Resource-Bounded Belief Revision

ILLC Dissertation Series 2000-01

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-5256090
fax: +31-20-5255101
e-mail: illc@wins.uva.nl

Resource-Bounded Belief Revision

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam,
op gezag van de Rector Magnificus
prof.dr. J.J.M. Franse
ten overstaan van een door het college voor promoties ingestelde
commissie in het openbaar te verdedigen in de
Aula der Universiteit
op donderdag 27 januari 2000 te 13.00 uur

door

Renata Wassermann

geboren te São Paulo, Brazilië.

Promotor: Prof. dr. H. Rott
Institute for Philosophy
University of Regensburg
93040 Regensburg
Germany

Co-promotor: dr. F. Voorbraak
Department of Mathematics, Computer Science, Physics and Astronomy
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam

The investigations were supported by CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

CIP gegevens

Cover design by Madalena Machado.
Printed and bound by Print Partners Ipskamp.

ISBN: 90-5776-040-1

“Don’t be in such a hurry to believe next time — I’ll tell you why — If you set to work to believe everything, you will tire out the muscles of your mind, and then you’ll be so weak you won’t be able to believe the simplest true things. Only last week a friend of mine set to believe Jack-the-giant-killer. He managed to do it, but he was so exhausted by it that when I told him it was raining (which was true) he *couldn’t* believe it, but rushed out into the street without his hat or umbrella, the consequence of which was his hair got seriously damp, and one curl didn’t recover its right shape for nearly two days.”
(Letter from Lewis Carroll to Mary Macdonald, 1864.)

Contents

Acknowledgments

Writing a PhD thesis is not a solitary work as many people claim (and complain). My experience is that it takes a good environment and many people around you to keep you working hard enough and keep you from only working. And once you find the balance, it can be a very enjoyable period.

There are many people that deserve credits here. First of all, I would like to thank my supervisors Hans Rott and Frans Voorbraak. I have learned a lot from them, and they formed a very complementary team, with different but compatible interests. I also want to thank the other members of the graduation committee: Johan van Benthem, Jeroen Groenendijk, Sven Ove Hansson, John-Jules Meyer, and Frank Veltman.

My decision to come to Amsterdam was strongly influenced by two Dutch friends that I met in Brazil and that gave me a very good impression of the people and of the University of Amsterdam, Richard Benjamins and Frank van Harmelen. Frank put me in touch with the ILLC and helped me with my application. Richard was moving back to Amsterdam on the week I arrived, so I already had my little circle of friends.

When I decided to come to Amsterdam, even though I was awarded a very good grant from CAPES, one of the Brazilian national funding agencies, and had signed a contract saying that I would go back after finishing my studies, the Dutch alien police still imagined that I was trying illegal immigration and gave me a hard time before letting me into the country. That the process of acquiring a visa lasted only seven months is due to the hard work of Erik-Jan van der Linden and Mirjam Adelaar, who sent thousands of faxes, letters and e-mails and called the people in charge of my process every day until they let me in. I also have to thank Henk Zeevat and Remko Scha for accepting me as a student. Henk taught me a lot, from a crash course on semantics to the best place to buy a second-hand bike. It was during my first year, chatting with Henk, that I heard the words "Belief Revision" for the first time. A visit to Noor van Leusen's office provided me enough literature on the subject to keep me busy for some months. And

reading all those papers, specially Gärdenfors book, I suddenly found out what I wanted to be when I grew up. I then joined Michiel van Lambalgen's PIONEER project "Reasoning with Uncertainty". I have seen Michiel during these years as a kind of academic father. Although he was not directly involved with my daily supervision, I always had the feeling that he was taking care of me, making sure that I had all I needed to complete my work.

Michiel encouraged me to spend some time abroad and I took the opportunity to visit Peter Gärdenfors in Lund. I would like to thank Peter for accepting me as a visitor and for the many tea-breaks where we exchanged ideas. These ideas led to two articles that are not part of the dissertation, but that I keep as directions for future work. During my stay in Sweden I met Sven Ove Hansson for a coffee and we immediately started a very fruitful collaboration. I have to thank him for his encouragement, and for letting me use our joint work as part of this dissertation.

During the last months of my PhD project, I decided to apply the theory I had developed to the area of model-based diagnosis. Frank van Harmelen and Annette ten Teije provided me the necessary references. Talking to Frank I finally convinced myself (and him, I hope) that the relationship between belief revision and diagnosis was deeper than what we thought.

Carlos Areces, Eduardo Fermé, Peter Krause, Francesco Orilia, and Maurice Pagnucco were kind enough to read the manuscript and help me to improve both the contents as the presentation.

Rosalie Iemhoff and Ingrid van Loon have helped me with the Dutch samenvatting. Nikos Massios and Marco Vervoort were my computer wizards. The cover was designed by my sister, Madalena Machado.

I could always count on Ingrid van Loon, Dick de Jongh, Marco de Vries, and later on Marjan Veldhuisen and Peter Blok to help me out of all kinds of problems during the time I was at the ILLC.

My four years in the Netherlands were supported by grants from CAPES, the Spinoza Project "Logic in Action", the Faculty of Humanities of the University of Amsterdam, and the NWO PIONEER-project "Reasoning with Uncertainty".

As I said in the beginning, not only the ones who made me work, but also the ones who kept me from working deserve words of gratitude. I will not be able to cite all of them, but here are some: my office-mates Rosalie Iemhoff, Nikos Massios and Marco Vervoort, Chris Albert, Marco Aiello, Carlos Areces, Mario Benevides, Raffaella Bernardi, Caterina Caracciolo, Rosella Gennari, Silvia Olabarriaga, Marc Pauly, and Paulo Santos. Visits from my parents, my brothers, my grandfather (my friends still talk about the caipirinhas he prepared in 96!), and many friends brought a bit of São Paulo to Amsterdam.

My family and my friends in Brazil provided me an incredible amount of e-mail emotional support, and every time I was there on holidays, they made me feel as if I had never left. During a particularly difficult moment when I had to depend on them for everything, they made me feel happy that I could depend on them for everything. From the Brazilian support team, I want to thank especially three wonderful women, Edith, Denise, and Claudia. Back in Europe, it took me some time to be able to become independent again. Silvia and Peter had to play the role of my family and they did a great job!

I dedicate this book to my father – it was the spirit I inherited from him that brought me here.

