) \cup Min_{\preceq}(\neg\alpha))) &\subseteq M(\gamma) \text{ and} \\ M(Th(M(K) \cup Min_{\preceq}(\neg\beta))) &\subseteq M(\gamma). \end{aligned}$$

If we can show that $M(K) \cup Min_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\gamma)$, it follows that $\gamma \in K - \alpha \wedge \beta$, which means we are done. We already have that $M(K) \subseteq M(\gamma)$, so it remains to be shown that $Min_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\gamma)$. Pick a $u \in Min_{\preceq}(\neg(\alpha \wedge \beta))$. It follows that either $u \in M(\neg\alpha)$ or $u \in M(\neg\beta)$. In the latter case, $u \in Min_{\preceq}(\neg\beta)$ and thus $x \in M(\gamma)$ since $Min_{\preceq}(\neg\beta) \subseteq M(\gamma)$. A similar argument holds in the former case.

For (K-8), suppose that $\beta \notin K - (\alpha \wedge \beta)$. If $\alpha \wedge \beta \notin K$, then $K - (\alpha \wedge \beta) = K$ by (K-3), and thus also $K = K - \beta$ (because $\beta \notin K - (\alpha \wedge \beta) = K$), from which the result follows. So we suppose that $\alpha \wedge \beta \in K$. Because $\beta \notin K - (\alpha \wedge \beta)$, there is a $u \in M(K) \cup Min_{\preceq}(\neg(\alpha \wedge \beta))$ such that $u \not\models \beta$. But $\alpha \wedge \beta \in K$, and so $u \notin M(K)$, which means that $u \in Min_{\preceq}(\neg(\alpha \wedge \beta))$. Now, pick a $\gamma \in K - (\alpha \wedge \beta)$, i.e. $M(K) \cup Min_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\gamma)$. We have to show that $M(K) \cup Min_{\preceq}(\neg\beta) \subseteq M(\gamma)$. We already have that $M(K) \subseteq M(\alpha)$. To show that $Min_{\preceq}(\neg\beta) \subseteq M(\gamma)$, pick a $v \in Min_{\preceq}(\neg\beta)$ and assume that $v \notin M(\gamma)$. Because $Min_{\preceq}(\neg(\alpha \wedge \beta)) \subseteq M(\gamma)$, it follows that $v \notin Min_{\preceq}(\neg(\alpha \wedge \beta))$. But then $u \prec v$ since $u \in Min_{\preceq}(\neg(\alpha \wedge \beta))$, contradicting the minimality of v in $M(\neg\beta)$. \square

To prove the “completeness” result for AGM contraction, we construct an appropriate faithful total preorder.

Definition A.2.2 Let $-$ be an removal, and let

$$Min_K = \{u \notin M(K) \mid u \in M(K - \alpha) \text{ for some } \alpha\}.$$

The *canonical relation* for $-$ is the binary relation \preceq on U containing just the ordered pairs sanctioned by conditions 1 to 5 below.

1. For every $u, v \in M(K)$, $u \preceq v$.

2. For every $u, v \in U \setminus (M(K) \cup \text{Min}_K)$, $u \preceq v$.
3. For every $u \in M(K)$ and $v \notin M(K)$, $u \prec v$.
4. For every $u \in \text{Min}_K$ and $v \in U \setminus (M(K) \cup \text{Min}_K)$, $u \prec v$.
5. For every $u, v \in \text{Min}_K$, $u \preceq v$ iff for every $\alpha \in K$, $v \in M(K - \alpha)$ and $u \in M(\neg\alpha)$ implies $u \in M(K - \alpha)$.

□

As we shall see below, the canonical relation for an AGM contraction $-$ is a total preorder on U with

- the models of K as the minimal elements,
- the elements that are neither models of K nor of some belief set obtained from K via $-$, as the maximal elements, and
- the rest of the elements of U in between.

We also need the following technical lemmas.

Lemma A.2.3 [Alchourrón et al., 1985] *If $-$ is a basic AGM contraction then the following is equivalent to (K-7): $(K - \alpha) \cap \text{Cn}(\alpha) \subseteq K - (\alpha \wedge \beta)$.*

Lemma A.2.4 *If \sim is a removal satisfying (K-1), (K-4), (K-6) and (K-8), then either $K \sim (\alpha \wedge \beta) \subseteq K \sim \alpha$ or $K \sim (\alpha \wedge \beta) \subseteq K \sim \beta$ for every $\alpha, \beta \in L$.*

Proof If $\models \alpha \wedge \beta$ then $\models \alpha$ and $\models \beta$, and by (K-6), $K \sim \alpha = K \sim \beta = K \sim (\alpha \wedge \beta) = K$, from which the result follows. So suppose that $\not\models \alpha \wedge \beta$. By (K-4), $\alpha \wedge \beta \notin K - (\alpha \wedge \beta)$, and so by (K-1), either $\alpha \notin K \sim (\alpha \wedge \beta)$ or $\beta \notin K \sim (\alpha \wedge \beta)$. The result then follows directly from (K-8). □

Lemma A.2.5 *Let $-$ be a basic AGM contraction. If $\alpha \in K$, $u \notin M(K)$ and $u \in M(K - \alpha)$ then $u \in M(\neg\alpha)$.*

Proof Suppose $\alpha \in K$, $u \notin M(K)$ and $u \in M(K - \alpha)$. By (K-6), $M((K - \alpha) + \alpha) = M(K)$, and because $M(K - \alpha) \cap M(\alpha) \subseteq M((K - \alpha) + \alpha)$, we have that $u \notin M(K - \alpha) \cap M(\alpha)$ and therefore $u \in M(\neg\alpha)$. □

Lemma A.2.6 *Let $-$ be an AGM contraction, and let \preceq be the canonical relation for $-$. If $\alpha \in K \setminus Cn(\top)$ and $u \in M(K - \alpha)$, then $u \preceq v$ for every $v \in M(\neg\alpha)$.*

Proof If $u \in M(K)$ then $u \preceq v$ for every $v \in U$, so suppose that $u \notin M(K)$. Because $u \in M(K - \alpha)$, $u \in Min_K$. Pick any $v \in M(\neg\alpha)$. Because $\alpha \in K$, $v \in Min_K$ or $v \in U \setminus (M(K) \cup Min_K)$. In the latter case the result follows from the definition of \preceq . For the former case, pick any $\beta \in K$ such that $u \in M(\neg\beta)$ and suppose that $v \in M(K - \beta)$. By lemma A.2.5, $v \in M(\neg\beta)$. We need to show that $u \in M(K - \beta)$. By lemma A.2.4, either $M(K - \alpha) \subseteq M(K - (\alpha \wedge \beta))$ or $M(K - \beta) \subseteq M(K - (\alpha \wedge \beta))$. In the former case, $\alpha \notin K - (\alpha \wedge \beta)$ by (K-4) and in the latter case $v \in M(K - (\alpha \wedge \beta))$, and because $v \in M(\neg\alpha)$, $\alpha \notin K - (\alpha \wedge \beta)$. So in either case $\alpha \notin K - (\alpha \wedge \beta)$ and thus, by (K-8), $K - (\alpha \wedge \beta) \subseteq K - \alpha$. Now assume that $u \notin M(K - \beta)$. Then there is a $\gamma \in K - \beta$, and thus $\beta \vee \gamma \in K - \beta$, such that $u \notin M(\gamma)$, which means, by lemma A.2.3, that $(\beta \vee \gamma) \in K - (\alpha \wedge \beta) \subseteq K - \alpha$. And because $u \in M(K - \alpha)$, we have that $u \Vdash \beta \vee \gamma$, contradicting the fact that $u \in M(\neg\beta)$ and $u \notin M(\gamma)$. \square

Lemma A.2.7 *The canonical relation \preceq for an AGM contraction $-$ is a total preorder.*

Proof It suffices to consider only interpretations in Min_K . For reflexivity, note that for every $\alpha \in K$, if $x \in M(K - \alpha)$ and $x \in M(\neg\alpha)$ then $x \in M(K - \alpha)$. For transitivity, pick any $x, y, z \in Min_K$ and suppose that $x \preceq y$ and $y \preceq z$. We need to show that $x \preceq z$. Pick a $\alpha \in K$ and suppose that $z \in M(K - \alpha)$ and $x \in M(\neg\alpha)$. By lemma A.2.5, $z \in M(\neg\alpha)$. We show that $x \in M(K - \alpha)$. Because $y \in Min_K$, there is a $\gamma \in K$ such that $y \in M(K - \gamma)$ and therefore, by lemma A.2.5, $y \in M(\neg\gamma)$. Furthermore, because $\gamma \in K$ and $\alpha \in K$ we have $\gamma \wedge \alpha \in K$. And because $M(\neg(\gamma \wedge \alpha)) = M(\neg\gamma) \cup M(\neg\alpha)$ it follows that $x, y, z \in M(\neg(\gamma \wedge \alpha))$. By lemma A.2.4, $M(K - \alpha) \subseteq M(K - \gamma \wedge \alpha)$ or $M(K - \gamma) \subseteq M(K - \gamma \wedge \alpha)$. In the former case, $z \in M(K - \gamma \wedge \alpha)$ because $z \in M(K - \alpha)$. So because $y \in M(\neg(\gamma \wedge \alpha))$, $\gamma \wedge \alpha \in K$ and $y \preceq z$, we have $y \in M(K - \gamma \wedge \alpha)$. In similar fashion, because $x \in M(\neg(\gamma \wedge \alpha))$ and $x \preceq y$, $x \in M(K - \gamma \wedge \alpha)$. In the latter case, because $y \in M(K - \gamma)$, we have $y \in M(K - \gamma \wedge \alpha)$ and then as before, $x \in M(K - \gamma \wedge \alpha)$. So either way $x \in M(K - \gamma \wedge \alpha)$. To show that $x \in M(K - \alpha)$, pick any $\beta \in K - \alpha$. We show that $x \in M(\beta)$. Because $\beta \in K - \alpha$, we also have that $\beta \vee \alpha \in K - \alpha$, so by lemma A.2.3, $\beta \vee \alpha \in K - \gamma \wedge \alpha$. So $x \Vdash \beta \vee \alpha$. But because $x \in M(\neg\alpha)$, we have that $x \in M(\beta)$.