Amsterdam, November of 1999.

Chapter 1

Introduction

The problem of belief revision has been extensively studied during the last twenty years. Given an agent with a set of (ascribed) beliefs, how should he change his beliefs when confronted with new information? This is the most general formulation of the problem of belief revision. An agent may be a human being, a computer program or any kind of system to which one can ascribe beliefs and from which one would expect rational reactions.

This is a multidisciplinary problem, with applications to several areas. We can give some examples of belief revision as it appears in:

- Daily life: I believed it was always raining in Amsterdam. One morning I woke up in Amsterdam and the sun was shining. I believed that on that day the weather was fine, contradicting my previous belief. I had to give up my belief that it always rained there.

- Databases: In the database containing data about the customers of a bookstore, there is an entry for John Smith, with his date of birth being 20/2/67. I get then a new order, where John Smith's date of birth is 20/2/76. I cannot add another date of birth and John's date of birth cannot have changed with time. I have to decide what to do. Keep the old data? Substitute it by the new? Or is it another John Smith after all, who should be added to the database?

- Robotics: A mobile robot has a map of the environment where it is supposed to move. On the map, there is nothing in front of it, so it should be able to move straight. But then its sensors indicate the presence of a big object in front of the robot. Should it doubt its sensors and continue trying to move straight? Or should it believe its sensors and doubt the map?

- Diagnosis: I believe that if I put an article at the right position on a properly working copying machine, I get copies of the article. Suppose I put an article at the right position, but all I get are blank pages. Should I give up my belief that I chose the right position? Or should I give up the belief that the copying machine is working properly?

Belief revision has been extensively studied in philosophy for extremely idealized agents. The agents considered are infinite beings, without any limitation of memory, time, or deductive ability. However, adapting these solutions to less idealized agents is far from trivial. In order to solve the problems cited above in a way which can be used by real agents, one has to consider that any realizable agent is a finite being and that calculations take time [Che86]. We need a theory which takes these characteristics — finiteness, memory and time limitations — into account.

Departing from the standard logical model for belief revision, the main goal of the present work is to find a theory that can be applied to more realistic agents. We stress here that our purpose is not to find a computational implementation of existing theories, but to elaborate a theory for less idealized agents.

In a recent paper, Chopra and Parikh [CP99] presented some desiderata for a belief revision formalism which we also see as our goals: distinction between explicit and implicit beliefs, no trivialization in the presence of inconsistencies, computational tractability, and minimal change.

The main achievements of our work are:

1. Formalization of a richer notion of belief state, based on the informal works of Harman and Cherniak (Chapter 4).

2. Generalization of standard results found in the literature, allowing for the use of more general logics (Chapter 5). This part is joint work with Sven Ove Hansson.

3. Design of a psychologically motivated, computationally efficient method for focussing on the relevant part of a belief state (Chapter 6).

4. Application of the developed framework to the problem of model-based diagnosis and use of the computational tools from model-based diagnosis for implementing belief revision operators (Chapter 7).

1.1 Organization of the Thesis

In the next chapter, we present an overview of some theories about resource-bounded agents. Although being very informal, these theories contain ideas which we will formalize in the following chapters.

In Chapter 3, we will present an overview of the main line of research in belief revision. We start by introducing the AGM paradigm and then present some variations on it that have been proposed in the literature.

In Chapter 4, we introduce our formal framework. We describe the belief state of a resource-bounded agent and operations that can be applied to it. We also show how our framework relates to the AGM paradigm and to Harman's informal proposal. Our framework classifies beliefs according to their status: whether they are explicitly believed or not, whether they are active or not and whether they are fully accepted or only provisional. All the operations affect the set of active beliefs. This set contains the beliefs which are available at a certain time and it changes according to the goals of the agent.

In Chapters 5 and 6, we turn to the problem of deciding which beliefs should become active for a belief change operation. Chapter 5 presents a logical solution, making use of the notion of local inference. We offer axiomatizations and representation results for generalizations of several belief change operations found in the literature.

In Chapter 6 we use extra-logical information in order to select which part of the belief state should become active. We present a computationally efficient method of retrieving the relevant part of an agent's beliefs and show how this method can be combined with the logical results obtained in Chapter 5.

In Chapter 7, we present an application of the formal framework developed in the preceding chapters. We show how to use extra-logical information present in the system descriptions used for model-based diagnosis in order to focus on a small relevant portion of the system. We also show how an algorithm used for finding minimal diagnoses can be adapted for implementing belief revision operators.

Finally, in Chapter 8 we present some conclusions and point toward future work.

Some of the results presented here have appeared somewhere else. Parts of Chapter 4 were published as [Was97] and [Was99b]. A preliminary version of Chapter 5 was presented as [HW98]. Most of Chapter 6 appeared as [Was98] and [Was99a].

1.2 Notation and Preliminaries

We consider L to be a propositional language closed under the usual truth-functional connectives $\neg$, $\vee$, $\wedge$, $\rightarrow$, $\leftrightarrow$, and containing a constant $\perp$ denoting falsum. We will use Greek lowercase letters $(\alpha, \beta, \gamma, \varphi, \psi, ...)$ to denote formulas of the language L. Sets of formulas will be denoted by uppercase letters $(A, B, K, X, Y, ...)$. Propositional letters of L are denoted by lowercase letters $(a, b, p, q, r, ...)$.

We call any total function taking sets of formulas to sets of formulas an *inference operation.* A Tarskian consequence operator is an inference operator C that satisfies monotony ($A \subseteq B \Rightarrow C(A) \subseteq C(B)$), inclusion ($A \subseteq C(A)$) and idempotency or iteration ($C(C(A)) = C(A)$). An example is the classical consequence operator Cn, defined by $Cn(A) = \{\alpha | A \vdash \alpha\}$, where $\vdash$ is the classical consequence relation for propositional logic. For simplicity, unless explicitly mentioned otherwise, we will work on the rest of this thesis with this classical consequence operator.

Chapter 2

Resource-Bounded Agents

In this chapter we will present some of the theories found in the literature that deal with the problem of bounded rationality. Instead of considering perfect reasoners as most theories based on logic do, the theories presented in this chapter try to analyze the effect of limited resources on the process of reasoning. These theories give some theoretical standards that we can use as a basis for a formalism, but they do not provide any sort of formal framework. In Chapter 4, we will present a formalization of several ideas from the informal theories described in this chapter.