To show that \preceq is a total preorder we still need to show that for every $x, y \in Min_K$, $x \preceq y$ or $y \preceq x$. Pick any $x, y \in Min_K$, and suppose that $x \not\preceq y$. Then there is an

$\alpha \in K$ such that $y \in M(K - \alpha)$ and $x \in M(\neg\alpha)$, but $x \notin M(K - \alpha)$. By lemma A.2.6 it then follows that $y \preceq x$. \square

We are now ready to prove the “completeness result” for AGM contraction.

Proposition A.2.8 *Every AGM contraction – can be defined in terms of a faithful total preorder using $(\text{Def} \sim \text{from } \preceq)$.*

Proof We show that the canonical relation \preceq for – is such a faithful total preorder. By lemma A.2.7, \preceq is a total preorder. To show that \preceq is faithful, we need only show that \preceq is smooth; the other conditions for faithfulness follow directly from the definition of \preceq . So pick any α . If $\neg\alpha \notin K$, the $M(\alpha)$ -smoothness of \preceq follows directly from the definition of \preceq , and if $\models \neg\alpha$ then $M(\alpha) = \emptyset$, and thus \preceq is $M(\alpha)$ -smooth. So suppose that $\neg\alpha \in K$, $\not\models \neg\alpha$ and pick any $y \in M(\alpha)$. We need to show that there is an x that is \preceq -minimal in $M(\alpha)$ such that $x \preceq y$. Because $\not\models \neg\alpha$ it follows from (K-2) and (K-4) that $M(K) \subset M(K - \neg\alpha)$. So there is an $x \in M(K - \neg\alpha)$ such that $x \notin M(K)$. By lemma A.2.5, $x \in M(\alpha)$ and by lemma A.2.6, $x \preceq y$ for every $y \in M(\alpha)$.

To show that – can be defined in terms of \preceq using $(\text{Def} \sim \text{from } \preceq)$, it suffices to show that $M(K - \alpha) = M(K) \cup \text{Min}_{\preceq}(\neg\alpha)$ for every $\alpha \in L$. Clearly, if $\models \alpha$ or $\alpha \notin K$, then $M(K) \cup \text{Min}_{\preceq}(\neg\alpha) = M(K)$, so we need only consider the case where $\alpha \in K \setminus \text{Cn}(\top)$. For the left-to-right inclusion, pick any $x \in M(K - \alpha)$. If $x \in M(K)$ then clearly $x \in M(K) \cup \text{Min}_{\preceq}(\neg\alpha)$, so suppose that $x \notin M(K)$. By lemma A.2.5, $x \in M(\neg\alpha)$. By lemma A.2.6 it follows that for every $y \in M(\neg\alpha)$, $x \preceq y$. So $x \in \text{Min}_{\preceq}(\neg\alpha)$ and thus $x \in M(K) \cup \text{Min}_{\preceq}(\neg\alpha)$.

For the right-to-left inclusion, note firstly that by (K-2), $M(K) \subseteq M(K - \alpha)$. Now pick any $x \in \text{Min}_{\preceq}(\neg\alpha)$. We need to show that $x \in M(K - \alpha)$. Because $\not\models \alpha$, it follows from (K-4) that $M(K - \alpha) \cap M(\neg\alpha) \neq \emptyset$. So pick any $y \in M(K - \alpha) \cap M(\neg\alpha)$. By lemma A.2.6, $y \preceq x$, and then $x \preceq y$ because $x \in \text{Min}_{\preceq}(\neg\alpha)$. Now, $x \notin M(K)$ because $\alpha \in K$ and $x \in M(\neg\alpha)$, and so $x \notin U \setminus (M(K) \cup \text{Min}_K)$ because $y \in \text{Min}_K$ and $x \preceq y$. So $x \in \text{Min}_K$ and therefore, from part (5) of definition A.2.2, it follows that $x \in M(K - \alpha)$. \square

We thus obtain the following representation theorem.

Theorem 3.2.6

1. *Every faithful total preorder defines an AGM contraction using $(Def \sim \text{from } \preceq)$.
Conversely, every AGM contraction can be defined in terms of a faithful total preorder using $(Def \sim \text{from } \preceq)$.*
2. *Every faithful total preorder defines an AGM revision using $(Def * \text{from } \preceq)$.
Conversely, every AGM revision can be defined in terms of a faithful total preorder using $(Def * \text{from } \preceq)$.*

Proof 1. The proof follows directly from propositions A.2.1 and A.2.8.

2. Follows from theorem 2.1.6, part (1) above, and proposition 3.2.8.

□

Appendix B

Proofs of some results in chapter 6

B.1 Results used in the proof of theorem 6.3.4

Proposition B.1.1 *Let \leq be a faithful modular weak partial order. The function $ss_K : L \rightarrow \wp U$, defined as: $ss_K(\alpha) = \nabla_{\leq}(\neg\alpha)$, is a saturatable selection function.*

Proof It follows trivially that $\alpha \equiv \beta$ implies $ss_K(\alpha) = ss_K(\beta)$, and that $\models \alpha$ implies $ss_K(\alpha) = \emptyset$. Now suppose that $\alpha \notin K$. Then $\not\models \alpha$, and it follows easily that $ss_K(\alpha) = \nabla_{\leq}(\neg\alpha) \subseteq \text{Min}_{\leq}(\neg\alpha) \subseteq M(K)$. Finally, suppose that $\alpha \in K$ and $\not\models \alpha$. By smoothness, $\emptyset \subset \text{Min}_{\leq}(\alpha) \subseteq \nabla_{\leq}(\neg\alpha)$, and so $ss_K(\alpha) \cap M(\neg\alpha) \neq \emptyset$. \square

Proposition B.1.2 *Every systematic withdrawal satisfies $(K \div 1)$ to $(K \div 10)$.*

Proof Let \div be a systematic withdrawal, and let \leq be a faithful modular weak partial order from which \div is obtained using (Def \sim from ∇_{\leq}). By proposition B.1.1 and definition 6.2.5, \div is a saturatable withdrawal and by theorem 6.2.6, it thus satisfies $(K \div 1)$ to $(K \div 6)$. For $(K \div 7)$, suppose that $\gamma \in K \div (\alpha \wedge \gamma)$. We only consider the case where $\not\models \alpha$ and $\alpha \wedge \gamma \in K$. Then $\gamma \in K$ by $(K \div 2)$, $\text{Min}_{\leq}(\neg(\alpha \wedge \gamma)) \subseteq M(\neg\alpha) \cap M(\gamma)$ and $\nabla_{\leq}(\neg\alpha) \subseteq M(\gamma)$. Now pick any $x \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma))$ and any $y \in \text{Min}_{\leq}(\neg(\alpha \wedge \gamma))$. (By smoothness, neither $\text{Min}_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma))$ nor $\text{Min}_{\leq}(\neg(\alpha \wedge \gamma))$ is empty.) It is clear that $y \not\leq x$. So $\nabla_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma)) \setminus \text{Min}_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma)) \subseteq M(\gamma)$. To show that $\gamma \in K \div (\alpha \wedge \beta \wedge \gamma)$, it thus remains to show that $\text{Min}_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma)) \subseteq M(\gamma)$. And if this were not the case, there would be a $z \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta \wedge \gamma))$ such that $z \in M(\neg\gamma)$. But then $z \in \text{Min}_{\leq}(\neg(\alpha \wedge \gamma))$, thus contradicting $\text{Min}_{\leq}(\neg(\alpha \wedge \gamma)) \subseteq M(\neg\alpha) \cap M(\gamma)$.

For $(K \div 8)$, suppose that $\beta \notin K \div (\alpha \wedge \beta)$. We have to show that $K \div (\alpha \wedge \beta) \subseteq K \div \beta$. If $\alpha \wedge \beta \notin K$, then by $(K \div 3)$, $K \div (\alpha \wedge \beta) = K$, and thus also $K = K \div \beta$ (because $\beta \notin K \div (\alpha \wedge \beta) = K$), from which the result follows. So we suppose that $\alpha \wedge \beta \in K$. Now, pick an $\alpha \in K \div (\alpha \wedge \beta)$. Then $M(K) \cup \nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$ and so $M(K) \subseteq M(\alpha)$ and $\nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$. We have to show that $M(K) \cup \nabla_{\leq}(\neg\beta) \subseteq M(\alpha)$. We already have that $M(K) \subseteq M(\alpha)$. To show that $\nabla_{\leq}(\neg\beta) \subseteq M(\alpha)$, it suffices to show that $\nabla_{\leq}(\neg\beta) \subseteq \nabla_{\leq}(\neg(\alpha \wedge \beta))$. If we can show that $\text{Min}_{\leq}(\neg\beta) \subseteq \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$, it immediately follows from (Def ∇_{\leq}) that $\nabla_{\leq}(\neg\beta) \subseteq \nabla_{\leq}(\neg(\alpha \wedge \beta))$. So pick any $y \in \text{Min}_{\leq}(\neg\beta)$ and assume that $y \notin \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$. Since $y \in M(\neg(\alpha \wedge \beta))$, it follows by the smoothness of \leq that there is an $x \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ such that $x < y$. Because $y \in \text{Min}_{\leq}(\neg\beta)$, it must be the case that $x \in M(\neg\alpha \wedge \beta)$, and since \leq is a modular weak partial order it then also follows that $\text{Min}_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\beta)$. Moreover, since $y \in \text{Min}_{\leq}(\neg\beta)$ and since $x < y$ it has to be the case that for every $v \in \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ and every $u \leq v$, $u \in M(\beta)$. But then $\nabla_{\leq}(\neg(\alpha \wedge \beta)) \in M(\beta)$, contradicting the supposition that $\beta \notin K \div (\alpha \wedge \beta)$. For $(K \div 9)$, suppose that $\alpha \in K$, $\alpha \vee \beta \in K \div \alpha$ and $\beta \notin K \div \alpha$. We only consider the case where $\not\models \alpha$. Then $\text{Min}_{\leq}(\neg\alpha) \subseteq M(\beta)$, $\nabla_{\leq}(\neg\alpha) \setminus \text{Min}_{\leq}(\neg\alpha) \subseteq M(\alpha)$, and $\nabla_{\leq}(\neg\alpha) \setminus \text{Min}_{\leq}(\neg\alpha) \not\subseteq M(\beta)$. So $\text{Min}_{\leq}(\neg(\alpha \wedge \beta)) < \text{Min}_{\leq}(\neg\alpha)$, and therefore $\nabla_{\leq}(\neg(\alpha \wedge \beta)) \subseteq M(\alpha)$, from which it follows that $\alpha \in K \div (\alpha \wedge \beta)$.¹ For $(K \div 10)$, suppose that $\not\models \alpha$ and $\beta \in K \div \alpha$. Then $\nabla_{\leq}(\neg\alpha) \subseteq M(\beta)$. Therefore $\text{Min}_{\leq}(\neg\alpha) \subseteq \text{Min}_{\leq}(\neg(\alpha \wedge \beta))$ and thus $\alpha \notin K \div (\alpha \wedge \beta)$. \square

Lemma B.1.3 *If $\not\models \alpha$ and \sim is a removal that satisfies $(K \div 1)$, $(K \div 4)$, $(K \div 5)$, $(K \div 7)$ and $(K \div 8)$, then $\{\beta \mid \beta \in K \sim (\alpha \wedge \beta)\} = \bigcap \{K \sim (\alpha \wedge \beta) \mid \beta \in L\}$.*

Proof Suppose $\beta \in K \sim (\alpha \wedge \beta)$. Now pick any γ . By $(K \div 7)$, $\beta \in K \sim (\alpha \wedge \gamma \wedge \beta)$, and by $(K \div 4)$, $\alpha \wedge \gamma \wedge \beta \notin K \sim (\alpha \wedge \gamma \wedge \beta)$. Therefore $\alpha \wedge \gamma \notin K \sim (\alpha \wedge \gamma \wedge \beta)$ by $(K \div 1)$, and so $K \sim (\alpha \wedge \gamma \wedge \beta) \subseteq K \sim (\alpha \wedge \gamma)$ by $(K \div 5)$ and $(K \div 8)$, from which it follows that $\beta \in K \sim (\alpha \wedge \gamma)$. So we have shown that $\{\beta \mid \beta \in K \sim (\alpha \wedge \beta)\} \subseteq \bigcap \{K \sim (\alpha \wedge \beta) \mid \beta \in L\}$. The converse is trivial. \square

Lemma B.1.4 *If \sim satisfies $(K \div 1)$ to $(K \div 10)$, the withdrawal $\ddot{-}$ defined in terms of \sim using (Def $\ddot{-}$ from \sim) is a severe withdrawal.*

¹See section 1.3 for an explanation of the convention of applying $<$ to sets of interpretations.

Proof (K $\ddot{-}$ 1) to (K $\ddot{-}$ 6) follow easily from lemma B.1.3, and (K $\ddot{-}$ 7) follows easily from (K \div 7). For (K $\ddot{-}$ 8), suppose that $\beta \notin K \ddot{-}(\alpha \wedge \beta)$. If $\not\models \alpha \wedge \beta$ then $\beta \notin K \sim (\alpha \wedge \beta \wedge \beta) = K \sim (\alpha \wedge \beta)$. And if $\models \alpha \wedge \beta$ then $\beta \notin K$, and thus $\beta \notin K \sim (\alpha \wedge \beta)$ by (K \div 2). So in either case, $K \sim (\alpha \wedge \beta) \subseteq K \sim \beta$ by (K \div 8). We need to show that $K \ddot{-}(\alpha \wedge \beta) \subseteq K \ddot{-}\beta$. The case where $\models \alpha \wedge \beta$ is trivial, and so we suppose that $\not\models \alpha \wedge \beta$. We only consider the case where $\not\models \beta$. We need to show that $\{\gamma \mid \gamma \in K \sim (\alpha \wedge \beta \wedge \gamma)\} \subseteq \{\gamma \mid K \sim (\beta \wedge \gamma)\}$. Suppose that $\gamma \in K \sim (\alpha \wedge \beta \wedge \gamma)$. If $\beta \wedge \gamma \notin K \sim (\alpha \wedge \beta \wedge \gamma)$, then $K \sim (\alpha \wedge \beta \wedge \gamma) \subseteq K \sim (\beta \wedge \gamma)$ by (K \div 8), and so $\gamma \in K \sim (\beta \wedge \gamma)$. So we consider the case where $\beta \wedge \gamma \in K \sim (\alpha \wedge \beta \wedge \gamma)$. Since $\gamma \in K \sim (\alpha \wedge \beta \wedge \gamma)$, it follows from (K \div 4) that $\alpha \wedge \beta \notin K \sim (\alpha \wedge \beta \wedge \gamma)$, and then by (K \div 8) that $K \sim (\alpha \wedge \beta \wedge \gamma) \subseteq K \sim (\alpha \wedge \beta)$. Because $K \sim (\alpha \wedge \beta) \subseteq K \sim \beta$ we then have that $\beta \wedge \gamma \in K \sim \beta$, and therefore $\beta \in K \sim \beta$, contradicting $\not\models \beta$ and (K \div 4). \square

Lemma B.1.5 *Let \sim be a withdrawal satisfying (K \div 1) to (K \div 10). Now define the removal $\ddot{-}$ in terms of \sim using (Def $\ddot{-}$ from \sim), and define the removal \div in terms of $\ddot{-}$ using (Def \div from $\ddot{-}$). Then \sim and \div are identical.*

Proof By combining (Def $\ddot{-}$ from \sim) and (Def \div from $\ddot{-}$) it suffices to show that

$$\beta \in K \sim \alpha \text{ iff } \begin{cases} \alpha \vee \beta \in K \sim (\alpha \wedge (\alpha \vee \beta)) \text{ and } \alpha \notin K \sim (\alpha \wedge \beta) \\ \text{if } \not\models \alpha, \not\models \beta, \alpha \in K, \\ \beta \in K \text{ otherwise.} \end{cases}$$

We only consider the case where $\not\models \alpha, \not\models \beta$ and $\alpha \in K$. If $\beta \in K \sim \alpha$ then $\alpha \vee \beta \in K \sim \alpha = K \sim (\alpha \wedge (\alpha \vee \beta))$ by (K \div 5), and $\alpha \notin K \sim (\alpha \wedge \beta)$ follows from (K \div 10). Conversely, if $\alpha \vee \beta \in K \sim (\alpha \wedge (\alpha \vee \beta)) = K \sim \alpha$, and $\alpha \notin K \sim (\alpha \wedge \beta)$, then $\beta \in K \sim \alpha$ by (K \div 9). \square

B.2 Theorems 6.5.12 and 6.5.14

Theorem 6.5.12 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. Then*

$$\beta \notin K \div \alpha \text{ iff } \begin{cases} \beta \notin K \text{ and } \models \alpha, \text{ or} \\ \beta \notin K \text{ and } \alpha \notin K, \text{ or} \\ \beta \sqsubseteq_{RE} \alpha \text{ and } \not\models \alpha, \text{ or} \\ \beta \not\sqsubseteq_{RE} \alpha \text{ and } \exists \gamma \not\sqsubseteq_{RE} \alpha \text{ such that } \{\beta, \gamma\} \models \alpha, \end{cases} \quad (6.1)$$

or equivalently,

$$\beta \in K \div \alpha \text{ iff } \begin{cases} \beta \in K \text{ and } \models \alpha, \text{ or} \\ \beta \in K \text{ and } \alpha \notin K, \text{ or} \\ \beta \not\sqsubseteq_{RE} \alpha \text{ and for every } \gamma \sqsubseteq_{RE} \alpha, \{\beta, \gamma\} \not\models \alpha. \end{cases} \quad (6.2)$$

Proof For the left-to-right direction of the proof of (6.1), suppose that $\beta \notin K \div \alpha$ and that none of the first three of the four required possible cases hold. That is, suppose that $(\beta \in K \text{ or } \not\models \alpha)$ and $(\beta \in K \text{ or } \alpha \in K)$ and $(\beta \not\sqsubseteq_{RE} \alpha \text{ or } \models \alpha)$. This means that $(\beta \in K \text{ or } (\not\models \alpha \text{ and } \alpha \in K))$ and $(\beta \not\sqsubseteq_{RE} \alpha \text{ or } \models \alpha)$, which, in turn, means that $(\beta \in K \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $(\beta \in K \text{ and } \models \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \models \alpha)$. Of these four cases, the fourth one is a logical contradiction, while the second one contradicts the supposition that $\beta \notin K \div \alpha$ (since $K \div \alpha = K$ if $\models \alpha$ by (K÷6)). So it has to be the case that $(\beta \in K \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$. If $\beta \in K$ and $\beta \not\sqsubseteq_{RE} \alpha$ then, since $\beta \notin K \div \alpha$, we have that $K \neq K \div \alpha$, from which it follows by (K÷3) that $\alpha \in K$. And thus, by part (1) of proposition 6.5.11, it follows that there is a $\gamma \sqsubseteq_{RE} \alpha$ such that $\{\beta, \gamma\} \models \alpha$. Similarly, if $(\not\models \alpha \text{ and } \alpha \in K)$ and $\beta \not\sqsubseteq_{RE} \alpha$ then, together with the supposition that $\beta \notin K \div \alpha$, it follows that there is a $\gamma \sqsubseteq_{RE} \alpha$ such that $\{\beta, \gamma\} \models \alpha$. For the right-to-left direction, note that if $\beta \notin K$ then it follows from (K÷2) that $\beta \notin K \div \alpha$. If $\beta \sqsubseteq_{RE} \alpha$ and $\not\models \alpha$ then by proposition 6.5.6, $\beta \notin K \div \alpha$. Finally, if $\beta \not\sqsubseteq_{RE} \alpha$ and there is a $\gamma \sqsubseteq_{RE} \alpha$ such that $\{\beta, \gamma\} \models \alpha$, then by part (2) of proposition 6.5.11, $\beta \notin K \div \alpha$.