We present theories by Harman [Har86], Cherniak [Che86], Levi [Lev91], and Russell and Wefald [RW91]. Of these, only Russell and Wefald show interest in implementing artificial agents. Harman and Cherniak are mainly concerned with human reasoning, although both make use of examples of artificial agents and AI literature. Levi, although mostly writing about idealized reasoning, presents some ideas about how a real agent deals with the problem of changing his belief state.

It is important to note that what we are looking for is not a limited implementation of a theory for ideal reasoning, but rather a theory for reasoners with limited resources, such as humans, computers, robots. As an example, consider the assumption very often made that the agent's beliefs are closed under logical rules. This leads not only to a problem from the computational point of view but there is also the question of why an agent would want to waste resources deriving all irrelevant consequences of his beliefs. The theories presented in this chapter are a step toward a more adequate account of how limited reasoners change their beliefs. Under certain assumptions (unlimited memory, time, logical ability) a theory for limited reasoners becomes a theory for perfect reasoners.

A theory about reasoning can be *normative* or *descriptive*. A normative theory prescribes the way an agent should reason, while a descriptive theory explains how the agent really reasons. The theories summarized here are basically normative, but the term is applied in a sense which is different from most works on epistemology since they prescribe how agents should reason given that their resources are limited.

2.1 Change in View

In this section we will present a summary of Harman's theory for belief change. The section is based on [Har86].

In his book, Harman tries to identify principles of reasoning, which, he claims, have nothing to do with principles of logic. Principles of logic, or deduction, do not show how one should change one's beliefs or intentions, so they cannot be taken as rules of reasoning. As an example, suppose someone believes p and $p \rightarrow q$. It is not always the case that one must accept q, one may decide to give up p or $p \rightarrow q$ instead. Harman presents the following example of this fact:

> **Example 1:** "Mary believes that if she looks in the cupboard, she will see a box of Cheerios. She comes to believe that she is looking in the cupboard and that she does not see a box of Cheerios. At this point, Mary's beliefs are jointly inconsistent and therefore *imply* any proposition whatsoever. This does not authorize Mary to *infer* any proposition whatsoever. Nor does Mary infer whatever she might wish to infer. Instead, she abandons her first belief, concluding that it is false after all.
>
> Furthermore, even before Mary fails to find any Cheerios in the cupboard, it would be silly for her to clutter her mind with vast numbers of useless logical implications of her beliefs, such as *either she will have Cheerios for breakfast or the moon is made of green cheese.*" ([Har86], pages 5-6)

Harman claims that there are two phases of reasoning to be distinguished: reflection (think about beliefs, plans, desires and various possibilities) and revision (actual changes), which may or may not follow reflection.

According to Harman, his theory is both normative and descriptive in a sense, since he observes how people reason in order to say how they should reason. The theory could be seen as a descriptive theory, but involving a certain amount of idealization.

His main concern is with human reasoning and not artificial intelligence. He bases his theory on intuitions and practical experience. The problem with the intuitions is that they tend to disregard limitations in time, memory or logical ability. What intuitively seems to be wrong may be classified as correct reasoning when one takes these limitations into account.

Harman assumes that there is a limit for the "storage capacity" of the beliefs, there are limits on the capacity of retrieval and it takes time and resources to add new beliefs to one's beliefs. This presupposes that some of one's beliefs are somehow explicitly represented in one's mind. The idea that beliefs are explicitly represented is quite controversial for natural agents (see for example the discussion about connectionist versus symbolic representation in [Bil98]). But for artificial agents it is clear that an explicit representation of beliefs exists. Not all of one's beliefs are explicitly stored; this would imply that one can only hold a finite amount of beliefs. Some beliefs are implicit in other beliefs. A belief being implicit does not mean that it is logically inferable from others. As an example of implicit belief that does not follow logically from others, Harman cites an example due to Dennett: one's implicit belief that elephants do not wear pajamas in the wild.

Besides the distinction between explicit and implicit beliefs, Harman introduces two other distinctions: beliefs can be unconscious or conscious and they can be occurrent or dispositional. A belief is unconscious if one is not aware of it and cannot easily become aware, otherwise it is available for consciousness. A believe is occurrent if it is somehow operative in guiding what one is doing or currently before one's consciousness, otherwise it is dispositional. All implicit beliefs are dispositional, but not all explicit beliefs are occurrent, only some of them are being used at a certain time.

Harman claims that one's beliefs are not always consistent. Sometimes one may realize that one has inconsistent beliefs but may not know how to "repair" it. Or one may lack time or logical ability to fix the inconsistency. In this case it may be rational to keep the inconsistency and avoid using it for inferences.

Harman does not consider (explicit) degrees of beliefs, but considers beliefs as an "all-or-nothing matter".

He distinguishes two different theories about belief dependencies and what could cause a belief to be given up or incorporated. The first, called *coherence theory*, claims that an agent does not keep track of all the belief dependencies. One does not have to remember the origin of one's beliefs, they are accepted as long as they are coherent with the rest of the agent's beliefs and as long as there is no evidence against them. On the other hand, according to the *foundations theory*, something can be believed as long as there is a valid justification associated to it. Justifications can only appeal to explicit beliefs. A justification must be acyclic and finite, which implies that there must be a set of "basic beliefs", beliefs that do not need any justification or are intrinsically justified.

Harman claims that foundations theory is closer to people's intuitions on what should be done, while coherence is closer to what people really do. In this sense, foundations theory could be seen as a normative theory while coherence theory would be a descriptive one. Harman cites psychological experiments that show that people do not keep track of all justifications for their beliefs. Foundations theory claims then that people should give up most of their beliefs, which sounds absurd. Keeping track of all dependencies of one's beliefs would clutter one's mind with unnecessary matters leaving no space for more important ones. For very idealized agents, foundations theory could be used as a normative theory indicating how to change one's beliefs. But for finite beings, Harman chooses the coherence theory as the norm.

The following example, extracted from [Har86] and based on psychological experiments shows that real agents do not keep all the dependencies between their beliefs:

> **Example 2:** "Consider Karen, who has taken an aptitude test and has just been told her results show she has a considerable aptitude for science and music but little aptitude for history and philosophy. This news does not correlate perfectly with her previous grades. She had previously done well not only in physics, for which her aptitude scores are reported to be high, but also in history, for which her aptitude scores are reported to be low. Furthermore, she had previously done poorly not only in philosophy, for which her aptitude scores are reported to be low, but also in music, for which her aptitude scores are reported to be high.
>
> After carefully thinking over these discrepancies, Karen concludes that her reported aptitude scores accurately reflect and are explained by her actual aptitudes; so she has an aptitude for science and music and no aptitude for history and philosophy; therefore her history course must have been an easy one, and also she did not work hard enough in the music course. She decides to take another music course and not to take any more history.
>
> ...
>
> Some days later she is informed that the report about her aptitude scores was incorrect! The scores reported were those of someone else whose name was confused with hers. Unfortunately, her own scores have now been lost. How should Karen revise her views, given this new information?" ([Har86], pages 33-34)

Foundations and coherence theories give completely different answers to this question. Although when asked most people think Karen should give up the beliefs based on the result of the test (like for example that her history course was an easy one), psychological experiments show that people in fact tend to keep these beliefs. This is due to the fact that limited reasoners usually lack the means for keeping track of all justifications for their beliefs.

Since Harman commits himself to coherence theory, he accepts that one should continue believing something as long as there is no evidence against it. But this applies only to what an agent fully believes. Harman distinguishes between fully accepted beliefs and working hypotheses. A working hypothesis can become a fully accepted belief if there is enough evidence for it and if it survives one's best attempts to refute it. A fully accepted belief may become a working hypothesis if there is evidence against it. Full acceptance ends inquiry — one should only fully accept a working hypothesis if one is convinced that further inquiry will not be worthwhile; inquiry will not, for example, reveal evidence against the working hypothesis which has not been investigated. Inquiry may be reopened later, if sufficiently strong evidence against a fully accepted belief appears. In this case, the fully accepted belief becomes a working hypothesis.

One reason why people tend to quickly transform working hypotheses into fully accepted beliefs is that keeping something as a working hypothesis means keeping track of dependencies, of how and where it was used in reasoning. Since people have limited storage capacity, inquiry must be limited and not everything can be seen as working hypotheses. The same argument against the foundations theory applies here.

The two competing goals of revision are: (i) to improve the coherence in one's beliefs and (ii) to change one's beliefs as little as possible. Harman proposes a very simplistic way of measuring changes: add the number of new (explicit) beliefs acquired to the number of old (explicit) beliefs given up. Certain beliefs may be accepted only for a moment, as intermediate steps of an argument. Once the conclusion has been reached, according to the coherence theory there is no more need to remember all steps that justified its acceptance. But by the simple measure given above, these momentary beliefs would be always counted twice, when accepted and when given up. Harman proposes not to include forgetting in the measure and maybe give less weight to momentary beliefs. But he argues that there is a good reason to try to minimize even short-term changes: they also consume resources and shorter arguments are always easier to handle than longer ones.

Harman states some principles that should be valid for any resource-bounded agent [Har86]:

1. *Clutter Avoidance:* "One should not clutter one's mind with trivialities" (page 12). This principles goes against trying to infer all consequences from one's beliefs. If the agent believes that p, he should not waste his resources inferring that $p \wedge p$, $p \vee q$, etc.

2. *Recognized Implication Principle:* "One has a reason to believe P if one recognizes that P is implied by one's view" (page 18). The agent may not be able to perform all logical inferences, but when he performs an inference and believes in the premises, he should accept the consequent. If the agent believes p and $p \rightarrow q$ and he infers q, then he should believe q.

3. *Recognized Inconsistency Principle:* "One has a reason to avoid believing things one recognizes to be inconsistent" (page 18). Since the agent may fail to perform some inferences, he may fail to realize that some of his beliefs are inconsistent. But once he realizes the inconsistency, the agent must try to avoid using it.

4. *Principle of Positive Undermining:* "One should stop believing P whenever one positively believes one's reasons for believing P are no good" (page 39). If the agent gets evidence against some of his beliefs (P) and he accepts the evidence, then he should stop (fully) believing P. Harman (page 44) presents as an example for this principle the story of William, who believes he has seen Connie, but later he finds out that she has an identical twin sister. His reasons for believing that he has seen Connie are no good.

5. *Principle of Conservatism:* "One is justified in continuing fully to accept something in the absence of a special reason not to" (page 46). This is a counterpart of the previous principle. One should only stop believing something if there is some evidence against it. If there is no such evidence, the agent should keep his beliefs.

6. *Interest Condition:* "One is to add a new proposition P to one's beliefs only if one is interested in whether P is true (and it is otherwise reasonable for one to believe P)" (page 55). The agent should not try to infer all kinds of things in which he is not interested. Harman defines some kinds of interest that determine what the agent should add to his beliefs: the interest in not being inconsistent, interest in the immediate environment, interest in facilitating reasoning (if the agent believes that knowing α would help him to obtain something he desires, he will be interested in α).

7. *Get Back Principle:* "One should not give up a belief one can easily (and rationally) get right back" (page 58). For instance, if the agent believes that his beliefs are inconsistent, he cannot solve the problem by just giving up this belief, since it can be immediately recovered.

In Chapter 4, we will develop a formal framework based on Harman's ideas. We distinguish between different kinds of beliefs, following Harman's distinction between explicit vs. implicit beliefs, accepted beliefs vs. working hypotheses, and occurrent vs. dispositional beliefs. In Section 4.4, we will interpret the principles above according to our framework.

2.2 Minimal Rationality

In this section we will present a summary of Cherniak's theory of minimal rationality [Che86].

Cherniak presented a theory for "minimal agents", i.e., agents that have the minimal abilities that are required for them to be called rational. His point of departure is the claim that any realizable agent is a finite object. Such finite "creatures" have limits in their cognitive resources, such as time, memory and deductive ability. His main hypothesis is that the definition of rationality which is universally assumed in philosophy is so idealized that it does not apply in any interesting way to human beings.

Cherniak defines a hierarchy of rationality concepts [Che86], on top of which appear ideal agents, with belief states that are deductively closed. Agents that are not able to perform any inference appear on the lowest level of the hierarchy. These last agents cannot be called rational. According to Cherniak, any rational agent (limited reasoners included) must satisfy at least the *minimal general rationality condition*: "If agent X has a particular belief-desire set, X would undertake some, but not necessarily all, of those actions that are apparently appropriate" (page 9). From this condition Cherniak derives the *minimal inference condition*: "X would make some, but not necessarily all, of the sound inferences from the belief set that are apparently appropriate" (page 10).