For the left-to-right direction of the proof of (6.2), suppose that $\beta \in K \div \alpha$ and that neither of the first two of the required three possible cases hold. That is, suppose that $(\beta \notin K \text{ or } \alpha)$ and $(\beta \notin K \text{ or } \alpha \in K)$. This means that $\beta \notin K$ or $(\not\models \alpha \text{ and } \alpha \in K)$. If $\beta \notin K$ then, since $\beta \in K \div \alpha$, it follows that $K \neq K \div \alpha$, and thus, by (K÷3) and (K÷6), that $\alpha \in K$ and $\not\models \alpha$. So, regardless of which of the two possibilities hold, it will be the case that $\alpha \in K$ and $\not\models \alpha$. Since $\beta \in K \div \alpha$ and $\not\models \alpha$, it follows from proposition 6.5.6 that $\beta \not\sqsubseteq_{RE} \alpha$. Now assume that there is a $\gamma \sqsubseteq_{RE} \alpha$ such that $\{\beta, \gamma\} \models \alpha$. Then, by part (2) of proposition 6.5.11, $\beta \notin K \div \alpha$, contradicting the supposition that $\beta \in K \div \alpha$. So it has to be the case that for every $\gamma \sqsubseteq_{RE} \alpha$, $\{\beta, \gamma\} \not\models \alpha$. For the right-to-left direction, note that if $\beta \in K$ and $\models \alpha$, or $\beta \in K$ and $\alpha \notin K$, respectively, then by (K÷6) or (K÷3) respectively, $\beta \in K \div \alpha$. So we need only consider the case in which these two possibilities do not hold. We have already seen above that if neither of these two possibilities hold, then $\alpha \in K$. Now suppose

that $\beta \not\sqsubseteq_{RE} \alpha$, and that $\{\beta, \gamma\} \not\models \alpha$ for every $\gamma \not\sqsubseteq_{RE} \alpha$. It then follows from part (1) of proposition 6.5.11 that $\beta \in K \div \alpha$. \square

Theorem 6.5.14 *Suppose that the RE-ordering \sqsubseteq_{EE} and the systematic withdrawal \div are semantically related. Then*

$$\beta \notin K \div \alpha \text{ iff } \begin{cases} \beta \notin K \text{ and } \models \alpha, \text{ or} \\ \beta \notin K \text{ and } \alpha \notin K, \text{ or} \\ \beta \sqsubseteq_{RE} \alpha \text{ and } \not\models \alpha, \text{ or} \\ \alpha \sqsubset_{RE} \beta \text{ and } \exists \gamma \in L \text{ such that} \\ \alpha \sqsubset_{RE} \gamma, \beta \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \{\beta, \gamma\} \models \alpha, \text{ or} \\ \alpha \parallel_{\sqsubseteq_{RE}} \beta \text{ and } \exists \gamma \in L \text{ such that} \\ \alpha \parallel_{\sqsubseteq_{RE}} \gamma, \beta \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \{\beta, \gamma\} \models \alpha, \end{cases} \quad (6.3)$$

or equivalently,

$$\beta \in K \div \alpha \text{ iff } \begin{cases} \beta \in K \text{ and } \models \alpha, \text{ or} \\ \beta \in K \text{ and } \alpha \notin K, \text{ or} \\ \alpha \sqsubset_{RE} \beta \text{ and } \forall \gamma \in L \text{ such that} \\ \alpha \sqsubset_{RE} \gamma \text{ and } \beta \parallel_{\sqsubseteq_{RE}} \gamma, \{\beta, \gamma\} \not\models \alpha, \text{ or} \\ \alpha \parallel_{\sqsubseteq_{RE}} \beta \text{ and } \forall \gamma \in L \text{ such that} \\ \alpha \parallel_{\sqsubseteq_{RE}} \gamma \text{ and } \beta \parallel_{\sqsubseteq_{RE}} \gamma, \{\beta, \gamma\} \not\models \alpha. \end{cases} \quad (6.4)$$

Proof To prove the left-to-right direction of (6.3), suppose that $\beta \notin K \div \alpha$ and that none of the first three of the five required possible cases hold. That is, suppose that $(\beta \in K \text{ or } \not\models \alpha)$ and $(\beta \in K \text{ or } \alpha \in K)$ and $(\beta \not\sqsubseteq_{RE} \alpha \text{ or } \models \alpha)$. This means that $(\beta \in K \text{ or } (\not\models \alpha \text{ and } \alpha \in K))$ and $(\beta \not\sqsubseteq_{RE} \alpha \text{ or } \models \alpha)$, which, in turn, means that $(\beta \in K \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $(\beta \in K \text{ and } \models \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \models \alpha)$. The fourth possibility above is a logical contradiction, while the second possibility contradicts the supposition that $\beta \notin K \div \alpha$ (since $K = K \div \alpha$ if $\models \alpha$, by (K \div 6)). So it must be the case that $(\beta \in K \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$ or $((\not\models \alpha \text{ and } \alpha \in K) \text{ and } \beta \not\sqsubseteq_{RE} \alpha)$. If $\beta \in K$, then $K \neq K \div \alpha$, and by (K \div 3), $\alpha \in K$. So in both cases, $\alpha \in K$ and $\beta \not\sqsubseteq_{RE} \alpha$. Now we can distinguish between two cases: Either $\alpha \sqsubseteq_{RE} \beta$ or $\alpha \not\sqsubseteq_{RE} \beta$. In the former case it follows that $\alpha \sqsubset_{RE} \beta$, and from part (1) of proposition 6.5.13 it then follows $\alpha \sqsubset_{RE} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$ for some $\gamma \in L$. In the latter case we have that $\alpha \parallel_{\sqsubseteq_{RE}} \beta$ and it then follows from part (2) of proposition 6.5.13 in that $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$ for some $\gamma \in L$. For the proof in the converse

direction note that if $\beta \notin K$ then it follows from $(K \div 2)$ that $\beta \notin K \div \alpha$. If $\beta \sqsubseteq_{RE} \alpha$ and $\not\models \alpha$, it follows from proposition 6.5.6 that $\beta \notin K \div \alpha$. If $\alpha \sqsubset_{RE} \beta$ and there is a γ such that $\alpha \sqsubset_{RE} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$, it follows from part (2) of proposition 6.5.11 that $\beta \notin K \div \alpha$. Similarly, if $\alpha \parallel_{\sqsubseteq_{RE}} \beta$ and there is a γ such that $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$, it follows from part (2) of proposition 6.5.11 that $\beta \notin K \div \alpha$.

To prove the left-to-right direction of (6.4), suppose that $\beta \in K \div \alpha$, and that neither of the first two of the four required possible cases hold. That is, suppose that $(\beta \notin K \text{ or } \not\models \alpha)$ and $(\beta \notin K \text{ or } \alpha \in K)$, which means that $\beta \notin K$ or $(\not\models \alpha \text{ and } \alpha \in K)$. If $\beta \notin K$ then, because $\beta \in K \div \alpha$, it follows that $K \neq K \div \alpha$, and thus, by $(K \div 6)$ and $(K \div 3)$, that $\not\models \alpha$ and $\alpha \in K$. So in both cases, $\not\models \alpha$ and $\alpha \in K$. Since $\beta \in K \div \alpha$ and $\not\models \alpha$, it follows from proposition 6.5.6 in that $\beta \sqsubseteq_{RE} \alpha$. We distinguish between two cases. Either $\alpha \sqsubseteq_{RE} \beta$, or $\alpha \not\sqsubseteq_{RE} \beta$. In the former case we get that $\alpha \sqsubset_{RE} \beta$. Now assume there is a γ such that $\alpha \sqsubset_{RE} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$. Then, by part (2) of proposition 6.5.11, it follows that $\beta \notin K \div \alpha$, contradicting the supposition that $\beta \in K \div \alpha$. So it has to be the case that $\{\beta, \gamma\} \not\models \alpha$ for every γ such $\alpha \sqsubset_{RE} \gamma$ and $\beta \parallel_{\sqsubseteq_{RE}} \gamma$. In the latter case, when $\alpha \sqsubseteq_{RE} \beta$, $\alpha \parallel_{\sqsubseteq_{RE}} \beta$. Now assume there is a γ such that $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$, $\beta \parallel_{\sqsubseteq_{RE}} \gamma$ and $\{\beta, \gamma\} \models \alpha$. Then, by part (2) of proposition 6.5.11, it follows that $\beta \notin K \div \alpha$, contradicting the supposition that $\beta \in K \div \alpha$. So it has to be the case that $\{\beta, \gamma\} \not\models \alpha$ for every γ such that $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$ and $\beta \parallel_{\sqsubseteq_{RE}} \gamma$. For the converse direction, note that if $\beta \in K$ and $\models \alpha$ (or $\beta \in K$ and $\alpha \notin K$), then it follows from $(K \div 6)$ (or $(K \div 3)$) that $\beta \in K \div \alpha$. So we need only consider the case in which these two possibilities don't hold. We've already seen above that if neither of these two possibilities hold, then $\alpha \in K$. Now, suppose that $\alpha \sqsubset_{RE} \beta$ and that $\{\beta, \gamma\} \not\models \alpha$ for every γ such that $\alpha \sqsubset_{RE} \gamma$ and $\beta \parallel_{\sqsubseteq_{RE}} \gamma$. Then it follows from part (1) of proposition 6.5.13 that $\beta \in K \div \alpha$. Similarly, if $\alpha \parallel_{\sqsubseteq_{RE}} \beta$ and $\{\beta, \gamma\} \not\models \alpha$ for every γ such that $\alpha \parallel_{\sqsubseteq_{RE}} \gamma$ and $\beta \parallel_{\sqsubseteq_{RE}} \gamma$, then it follows from part (2) of proposition 6.5.13 that $\beta \in K \div \alpha$. \square