Cherniak defends that rational agents should use some heuristics to decide which inferences are worth making. No finite agent is able to calculate the deductive closure of his own beliefs, unless his logic is trivial and does not allow for interesting deductions. And even if he were able to find all consequences of his beliefs, he should not waste resources calculating useless logical consequences of his beliefs. According to Cherniak:

> "Each inquiry, deductive or otherwise, has costs; the heuristic imbecile would squander its limited cognitive resources on such valueless inferences and would therefore be paralyzed for appropriate inferences. Thus, for creatures with limited resources (such as time pressures), heuristic imbecility by itself entails complete logical incompetence." (page 11)

Cherniak tries to find an appropriate "threshold of minimal rationality" above which some seeming irrationalities could be explained by the limitations of the agent.

Another claim in the book is that not everything is considered at the same time. He distinguishes between beliefs that are "activated" or under consideration and beliefs that are inactive. An easy inference may be more difficult if not all the premises are activated. The subset of activated beliefs is subject to stricter rules of rationality. Cherniak defends that it is perfectly rational not to make all of the sound inferences one could possibly do.

Cherniak proposes that inferences are ordered according to their feasibility. But this order is not fixed; it varies from one agent to the other and also from one agent to the same agent at another instant.

Cherniak argues for adoption of a model of memory which is more psychologically adequate than the one commonly used in epistemology. Drawing on ideas of cognitive psychology, he claims that a belief state consists of two parts: a long-term memory and a short-term one. The long-term memory has no practical limitation in terms of size, but beliefs that are inactive (not in the short-term memory) cannot be used for reasoning. All reasoning is done in the short-term memory, which has a strict size limitation. In order to select which beliefs from the long-term memory have to become active and be copied into short-term memory, one needs a very efficient retrieval procedure. It is not feasible to examine all the beliefs in the long-term memory, at least not if there is any time constraint. To allow for efficient retrieval, long-term memory must be organized in "compartments" according to some relevance criteria. These criteria vary from one agent to the other. Beliefs are in the same compartment if they tend to be retrieved ("remembered", activated) at the same time. This explains why some inconsistencies are not immediately recognized, and why some "obvious" inferences are not made. Only beliefs that are active can be used as premises for an inference. In order to retrieve the relevant subset of the long-term memory, Cherniak defends what he calls "limited search":

> "... the required strategy must be better than chance, but need not, of course, be perfect; the latter would require prescience. Searches can be expected to fail frequently in either possible way: beliefs that turn out not to be currently relevant may be checked, and beliefs that turn out to be useful may be skipped." (page 65)

Thus, the fact that the search will not be perfect is the price to pay for quick retrieval.

> "We can now appreciate both the costs and the benefits of this strategy; prima facie, the resulting behavior can be characterized as departures from rationality, but on the assumption that exhaustive memory search is not feasible, such memory organization is advisable overall, in the long run, despite its costs." (page 67)

Cherniak attacks the commonly accepted idea that any agent must reason according to some sound and complete logic. He cites results obtained in computational complexity theory as important measures of feasibility of deductive procedures. He says that heuristics which would be considered as formally incorrect are perfectly rational, since "they are a means of avoiding computational paralysis while still doing better than guessing". He points that a distinction must be made between theoretical adequacy of a system and what he calls "practical adequacy". Moreover, formally incorrect deductive procedures may be sound for the range of the agent's needs.

According to Cherniak, an agent is not expected to check all facts that could undermine an assertion, but at least some of the relevant possibilities should be considered. The procedure of eliminating counter-possibilities is limited due to the resource-boundedness.

Like Harman, Cherniak is looking for a normative theory for limited agents, which must be different from theories for ideal agents, since one cannot expect agents to reason in a way that is impossible for their architectures. As in Harman, several concepts are introduced (active/inactive beliefs, feasible inferences, compartments), but no formalization is given. The framework presented in Chapter 4 gives a formal interpretation for these concepts.

2.3 Do the Right Thing

In the book of Russell and Wefald [RW91], an agent is defined as a "system that senses its environment and acts upon it". The theory concerns mainly artificial agents. The authors complain that most approaches to artificial intelligence consider only the quality of the outcome of a process and not the process itself. They defend an approach that allows meta-reasoning in order to decide the best way to achieve the best result which is possible given the resource limitations. An agent can be viewed in different ways — as a function mapping its inputs into its actions, as a program implementing this mapping or as the behavior of such a program.

They say that the problem in artificial intelligence is not to find better implementations for formal models, but to design formal models that are adequate, that take into account computational limitations. The problem is thus the lack of a theoretical basis for artificial intelligence.

One should not expect to create a computer program that always chooses immediately the best moves in any possible chess position. There is a necessary trade-off between optimality and speed that should be accounted for by theories of intelligent agents. They define *bounded optimality* as "doing as well as possible given what resources one has". They look for a theory for a limited rational agent. Such a limited agent will not always take the same decisions as the ideal agent, but will do so when its internal operations are so fast that they can be ignored.

The authors refer to Herbert Simon's work in the area of economics as showing that perfect deliberative rationality was not adequate as a theoretical basis for economics, where computational constraints and actual human behavior should be taken into account (cf. Chapters 7 and 8 of [Sim82]). The authors argue that a system should take the decision that would give "the highest return in the shortest time". They claim that most decisions taken by humans would be seen as irrational given the standard criteria for rationality.

2.4 Doxastic Performance

In this section we will present some remarks by Levi on non-ideal agents. Although the theory he presents in [Lev91] is a normative theory for idealized agents, he makes clear that what the theory predicts is what an agent should try to approximate, even if most agents will fail in matching the prediction.

According to the Peircean belief-doubt model[1] that Levi follows, at any instant an agent must have a clear distinction between what is conjecture and what is settled. Of course, this distinction can change with time.

In [Lev91], Levi distinguishes *doxastic commitments* from *doxastic performance*. He gives the following example in order to illustrate the difference (page 6):

> **Example 3:** At time t_1, X fully believes that Albany is north of New York and that being north of is a transitive relation.
>
> At time t_2, X comes to believe fully, in addition to what he believes fully at t_1, that Montreal is north of Albany.
>
> At time t_3, X comes to believe fully that Montreal is north of New York.

According to Levi, the change that occurs between t_1 and t_2 is a change of doxastic commitment while between t_2 and t_3 there is a change of doxastic performance. At t_2, the agent is already committed to believing that Montreal is north of New York, since this follows from his beliefs at this instant, but only at t_3 the agent recognizes this commitment and fulfills it. Only at this instant does his doxastic performance "catch up" with his doxastic commitments. The distinction between doxastic performance and doxastic commitments corresponds roughly to the distinction between the agent's psychological state and the ideal epistemic state represented by belief sets in AGM theory, as will be seen in Chapter 3.

An agent may fail to live up to his commitments, as long as he has a good excuse. As good excuses, Levi cites failure of memory, emotional distress, lack of mathematical and logical training, lack of access to encyclopedias, computing facilities, and other methods of repairing the agent's disabilities.

Levi (like the AGM theory explained in Chapter 3) is concerned with changes in the doxastic commitments of an agent. The doxastic commitments of an agent are logically closed. However, for Levi an agent is "no ideally situated, rational angel". An agent may fail to recognize some of the consequences of his doxastic commitments, a performance mistake.

An agent can undertake a doxastic commitment even though he is certain that his performance will not fulfill the commitments. The agent's duty is actually "to fulfill the commitments insofar as his computational abilities, memory, and emotional state permit and insofar as various technologies can be deployed at a reasonable cost" ([Lev91], page 38).

For Levi, his theory, as well as others based on belief-desire models, are intended to be normative or prescriptive and not explanatory or descriptive. He

[1] In [Pei77], Peirce claims that beliefs are like habits, in that they persist in the absence of doubt. Doubt is more like an "irritation" that one tries to get rid of by means of inquiry. One cannot obtain full knowledge, but should pursue inquiry until one gets an opinion which one believes to be true.

claims that an agent should change his performance in order to fulfill his doxastic commitments and reach doxastic equilibrium. The agent is obliged to pursue this equilibrium unless the changes needed are not feasible or too expensive. Levi does not clarify which changes are feasible and how one measures the costs related to changes in performance.

An agent can fall into two types of inconsistencies: (i) The doxastic performance may be inconsistent, and can thus be solved by "logical therapy". (ii) The doxastic commitment may be inconsistent, which can only be solved by contraction.

Levi claims that all demands on doxastic rationality that he defends must be seen as demands on doxastic commitment and not on doxastic performance. Human beings cannot satisfy the rationality demands for ideal agents (for instance, having a deductively closed set of beliefs) applied to performance.

Levi distinguishes two sorts of expansion: deliberate and routine expansion. Routine expansion is immediate and usually triggered by new information from the outside world. Deliberate expansion is inferential, where inferential includes inductive and abductive inferences. Both can lead to inconsistency, but if the agent is living up to his doxastic commitments, deliberate expansion will not. Routine expansion leads very often to an inconsistent state, which is later contracted to restore consistency.

Levi disagrees with Gärdenfors' view (in [Gär88]) that expansion into inconsistency is never legitimate and that the only admissible expansions are those in which the new piece of information is consistent with the previous beliefs. Levi argues for the common process of acquiring new information, expanding into inconsistency and then doubting both the new information and the previous beliefs. He uses the following example to illustrate this process:

> **Example 4:** "Suppose I am walking on 72nd Street and Columbus Avenue and see someone who is a dead ringer for Victor Dudman. I am sure that Dudman is safely ensconced in Sydney, New South Wales. But my initial confidence in my eyesight and in Dudman's appearance is reflected in a commitment to add the doxastic proposition that Dudman is on 72nd and Columbus to my state of full belief. I have expanded into inconsistency.
>
> (...)
>
> ... in the confrontation with the Dudman look-alike, I would look again. This would make little sense if I took for granted the deliverances of my first observation or if I remained convinced that Dudman is in Sydney. Looking again reveals that I have contracted by questioning part of my initial background assumptions as well as the testimony of my senses." (page 76)

Contraction that arises from the need to restore consistency after an expansion is a common operation.

2.5 Doing the Right Revision?

In this section we will describe the agent we will be talking about in the rest of this thesis. We are only concerned with belief change, and not with other kinds of reasoning. From the works presented in this chapter, only Harman deals with the particular problem of belief revision for non-ideal agents, which explains why his theory is more extensively discussed than the others. What we want is an agent that can decide what to do with incoming information. Because of resource limitations, the agent will not be able to always take the decision which would be expected from an ideal reasoner. But the right decision for a resource-bounded agent should be as close to the ideal as possible, as described in [RW91].

We will consider agents to be either natural or artificial beings with limits in memory and logical abilities. Following Harman, some of the agent's beliefs will be considered to be explicitly represented. The number of explicit beliefs is finite, but large enough so that there is no practical limit, as assumed by Cherniak for long-term memory. Agents may not be able to perform all kinds of inferences. Each agent will have an associated inference operation which gives all the inferences he can make in one step. The set of implicit beliefs (the doxastic commitments of the agent) is given by the closure of the explicit beliefs under the agent's inference function, i.e., by repeatedly applying the agent's inference operator until a fixed point is reached.

All reasoning happens inside a small part of the belief state, which we will call active beliefs. The active beliefs are roughly equivalent to what Harman calls occurrent beliefs and Cherniak calls short-term memory. In the set of active beliefs, besides the explicit beliefs which are in use, there are also conjectures (Levi), or working hypotheses (Harman), in which the agent does not yet fully believe, but which he is considering for acceptance.

The explicit beliefs (active and inactive parts) are organized like a web, where links denote some kind of relevance. The beliefs are thus grouped in "chunks" or "compartments", and beliefs in the same compartments tend to be activated together. Since the set of explicit beliefs may be very big, an efficient method for retrieving beliefs that will be activated is needed. In Chapter 5 we present a retrieval method based on logic. In Chapter 6, a more efficient model of retrieval is presented. This method does not guarantee that all relevant beliefs will be retrieved, but it is based on "quick and dirty" heuristics as defended by Cherniak. It allows the agent to do as well as he can, using what Russell and Wefald call an interruptible anytime method, which means the longer you let it run, the better the result it gives.