Appendix C

List of identities

(Def $*$ from \sim)	page 21
(Def $-$ from $*$)	page 21
(Def \sim from s_K)	page 22
(Def $-$ from M)	page 22
(Def s_K from \subseteq)	page 23
(Def $-$ from \sqsubseteq_{EE})	page 25
(Def \sqsubseteq_{EE} from \sim)	page 25
(Def $*$ from \sqsubseteq_{GE})	page 26
(Def \sqsubseteq_E from \sqsubseteq_G)	page 27
(Def \sqsubseteq_G from \sqsubseteq_E)	page 27
(Def K/α from \sqsubset_H)	page 28
(Def $-$ from \sqsubset_H)	page 28
(Def \sim from sm_K)	page 40
(Def $*$ from sm_K)	page 40
(Def $*$ from \mathcal{S})	page 42
(Def $*$ from \preceq)	page 43
(Def \sim from \preceq)	page 43
(Def $*$ from B)	page 47
(Def \sqsubseteq_E from \preceq)	page 48
(Def \sqsubseteq_G from \preceq)	page 48
(Def \sqsubseteq_{GE} from $*$)	page 49
(Def \ll from \subseteq)	page 51
(Def $-$ from \ll)	page 51

(Def \sqsubseteq from \preceq)	page 52
(Def \sqsubseteq_{EE} from \sqsubseteq_H)	page 53
(Def \sim_P from P)	page 65
(Def \sim from \preceq)	page 72
(Def E from \sim)	page 72
(Def \sim from s_K)	page 73
(Def \sim from $*$)	page 75
(Def $*$ from \sim)	page 75
(Def sc_{\sqsubseteq})	page 90
(Def ∇_{\preceq})	page 90
(Def \preceq from κ)	page 96
(Def \sqsubseteq_{κ} from κ)	page 97
(Def \sqsubseteq_P from κ)	page 98
(Def \sqsubseteq_R from \preceq)	page 99
(Def \leq from \preceq)	page 106
(Def \preceq from \leq)	page 106
(Def \sqsubseteq_{RE} from \sqsubseteq_{EE})	page 112
(Def \sqsubseteq_{EE} from \sqsubseteq_{RE})	page 112
(Def \sqsubseteq_{RE} from \sim)	page 117
(Def $-$ from \sqsubseteq_{RE})	page 117
(Def $*$ from \sqsubseteq_{RG})	page 126
(Def \sqsubseteq_{GE} from \sqsubseteq_{RG})	page 127
(Def \sqsubseteq_{RG} from \sqsubseteq_{GE})	page 127
(Def \sim from \sqsubseteq)	page 127
(Def \sqsubseteq_C from \sqsubseteq_{EE} and \sqsubseteq_{GE})	page 129
(Def \sim from $-$ and s)	page 148
(Def \sim from ∇_{\preceq})	page 152
(Def $\ddot{-}$ from $\ddot{-}$)	page 156
(Def $\ddot{-}$ from \preceq)	page 157
(Def $-$ from \sim)	page 160
(Def $\ddot{-}$ from \sim)	page 160
(Def $\ddot{-}$ from \sim (v2))	page 160
(Def \div from $-$)	page 162
(Def \div from $\ddot{-}$)	page 162

(Def $\dot{-}$)	page 169
(Def \sim from \mathcal{V})	page 172
(Def \mathcal{V}_D from \mathcal{V})	page 173
(Def $\ddot{-}$ from \sqsubseteq_{EE})	page 174
(Def \sqsubseteq_{EE} from $\ddot{-}$)	page 175
(Def $\ddot{-}$ from \sqsubseteq_{RE})	page 176
(Def \div from \sqsubseteq_{EE})	page 177
(Def \div from \sqsubseteq_{RE})	page 178
(Def cug_{\sqsubseteq})	page 191
(Def \diamond_{\sqsubseteq})	page 191
(Def \ast from κ)	page 203
(Def \ast from κ)	page 204
(Def \sim from κ)	page 204
(Def \star from κ)	page 204
(Def \ast)	page 212
(Def \oplus)	page 217
(Def \ast from \approx)	page 219
(Def \ast_{\triangleright})	page 225
(Def \ast_{\triangleleft})	page 229
(Def \otimes_{\leftarrow})	page 233
(Def \otimes_{\rightarrow})	page 233
(Def \otimes_{\triangleleft})	page 234
(Def \otimes_{\triangleright})	page 234
(Def $K(\hat{\otimes})$)	page 235
(Def \preceq from IB)	page 245
(Def $-_{IB}$ from IB)	page 246
(Def \otimes from \ominus)	page 257
(Def \ast_{IB} from IB)	page 257
(Def \ast from \otimes)	page 264

Bibliography

Ernest Adams. *The logic of conditionals*. D. Reidel, Dordrecht, Netherlands, 1975.

Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50:510–530, 1985.

Carlos E. Alchourrón and David Makinson. Hierarchies of regulations and their logic. In R. Hilpinen, editor, *New Studies in Deontic Logic*, pages 125–148. D. Reidel, Dordrecht, 1981.

Carlos E. Alchourrón and David Makinson. On the logic of theory change: Contraction functions and their associated revision functions. *Theoria*, 48:14–37, 1982.

Carlos E. Alchourrón and David Makinson. On the logic of theory change: Safe contraction. *Studia Logica*, 44:405–422, 1985.

Carlos E. Alchourrón and David Makinson. Maps between some different kinds of contraction function: the finite case. *Studia Logica*, 45:187–198, 1986.

Carlos Areces and Veronica Becher. Iterable AGM functions. In *Proceedings of BR '98. Belief Revision Workshop. Trento, Italy*, 1998.

C. Baral, S. Kraus, and J. Minker. Combining multiple knowledge bases. *IEEE Transactions on Knowledge and Data Engineering*, 3(2):208–220, 1991.

C. Baral, S. Kraus, J. Minker, and V.S. Subrahmanian. Combining multiple knowledge bases consisting of first-order theories. *Computational Intelligence*, 8(1):45–71, 1992.

Jon Barwise and J. Seligman. *Information Flow. The Logic of Distributed Systems*. Cambridge University Press, New York, 1997.

- Patrick Blackburn, Jan Jaspars, and Maarten de Rijke. Reasoning about changing information. *South African Computer Journal*, (19):2–26, 1997.
- A. Borgida. Language features for flexible handling of exceptions in information systems. *ACM Transactions on Database Systems*, 10:563–603, 1985.
- A. Borgida and T. Imielinski. Decision making in committees: A framework for dealing with inconsistency and non-monotonicity. In *Non-Monotonic Reasoning Workshop (1984 : New Paltz, N.Y.)*, pages 21–32, Menlo Park, CA, 1984. American Association for Artificial Intelligence.
- Craig Boutilier. Conditional logics of normality as modal systems. In *AAAI 90: Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 594–599, Cambridge, MA, 1990. AAAI Press/MIT Press.
- Craig Boutilier. Epistemic entrenchment in autoepistemic logic. *Fundamenta Informaticae*, 17:5–29, 1992.
- Craig Boutilier. Revision sequences and nested conditionals. In R. Bajcsy, editor, *IJCAI-93. Proceedings of the 13th International Joint Conference on Artificial Intelligence held in Chambéry, France, August 28 to September 3, 1993*, volume 1, pages 519–525, San Mateo, CA, 1993. Morgan Kaufmann.
- Craig Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68:33–85, 1994.
- Craig Boutilier. Iterated revision and minimal changes of conditional beliefs. *Journal of Philosophical Logic*, 25(3):263–305, 1996.
- John P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22(1):76–84, 1981.
- John Cantwell. Some logics of iterated belief change. *Studia Logica (to appear)*, 1999.
- Rudolf Carnap and Yehoshua Bar-Hillel. An outline of a theory of semantic information. Technical Report 247, The Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.
- Rudolf Carnap and Yehoshua Bar-Hillel. Semantic information. *British Journal for the Philosophy of Science*, 4:147–157, 1953.

- Gregory J. Chaitin. *Algorithmic information theory*. Cambridge University Press, New York, 1987.
- Mukesh Dalal. Investigations into a theory of knowledge base revision. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, Saint Paul, Minnesota*, pages 475–479, 1988.
- Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. In Ronald Fagin, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 5–23, Pacific Grove, CA, 1994. Morgan Kaufmann.
- Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
- James P. Delgrande and Torsten H. Schaub. Compiling specificity into approaches to nonmonotonic reasoning. *Artificial Intelligence*, 90:301–348, 1997.
- Keith Devlin. *Logic and information*. Cambridge University Press, Cambridge, 1991.
- Jon Doyle. A truth maintenance system. *Artificial Intelligence*, 12(2):231–272, 1979.
- Jon Doyle. Reason maintenance and belief revision: Foundations versus coherence theories. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 29–51. Cambridge University Press, Cambridge, 1992.
- W. Edwards. Conservatism in human information processing. In B. Kleinmuntz, editor, *Formal Representation of Human Judgment*. Holt, Rinehart and Winston, New York, 1968.
- H. Einhorn and R. Hogarth. Confidence in judgment: Persistence in the illusion of validity. *Psychological Review*, 85:395–416, 1978.
- Herbert B. Enderton. *Mathematical Introduction to Logic*. Academic Press, 1972.
- Eduardo Fermé. On the Logic of Theory Change: Contraction Without Recovery. *Journal of Logic, Language and Information*, 7(2):122–137, April 1998.
- Eduardo Fermé and Ricardo Rodríguez. Semi-Contraction: Axioms and Construction. Unpublished manuscript, Universidad de Buenos Aires, April 1998.

- Melvin Fitting. *First-Order Logic and Automated Theorem Proving*. Springer, New York, 1996.
- Michael Freund and Daniel Lehmann. Belief revision and rational inference. Technical Report TR 94-16, The Leibniz Centre for Research in Computer Science, Institute of Computer Science, Hebrew University of Jerusalem, 1994.
- André Fuhrmann. Theory contraction through base contraction. *Journal of Philosophical Logic*, 20:175–203, 1991.
- André Fuhrmann and Sven Ove Hansson. A survey of multiple contractions. *Journal of Logic, Language and Information*, 3:39–76, 1994.
- André Fuhrmann and Isaac Levi. Undercutting and the Ramsey Test for Conditionals. *Synthese*, 101:157–169, 1994.
- Dov M. Gabbay. Theoretical foundations for non-monotonic reasoning in expert systems. In K. R. Apt, editor, *Proceedings NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, pages 439–457. Springer, Berlin, 1985.
- Peter Gärdenfors. Conditionals and changes of belief. In I. Niiniluoto and R. Tuomela, editors, *The Logic and Epistemology of Scientific Change: Acta Philosophica Fennica*, volume 30, pages 381–404. 1978.
- Peter Gärdenfors. Rules for rational changes of belief. In T. Pauli, editor, *320211: Philosophical Essays Dedicated to Lennart Aqvist on his Fiftieth Birthday*, pages 88–101. Philosophical studies No. 34, Department of Philosophy, University of Uppsala, Uppsala, 1982.
- Peter Gärdenfors. Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy*, 62:136–157, 1984.
- Peter Gärdenfors. *Knowledge in Flux : Modeling the Dynamics of Epistemic States*. The MIT Press, Cambridge, Massachusetts, 1988.
- Peter Gärdenfors. Belief revision: An introduction. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 1–28. Cambridge University Press, Cambridge, 1992.

- Peter Gärdenfors and David Makinson. Revisions of knowledge systems using epistemic entrenchment. In M.Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning About Knowledge*, pages 83–95. Morgan Kaufmann, Los Altos, California, 1988.
- Peter Gärdenfors and David Makinson. Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65:197–245, 1994.
- Peter Gärdenfors and Hans Rott. Belief revision. In Dov M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming: Epistemic and Temporal Reasoning*, volume 4, pages 35–132. Clarendon Press, Oxford, 1995.
- Hector Geffner. *Default Reasoning: Causal and Conditional Theories*. The MIT Press, Cambridge, Massachusetts, 1992.
- Hector Geffner and Judea Pearl. Conditional entailment: bridging two approaches to default reasoning. *Artificial Intelligence*, 53:209–244, 1992.
- Matthew L. Ginsberg. Counterfactuals. *Artificial Intelligence*, 30:35–79, 1986.
- Matthew L. Ginsberg. Introduction. In Matthew L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, chapter 1, pages 1–23. Morgan Kaufmann, San Mateo, California, 1987.
- Clark Glymour and Richmond H. Thomason. Default Reasoning and the Logic of Theory Perturbation. In *Non-Monotonic Reasoning Workshop (1984 : New Paltz, N.Y.)*, pages 93–102, Menlo Park, CA, 1984. American Association for Artificial Intelligence.
- Moisés Goldszmidt and Judea Pearl. On the relation between rational closure and system Z. In *Proceedings: Third International Workshop on Nonmonotonic Reasoning, South Lake Tahoe, CA*, pages 130–140, 1990.
- Moisés Goldszmidt and Judea Pearl. A maximum entropy approach to nonmonotonic reasoning. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15:220–232, 1993.
- Moisés Goldszmidt and Judea Pearl. Qualitative probabilities for default reasoning, belief revision and causal modeling. *Artificial Intelligence*, 84:57–112, 1996.

- Russel Greiner. The complexity of theory revision. *Artificial Intelligence*, 107(2):175–218, 1999.
- Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.
- Sven Ove Hansson. New operators for theory change. *Theoria*, 55:114–132, 1989.
- Sven Ove Hansson. Belief contraction without recovery. *Studia Logica*, 50:251–260, 1991.
- Sven Ove Hansson. A dyadic representation of belief. In Peter Gärdenfors, editor, *Belief Revision*, pages 89–121. Cambridge University Press, Cambridge, 1992a.
- Sven Ove Hansson. In defense of base contraction. *Synthese*, 91:239–245, 1992b.
- Sven-Ove Hansson. Changes of disjunctively closed bases. *Journal of Logic, Language and Information*, 2(4):255–284, 1993a.
- Sven Ove Hansson. Reversing the Levi identity. *Journal of Philosophical Logic*, 22:637–669, 1993b.
- Sven Ove Hansson. Theory contraction and base contraction unified. *The Journal of Symbolic Logic*, 58(2):602–625, 1993c.
- Sven Ove Hansson. Knowledge-level analysis of belief base operations. *Artificial Intelligence*, 82:215–235, 1996.
- Sven Ove Hansson. *A Textbook of Belief Dynamics, Theory Change and Database Updating (to appear)*. Kluwer Academic Publishers, Dordrecht, 1999.
- Sven Ove Hansson and Erik J. Olsson. Levi contractions and AGM contractions: A Comparison. *Notre Dame Journal of Formal Logic*, 36(1):103–119, 1995.
- Sven Ove Hansson and Hans Rott. How not to change the theory of theory change: A reply to Tennant. *British Journal for the Philosophy of Science*, 46:381–383, 1995.
- G. Harman. *Change in View: Principles of Reasoning*. The MIT Press, Cambridge, Massachusetts, 1986.

- W. L. Harper. Rational conceptual change. In *PSA 1976, Volume 2*, pages 462–494. Philosophy of Science Association, East Lansing, MI, 1977.
- E. Hoenkamp. An analysis of psychological experiments on non-monotonic reasoning. In *Proceedings of the 7th Biennial Meeting of the Canadian Society for the Computational Study of Intelligence*, 1988.
- William James. *Pragmatism — A New Name for Some Old Ways of Thinking*. Longmans, Green, and Co., New York, 1907.
- H. Katsuno and A.O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.
- H. Katsuno and K. Satoh. A unified view of consequence relation, belief revision and conditional logic. In *IJCAI-91. Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 406–412, San Francisco, CA, 1991. Morgan Kaufmann.
- Hirofumi Katsuno and Alberto Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 183–203. Cambridge University Press, Cambridge, 1992.
- A.M. Keller and M. Winslett Wilkins. On the use of an extended relational model to handle changing incomplete information. *IEEE Transactions on Software Engineering*, SE-11(7):620–633, 1985.
- N.E. Kfir-Dahav and M. Tennenholtz. Multi-agent belief revision. In Yoav Shoham, editor, *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Sixth Conference (TARK 1996)*, pages 175–194, San Francisco, California, 1996. Morgan Kaufmann.
- Sébastien Konieczny and Ramón Pino-Pérez. On the logic of merging. In A. G. Cohn, L. Schubert, and S. C. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference (KR '98)*, pages 488–498, San Francisco, California, 1998. Morgan Kaufmann.
- Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

- Daniel Lehmann. What Does a Conditional Knowledge Base Entail? In Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editors, *Proceedings First International Conference on Principles of Knowledge Representation and Reasoning, Toronto, Ontario*, pages 212–222, San Mateo, California, 1989. Morgan Kaufmann.
- Daniel Lehmann. Belief revision, revised. In C. S. Mellish, editor, *IJCAI-95. Proceedings of the 14th International Joint Conference on Artificial Intelligence held in Montreal, PQ, August 20–25, 1995*, volume 2, pages 1534–1540, San Mateo, CA, 1995. Morgan Kaufmann.
- Daniel Lehmann and Menachem Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.
- Isaac Levi. *Gambling with Truth*. The MIT Press, Cambridge, Massachusetts, 1973. Previous issue: 1967, Knopf, New York.
- Isaac Levi. Subjunctives, dispositions and changes. *Synthese*, 34:423–455, 1977.
- Isaac Levi. *The Enterprise of Knowledge*. The MIT Press, Cambridge, Massachusetts, 1980.
- Isaac Levi. Iteration of conditionals and the Ramsey test. *Synthese*, 76:49–81, 1988.
- Isaac Levi. *The Fixation of Belief and Its Undoing*. Cambridge University Press, Cambridge, 1991.
- Isaac Levi. *For the sake of argument*. Cambridge University Press, Cambridge, Massachusetts, 1996.
- Isaac Levi. Contraction and informational value. Unpublished manuscript, Columbia University, February 1998.
- D. K. Lewis. Counterfactuals. *Journal of Philosophy*, 70:556–567, 1973.
- Paolo Liberatore and Marco Schaerf. Arbitration (or How to Merge Knowledge Bases). *IEEE Transactions on Knowledge and Engineering*, 10(1):76–90, January/February 1998.
- Vladimir Lifschitz. On the satisfiability of circumscription. *Artificial Intelligence*, 28:17–27, 1986.