There is no consistency requirement. Like Levi, we allow for routine expansion into inconsistency. All new beliefs have to survive inquiry before being fully accepted.

In the next chapter, we will present the AGM theory for belief revision, which deals with the problem of belief change for ideal agents.

In Chapter 4, a formal framework for resource-bounded belief change based on some of the concepts introduced in this chapter. We will introduce basic operations that can be used as building blocks for more complex operations. This allows us to see the whole process of belief change instead of only the final outcome. Although the framework that we will present is general enough to account for changes in the agent's inference operation, we will not deal with this in the present work.

What we will present in Chapters 4, 5, and 6 is not another informal theory about limited reasoners, but a formalization of some ideas about the behavior of resource-bounded agents concerning revision. Some of the ideas formalized were (informally) proposed in the literature and were presented in this chapter.

It is important to remark that this is not the first attempt to formalize resource-bounded reasoning. Some other approaches will be discussed in Chapter 6. We will not discuss so-called resource logics [Gab96], since their motivation is quite different from ours: the resources that they limit are the number of times one can use the premises and when one can use the inference rules.

Although we are interested in eventually having a theory for resource-bounded belief change which can be efficiently implemented, in this thesis we will not analyze the computational complexity of the inference operators used. We will not examine any quantitative approach nor discuss the decision processes involved in belief change.

Chapter 3

The AGM Paradigm

In this chapter we are going to briefly present the by now standard model for belief revision. For a more detailed exposition, the reader is referred to [Gär88],[GR95] or [Han99b]. This model became known as *AGM* due to the initials of the authors of the seminal paper [AGM85], Carlos Alchourrón, Peter Gärdenfors, and David Makinson. We will first present the original AGM framework. Later in this chapter we will present some of its extensions that are relevant for our work. For the sake of simplicity, we present the results for the classical consequence operator Cn, although the original AGM results were proved for slightly more general consequence operators.[1]

The original AGM framework is a theory about how highly idealized rational agents should revise their beliefs when receiving new information. The agents are idealized in that they have unlimited memory and ability of inference. They are able to immediately close a set under deductive inference and to keep a belief set that is usually infinite. Beliefs are represented by formulas of a propositional language and the logic underlying the agents' reasoning is classical.

In the AGM model, belief states are represented by theories (possibly together with some selection mechanism), that is, sets of formulas K such that $Cn(K) = K$. These theories are called *belief sets*. There is only one inconsistent belief set, the set of all formulas of L, and this is sometimes represented by $K_{\perp}$.

Classical AGM theory can be seen as modeling only the changes that occur in belief sets due to new information about a static world (i.e, the world does not change, only the information that the agent has about it grows, cf. [KM92]). It does not model all changes of an agent's epistemic state, but only those that affect the belief set.

In AGM theory, there are three operations that can be performed on belief

[1] The consequence operator C used in [AGM85] is supposed to satisfy the following conditions: (i) $X \subseteq C(X)$; (ii) $C(X) = C(C(X))$; (iii) $C(X) \subseteq C(Y)$ if $X \subseteq Y$; (iv) $Cn(X) \subseteq C(X)$; (v) If $\alpha \in C(X)$, then for some finite $X' \subseteq X$, $\alpha \in C(X')$; and (vi) $\alpha \in C(X \cup \{\beta_1 \vee \beta_2\})$ whenever $\alpha \in C(X \cup \{\beta_1\})$ and $\alpha \in C(X \cup \{\beta_2\})$.

sets: contraction, expansion and revision. Contraction consists of giving up (at least) as many beliefs as it is needed so that the new belief set does not imply (and so, does not contain) a specified sentence. Expansion consists of adding new information to the belief set. If the old and the new information are not logically compatible, then the new belief state after expansion will be inconsistent. Revision is consistent incorporation of new information, i.e., if the input sentence is consistent, then the new belief set will be consistent (even if the old belief set was not). If necessary, consistency is obtained by deleting parts of the original belief set.

3.1 Postulates

Of the three AGM operations, only expansion is characterized in a unique way. When a belief set K is expanded with a proposition φ, the resulting set $K + \varphi$ is obtained by simply adding the new belief to the old belief set and taking the logical consequences of the resulting set:

$$K + \varphi = Cn(K \cup \{\varphi\}).$$

The name expansion is justified by the fact that $K \subseteq K + \varphi$.

Contraction and revision operations are not directly defined, but constrained by a set of rationality postulates.[2] For the contraction of a belief set K in relation to a sentence φ (denoted $K - \varphi$), six basic postulates are given [AGM85] ($\vdash$ is the consequence relation associated with Cn):

(K-1) $K - \varphi$ is a belief set (*closure*)
(K-2) $K - \varphi \subseteq K$ (*inclusion*)
(K-3) If $\varphi \notin K$, then $K - \varphi = K$ (*vacuity*)
(K-4) If not $\vdash \varphi$, then $\varphi \notin K - \varphi$ (*success*)
(K-5) $K \subseteq (K - \varphi) + \varphi$ (*recovery*)
(K-6) If $\vdash \varphi \leftrightarrow \psi$, then $K - \varphi = K - \psi$ (*equivalence*)

These postulates are supposed to capture the intuition behind the operation of giving up a belief in a rational way. Postulate **(K-1)** says that the result of contracting a belief set by a formula should again be a belief set. The next postulate assures that in an operation of contraction no new formulas are added to the initial belief set. If the formula to be contracted is not an element of the initial belief set, then by **(K-3)** nothing changes. Postulate **(K-4)** says that unless the sentence to be contracted is logically valid (and hence, an element of every theory), it is not an element of the resulting belief set. The recovery

[2] Originally, expansion was also defined by means of a set of postulates, but it can be completely determined by the postulates.

postulate **(K-5)** is the most controversial one [Mak87]. It says that a contraction should be recoverable, that is, that the original belief set should be recovered by expanding by the formula that was contracted. The last postulate assures that contraction by logically equivalent sentences produces the same output.

Besides these six basic postulates, there are two additional ones:

(K-7) $K - \varphi \cap K - \psi \subseteq K - (\varphi \wedge \psi)$
(K-8) If $\varphi \notin K - \varphi \wedge \psi$, then $K - \varphi \wedge \psi \subseteq K - \varphi$

Postulate **(K-7)** assures that beliefs that are both in $K - \varphi$ and in $K - \psi$ must be in $K - (\varphi \wedge \psi)$. If, when contracting by $\varphi \wedge \psi$ one has to give up φ, then by postulate **(K-8)**, any belief deleted when removing φ will also be deleted when removing $\varphi \wedge \psi$.