- Vladimir Lifschitz. Computing circumscription. In Matthew L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 167–173. Morgan Kaufmann, San Mateo, California, 1987.
- S. Lindström and W. Rabinowicz. Epistemic entrenchment with incomparabilities and relational belief revision. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change: Workshop, Konstanz, FRG, October 1989, Proceedings*, volume 465 of *Lecture Notes in Artificial Intelligence*, pages 93–126, Berlin, 1991. Springer-Verlag.
- Sten Lindström. A semantic approach to nonmonotonic reasoning: Inference operations and choice. Technical Report 6, Uppsala Prints and Preprints in Philosophy, Department of Philosophy, University of Uppsala, 1991.
- David Makinson. How to give it up. *Synthese*, 62:347–363, 1985.
- David Makinson. On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic*, 16:383–394, 1987.
- David Makinson. General theory of cumulative inference. In M. Reinfrank, J. de Kleer, M.L. Ginsberg, and E. Sandewall, editors, *Non-Monotonic Reasoning, 2nd International Workshop, Grassau, FRG, June 1988, Proceedings*, volume 346 of *Lecture Notes In Artificial Intelligence*, pages 1–18, Berlin, 1989. Springer-Verlag.
- David Makinson. Five faces of minimality. *Studia Logica*, 52:339–379, 1993.
- David Makinson. General patterns in nonmonotonic reasoning. In Dov M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Nonmonotonic reasoning and Uncertain Reasoning*, volume 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 35–103. Clarendon Press, Oxford, 1994.
- David Makinson. On the force of some apparent counterexamples to recovery. In E. G. Valdés, W. Krawietz, G. H. von Wright, and R. Zimmerling, editors, *Normative Systems in Legal and Moral Theory*. Dunker and Humboldt, Berlin, 1997.
- David Makinson and Peter Gärdenfors. Relations between the logic of theory change and nonmonotonic logic. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change: Workshop, Konstanz, FRG, October 1989, Proceedings*, volume 465

- of *Lecture Notes in Artificial Intelligence*, pages 185–205. Springer-Verlag, Berlin, 1991.
- João P. Martins and Stuart C. Shapiro. A model for belief revision. *Artificial Intelligence*, 35:25–79, 1988.
- John McCarthy. Circumscription — a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980.
- John McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28:89–116, 1986.
- Drew McDermott and Jon Doyle. Nonmonotonic logic I. *Artificial Intelligence*, 13:41–72, 1980.
- Drew McDermott and Jon Doyle. Nonmonotonic logic II: Nonmonotonic modal theories. *Journal of the ACM*, 29:33–57, 1982.
- Thomas A. Meyer, Willem A. Labuschagne, and Johannes Heidema. Infobase Change: A First Approximation. *Journal of Logic, Language and Information (to appear)*, 1999a.
- Thomas A. Meyer, Willem A. Labuschagne, and Johannes Heidema. Refined Epistemic Entrenchment. *Journal of Logic, Language and Information (to appear)*, 1999b.
- Robert Moore. Possible-world semantics for autoepistemic logic. In *Proceedings of the AAAI Workshop on Non-Monotonic Reasoning, New Paltz, NY*, pages 344–354, 1984.
- Robert Moore. Semantical considerations on non-monotonic logic. *Artificial Intelligence*, 25:75–94, 1985.
- Michael Morreau. Epistemic semantics for counterfactuals. *Journal of Philosophical Logic*, 21:33–62, 1992.
- Yoram Moses and Yoav Shoham. Belief as defeasible knowledge. *Artificial Intelligence*, 64:299–321, 1993.
- Abhaya C. Nayak. Foundational belief change. *Journal of Philosophical Logic*, 23:495–533, 1994a.

- Abhaya C. Nayak. Iterated belief change based on epistemic entrenchment. *Erkenntnis*, 41:353–390, 1994b.
- Abhaya C. Nayak, Paul Nelson, and Hanan Polansky. Belief change as change in epistemic entrenchment. *Synthese*, 109:143–174, 1996.
- Bernhard Nebel. A knowledge level analysis of belief revision. In Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, pages 301–311, San Mateo, CA, 1989. Morgan Kaufmann.
- Bernhard Nebel. *Reasoning and Revision in Hybrid Representation Systems*, volume 422 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Berlin, 1990.
- Bernhard Nebel. Belief revision and default reasoning: Syntax-based approaches. In James Allen, Richard Fikes, and Erik Sandewall, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference KR '91*, pages 417–428. Morgan Kaufmann, San Francisco, California, 1991.
- Bernhard Nebel. Syntax-based approaches to belief revision. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 52–88. Cambridge University Press, Cambridge, 1992.
- Alan Newell. The knowledge level. *Artificial Intelligence*, 18:87–127, 1982.
- R Niederee. Multiple contraction. A further case against Gärdenfors's principle of recovery. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change: Workshop, Konstanz, FRG, October 1989, Proceedings*, volume 465 of *Lecture Notes in Artificial Intelligence*, pages 322–334, Berlin, 1991. Springer-Verlag.
- Maurice Pagnucco. *The Role of Abductive Reasoning within the Process of Belief Revision*. PhD thesis, Basser Department of Computer Science, University of Sydney, Australia, 1996.
- Odile Papini. Iterated revision operations stemming from the history of an agent's observations. In *Proceedings of BR '98. Belief Revision Workshop. Trento, Italy*, pages 1–15, 1998.

- Odile Papini. Iterated revision operations stemming from the history of an agent's observations. In Hans Rott and Mary-Anne Williams, editors, *Frontiers of Belief revision (to appear)*. Kluwer Academic Publishers, Dordrecht, 1999.
- Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California, 1988.
- Judea Pearl. Probabilistic semantics for nonmonotonic reasoning: a survey. In Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editors, *Proceedings First International Conference on Principles of Knowledge Representation and Reasoning, Toronto, Ontario*, pages 505–516, San Mateo, California, 1989. Morgan Kaufmann.
- Judea Pearl. System Z: a natural ordering of defaults with tractable applications to non-monotonic reasoning. In R. Parikh, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 121–135, San Mateo, CA, 1990. Morgan Kaufmann.
- Francis Jeffery Pelletier and Renee Elio. What should default reasoning be, by default? *Computational Intelligence*, 13(2):165–187, 1997.
- Pavlos Peppas and Paul Sprakis. Smoothness Conditions In Belief Revision. *Studia Logica (submitted)*, 1999.
- Pavlos Peppas and Mary-Anne Williams. Constructive modellings for theory change. *Notre Dame Journal of Formal Logic*, 36(1):120–133, 1995.
- David Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36: 27–47, 1988.
- Wlodzimierz Rabinowicz. Stable Revision, or is Preservation Worth Preserving? In André Fuhrmann and Hans Rott, editors, *Logic Action Information*, pages 101–128. Walter de Gruyter and Co., Berlin, 1995.
- Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.
- Raymond Reiter and Johan de Kleer. Foundations of assumption-based truth maintenance systems: Preliminary report. In *Proceedings 6th NCAI*, pages 183–188, 1987.

- P. Z. Revesz. On the Semantics of Theory Change: Arbitration between Old and New Information. In *Proceedings PODS '93, 12th ACM SIGACT SIGMOD SIGART Symposium on the Principles of Database Systems*, pages 71–82, 1993.
- L. Ross and M. Lepper. The perseverance of beliefs: empirical and normative considerations. In R. Schweder, editor, *Fallible Judgment in Behavioral Research*. Jossey-Bass, San Francisco, California, 1980.
- Hans Rott. Two methods of constructing contractions and revisions of knowledge systems. *Journal of Philosophical Logic*, 20:149–173, 1991.
- Hans Rott. Modellings for belief change: Prioritization and entrenchment. *Theoria*, 58 (1):21–57, 1992a.
- Hans Rott. On the Logic of Theory Change: More Maps Between Different Kinds of Contraction Function. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 122–141. Cambridge University Press, Cambridge, 1992b.
- Hans Rott. Preferential belief change using generalized epistemic entrenchment. *Journal of Logic, Language and Information*, 1:45–78, 1992c.
- Hans Rott. Just because. Taking belief bases very seriously. In Sven Ove Hansson and Wlodimierz Rabinowicz, editors, *Logic for a Change*, volume 9, pages 106–124. Uppsala Prints and Preprints in Philosophy, 1995.
- Hans Rott and Maurice Pagnucco. Severe withdrawal (and Recovery). *Journal of Philosophical Logic (to appear)*, 1999.
- K. Satoh. Nonmonotonic reasoning by minimal belief revision. In *Proceedings International Conference on Fifth Generation Computer Systems, Tokyo*, pages 455–462, 1988.
- Claude Elwood Shannon. *The mathematical theory of communication*. Illinois U.P., 1964.
- Claude Elwood Shannon. *Claude Elwood Shannon: collected papers*. IEEE Press, New York, 1993.

- Yoav Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, Cambridge, Massachusetts, 1987a.
- Yoav Shoham. A semantical approach to nonmonotonic logics. In Matthew L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 227–250, San Mateo, California, 1987b. Morgan Kaufmann.
- Wolfgang Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In William L. Harper and Brian Skyrms, editors, *Causation in Decision: Belief, Change and Statistics: Proceedings of the Irvine Conference on Probability and Causation: Volume II*, volume 42 of *The University of Western Ontario Series in Philosophy of Science*, pages 105–134, Dordrecht, 1988. Kluwer Academic Publishers.
- Wolfgang Spohn. A Reason for Explanation: Explanations Provide Stable Reasons. In Wolfgang Spohn, Bas C. Van Fraassen, and Brian Skyrms, editors, *Existence and Explanation: Essays presented in Honor of Karel Lambert*, volume 49 of *University of Western Ontario series in philosophy of science*, pages 165–196. Kluwer Academic Publishers, Dordrecht, 1991.
- Robert Stalnaker, W. Harper, and G. Pearce. *Ifs*. D. Reidel Publishing Company, Dordrecht, 1981.
- V.S. Subrahmanian. Amalgamating knowledge bases. *ACM Transactions on Database Systems*, 19(2):291–331, 1994.
- Neil Tennant. Changing the theory of theory change: Towards a computational approach. *British Journal for the Philosophy of Science*, 45:865–897, 1994.
- Neil Tennant. On having bad contractions, or : no room for recovery. *Journal of Applied Non-Classical Logics*, 7(2):241–266, 1997.
- Johan van Benthem. Foundations of conditional logic. *Journal of Philosophical Logic*, 13:303–349, 1984.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25: 221–261, 1996.

- A. Weber. Updating propositional formulas. In *Proceedings First Conference on Expert Database Systems*, pages 487–500, 1986.
- Mary-Anne Williams. Transmutations of knowledge systems. In Jon Doyle, Erik Sandewall, and Pietro Torasso, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth International Conference (KR '94)*, pages 619–629, San Francisco, California, 1994. Morgan Kaufmann.
- M. Winslett. Reasoning about action using a possible models approach. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, Saint Paul, Minnesota*, pages 89–93, Palo Alto, CA, 1988. American Association for Artificial Intelligence, Morgan Kaufmann.

Index

- \models , *see* satisfaction
- \models , *see* content relation
- \models , *see* semantic entailment
- \vdash , *see* abstract consequence relation
- \sim , *see* nonmonotonic consequence relation
- \preceq_{IB} , *see* IB -induced faithful total pre-order
- \preceq -minimal, 42
- ∇_{\preceq} , *see* downset
- \diamond_{\sqsubseteq} , *see* upset
- $K \perp \alpha$, *see* remainder
- $K \top \alpha$, *see* entailment set
- \sqsubset_H , *see* hierarchy
- \sqsubset_P , *see* Spohn's plausibility ordering
- \sqsubseteq_C , *see* combined entrenchment
- \sqsubseteq_{CR} , *see* CR-ordering
- \sqsubseteq_{EE} , *see* epistemic entrenchment
- \sqsubseteq_{GE} , *see* GE-ordering
- \sqsubseteq_{GRE} , *see* GRE-ordering
- \sqsubseteq_{κ} , 97
- \sqsubseteq_R , *see* refined ordering
- \sqsubseteq_{RE} , *see* refined entrenchment
- \parallel_{\sqsubseteq} , 14
- $-_{IB}$, *see* IB -induced theory contraction
- $*_{IB}$, *see* IB -induced theory revision
- $*_{\triangleright}$, *see* P_{\triangleright} -revision
- $*_{\triangleleft}$, *see* P_{\triangleleft} -revision
- \otimes -associated revision, *see* associated revision on epistemic states
- $\overleftarrow{\sigma}$, 244
- $|\sigma|$, 244
- A_{PL} , 11
- A -
 - believed, 14
 - disbelieved, 14
 - established, 14
 - neutral, 14
 - refuted, 14
 - undecided, 14
- $\hat{\alpha}$, *see* l-model
- α -
 - contraction, 17
 - discarded wffs, *see* discarded wffs
 - removal, 17
 - retained wffs, *see* retained wffs
 - revision, 17
 - weakened version of wff, *see* weakened version of wff
- adjustment, 204
- arbitration, 235
- associated
 - belief base with a belief set, 240
 - infatom, 34
 - infatom semantics, 34

- infobase and a belief set, 244
- revision on epistemic states, 265
- theory contraction with a base contraction, 242
- theory contraction with infobase IB-contraction, 245
- theory revision with infobase revision, 257
- valuation, 34
- valuation semantics, 34
- assumption, 85
- Assumption Based Truth Maintenance Systems, 2
- axiomatisation
 - finite, of a belief set, 13
- axiomatise
 - belief set, 13
 - set of infatoms, 37
 - set of interpretations, 13
- background context, 60
- base change, 240
- belief, 1
- belief base, 240
- belief change, 1
 - iterated, 201
- belief revision, 1
- belief set, 13
- belief update, 2
- C , *see* semantic content
- choice function, 261
- Cleopatra
 - example, 141, 166
- closest upper gate, 191
- Cn , 11
- co-atom, 96
- coherentist, 2
 - basic infobase revision, 268
- combined entrenchment, 129
- commensurability thesis, 7
- compactness, 11
- conditional
 - assertion, 82
 - belief, 212
 - knowledge base, 82
 - logic, 71
- conditionalisation, 203
- connected, 14
- consequence relation
 - abstract, 12
 - cumulative, 62
 - expectation based, 72
 - E -based, 75
 - loop-cumulative, 62
 - nonmonotonic, 61
 - P -induced, 65
 - preferential, 62
 - rational, 62, 67
- conservative core, 82
- content
 - bit, 34
 - element, 33
 - relation, 33
- contraction, 7
 - AGM, 19
 - on epistemic states, 218
 - basic AGM, 18
 - basic infobase, 252

- full meet, 22
- maxichoice, 22
- mild, 173
- package, 280
- partial meet, 22
 - connectively relational, 23
 - relational, 23
 - transitively relational, 23
- pseudo-, 243
- safe, 28
- saturatable, 144
- semi-, 148
- contraction postulates
 - basic AGM, 18
 - supplementary AGM, 19
- converse complement, 103
- core, 91
- corroborating evidence, 223
- counterfactual reasoning, 71
- countermodel, 11
- CR-ordering, 130
- cug_{\sqsubseteq} , *see* closest upper gate
- DAG, *see* directed acyclic graph
- default information, 81
- defeasible reasoning, 59
- dependent, 256
- directed acyclic graph, 196
- discarded wffs, 246
- downset, 90
- DP-
 - postulates, 210
 - revision, 210
- EE-ordering, *see* epistemic entrenchment
- element-equivalent, 244
- elementarily equivalent, 12
- entailment
 - ϵ -entailment, 83
 - preferential, 82
 - rational, 83
 - semantic, 11
- entailment set, 27
- epistemic
 - entrenchment, 25
 - relevance, 259
 - state, 1
- equivalent
 - logically, 14
 - X -equivalent, 248
- evidence, 60
- expansion, 14
 - on epistemic states, 217
- expectation state, 76
- expectations, 72
- faithful, 42, 45
- fallback, 91
 - based withdrawal, 163
 - families, 103
- filtering condition, 256
- fixed information, 60
- fixed point ordering
 - revision operation, 234
- foundationist, 2
- FPO, *see* fixed point ordering
- Gärdenfors triviality result, 82
- GE-ordering, 26
- GRE-ordering, 132

- hamburger example
 - Hansson's, 240
- Harper Identity, 21
- hierarchy, 28
 - continuing down, 28
 - continuing up, 28
 - regular, 28
 - virtually connected, 29
- i -containing, 34
- (IB, α) -relevant, *see* relevant
- IB -dependent, *see* dependent
- IB -induced
 - faithful total preorder, 245
 - theory contraction, 246
 - theory revision, 257
- IB -number, 245
- Inf, 33
- infatom, 32
- infobase change operations, 244
- infon, 33
- informational value, 171
 - damped, 173
 - undamped, 172
- interchangeable, 21, 25, 56, 113, 118, 140, 161, 163, 174, 175
- interpolation thesis, 157
- interpretation, 11
- introspective beliefs, 273
- irrelevance, 69
- iterated revision
 - dynamic view of, 206
 - static view of, 205
- K_κ , 96
- K -linear order, 114
- KLM approach, 61
- knowledge, 1
- knowledge level, 3
- L , formal object language, 10
- l -model, 65
- L -revision, 223
- labelling function, 65
- Levi Identity, 21
- limit assumption, 42
- LR-entrenchment, 102
- LR-ordering, 101
- merging, 231
- Min_\preceq , 42, 234
- minimal-equivalent, 54
- model, 11
- modular
 - strict partial order, 67
 - weak partial order, 106
- multiple
 - change, 280
 - revision, 280
 - withdrawal, 142
- N_{PL} , 11
- N_β^R , *see* neutralised
- N_{IB}^R , *see* neutralised
- n -
 - reasoning context, 86
 - refined epistemic state, 86
- natural conditional function, 202
- neutralised, 249
- nonmonotonic, 59

- nonmonotonic reasoning
 - dynamic flavour of, 79
 - dynamic view of, 77
 - static view of, 80
- OCF, *see* ordinal conditional function
- ordinal conditional function, 96, 202
- P_{\triangleright} -revision, 226
- P_{\triangleleft} -revision, 230
- partial meet Levi-contraction, 145
- plausibility ordering
 - Grove, 26
 - Spohn, 98
- plausible, 59
- plausible consequence, 61
- power order, 48
- preferential model, 65
- preorder, 14
 - B-faithful, 46
 - faithful, 42
 - KM-faithful total, 46
 - layered, 131
 - total, 14
- principle of
 - Categorical Matching, 4
 - Conservatism, 5
 - Identity of Indiscernibles, 39
 - Indifference, 5
 - Informational Economy, 5
 - Irrelevance of Syntax, 4
 - Minimal Change, 4
 - Preference, 5
 - Reductionism, 3
- propositional atoms, 11
- pseudo-contraction, *see* contraction
- (R, α, β) -neutralised, *see* neutralised
- (R, α) -neutralised, *see* neutralised
- R-ordering, *see* refined ordering
- ranked model, 68
- rational closure, 83
- RE-ordering, *see* refined entrenchment
- reason maintenance, 240
- reasoning
 - cautious and bold, 87
- refined entrenchment, 107
- refined ordering, 99
- relevance selection function, 252
- relevant, 248
- remainder, 22, 259
- retained wffs, 246
- revision
 - AGM, 20
 - basic AGM, 20
 - basic infobase, 257
 - multiple, 280
 - on epistemic states, 207
- revision postulates
 - basic AGM, 20
 - supplementary AGM, 20
- revision-equivalent, 155
- revision-equivalent class
 - principled, 155
- RG-ordering, 124
- $S(\sigma)$, 244
- satisfaction, 11
- satisfiable, 13
- sc_{\sqsubseteq} , *see* strict cut

- selection function, 22
 - saturatable, 146
 - semantic, 40
 - semi-, 148
- semantic content, 34
- semantically related, 44, 50, 56, 106, 107, 118, 124, 131, 132, 135, 156, 161, 174
- semantics
 - infatom, 34
 - possible-worlds, 11
 - valuation, 12
 - classical, 38
- semi-contraction, *see* contraction
- smooth, 42
- specificity, 83
- sphere-semantics, 41
- state description, 38
- stopperedness, 42
- strict cut, 90
- subsequence, 244
 - ordered, 244
- symbol level, 4
- system of spheres, 41
- system-Z, 84
- Th*, *see* theory
- theory, 13
 - determined by, 13
 - generated by, 34
- transitively relational
 - binary relation, 73
- transmutation, 203
- transparent propositional language, 14
- Truth Maintenance Systems, 2
- Tweety, 59
- U , set of interpretations, 11
- u_{IB} , *see* IB -number
- upset, 191
- V , set of valuations, 11
- valuation, 11
- W -smooth, 42
- weakened version of wff, 250
- well-founded, 64
- widening ranked models, 224
- withdrawal, 19
 - fallback-based, 163
 - methodical, 171
 - multiple, 142
 - proper, 144
 - reasonable, 157
 - saturatable, 146
 - sensible, 147
 - severe, 152
 - on epistemic states, 218
 - systematic, 153
- X -equivalent, *see* equivalent