For the operation of revision, i.e., adding a new belief to a belief set in such a way that the resulting belief set is consistent if the new belief is consistent, the following postulates are presented [Gär88][3]:

(K*1) $K * \varphi$ is a belief set (*closure*)
(K*2) $\varphi \in K * \varphi$ (*success*)
(K*3) $K * \varphi \subseteq K + \varphi$ (*inclusion*)
(K*4) If $\neg\varphi \notin K$, then $K + \varphi \subseteq K * \varphi$ (*preservation*)
(K*5) $K * \varphi = L$ if and only if $\vdash \neg\varphi$ (*consistency*)
(K*6) If $\vdash \varphi \leftrightarrow \psi$, then $K * \varphi = K * \psi$ (*equivalence*)

Note that the same name is given to different, but related principles. This is common practice in the literature and we will maintain the usual names in this thesis. Postulate **(K*1)** says that the result of revising a belief set by a formula should be a belief set. Postulate **(K*2)** says that the formula by which a belief set is revised must be an element of the revised belief set. The next postulate assures that no information other than the input of the revision (together with its consequences) is added. If the new formula is consistent with the belief set, postulate **(K*4)**, together with **(K*3)** imply that nothing is given up, i.e., revision in this case equals expansion. Postulate **(K*5)** says that, unless the input formula is inconsistent, the revised belief set is consistent. Finally, postulate **(K*6)** says that revising a belief set by logically equivalent formulas should produce the same output.

Two extra postulates deal with revision by conjunctions:

(K*7) $K * (\varphi \wedge \psi) \subseteq (K * \varphi) + \psi$
(K*8) If $\neg\psi \notin K * \varphi$, then $(K * \varphi) + \psi \subseteq K * (\varphi \wedge \psi)$

[3]The postulates were already presented in [AGM85], but in a slightly different formulation.

Contraction and revision can be defined in terms of each other via the Harper or the Levi identities [Gär88]. Revising with a belief φ corresponds to contracting by the negation of φ and then expanding with φ:

$$K * \varphi = (K - \neg\varphi) + \varphi \qquad \text{(Levi identity)}$$

The thus defined operator $*$ will be called the revision operator associated with the contraction operator $-$. Analogously, the following identity defines a contraction operator associated with a revision operator:

$$K - \varphi = (K * \neg\varphi) \cap K \qquad \text{(Harper identity)}$$

3.1.1. THEOREM. *[Gär88] If $-$ is a contraction function satisfying* **(K-1)**-**(K-4)** *and* **(K-6)**, *then its associated revision function satisfies* **(K*1)**-**(K*6)**.

Note that the recovery postulate is not needed. Contraction operators that satisfy **(K-1)**-**(K-4)** and **(K-6)** but not **(K-5)** have been called *withdrawal* operators [Mak87].

3.1.2. THEOREM. *[AGM85][4] If $-$ is a contraction function satisfying* **(K-1)**-**(K-4)** *and* **(K-6)**, *then (a) if* **(K-5)** *and* **(K-7)** *are satisfied, the associated revision function satisfies* **(K*7)** *and (b) if* **(K-8)** *is satisfied, the associated revision function satisfies* **(K*8)**.

3.1.3. THEOREM. *[Gär88] If $*$ is a revision function satisfying* **(K*1)**-**(K*6)**, *then the associated contraction function satisfies* **(K-1)**-**(K-6)**.

3.1.4. THEOREM. *[Gär88] If $*$ is a revision function satisfying* **(K*1)**-**(K*6)**, *then (a) if* **(K*7)** *is satisfied, the associated contraction function satisfies* **(K-7)** *and (b) if* **(K*8)** *is satisfied, the associated contraction function satisfies* **(K-8)**.

3.2 Constructions

The postulates above do not determine unique contraction or revision operators for a belief set, but only restrict the set of possible such operators. In [AGM85] a particular construction is presented that, given a belief set and an input belief, returns the result of contracting (or revising) the given set by the input.

This construction makes use of the concept of a *remainder set*, the set of maximal subsets of a given set not implying a given sentence. Formally:

[4] Corrected version in [Fer99].

3.2.1. Definition. *[AM82] Let X be a set of formulas and α a formula. The* **remainder set** *$X \bot \alpha$ of X and α is defined as follows. For any set Y, $Y \in X \bot \alpha$ if and only if:*

- $Y \subseteq X$
- $Y \nvdash \alpha$
- *For all Y' such that $Y \subset Y' \subseteq X$, $Y' \vdash \alpha$.*

3.2.2. Observation. *[AGM85] If K is closed under logical deduction, then so are the elements of $K \bot \alpha$.*

It is assumed that there is some way of picking out the best (in some sense) elements of a remainder set. This is formalized by means of a selection function:

3.2.3. Definition. *[AGM85] A* **selection function** *for X is a function γ such that:*

- *If $X \perp \alpha \neq \emptyset$, then $\emptyset \neq \gamma(X \perp \alpha) \subseteq X \perp \alpha$.*
- *Otherwise, $\gamma(X \perp \alpha) = \{X\}$.*

A contraction is obtained by taking the intersection of the best subsets of K that do not imply α:

3.2.4. Definition. *[AGM85] For any sentence α, the operation of* **partial meet contraction** *over a belief set K determined by the selection function γ is given by:*

$K -_\gamma \alpha = \bigcap \gamma(K \perp \alpha)$

Partial meet revision is obtained from partial meet contraction and expansion by means of the Levi identity:

3.2.5. Definition. *Let K be a belief set and γ a selection function. For any sentence α, the operation of* **partial meet revision** *over K determined by γ is given by:*

$K *_\gamma \alpha = Cn(\bigcap \gamma(K \perp \alpha) \cup \{\alpha\})$

In their paper [AGM85], Alchourrón, Gärdenfors and Makinson show that partial meet constructions bear a very special relation to the contraction and revision postulates. They prove the following representation results:

3.2.6. Theorem. *[AGM85] Let $-$ be a function which, given a formula α, takes a belief set K into a new belief set $K - \alpha$. For every theory K, $-$ is a partial meet contraction operation over K if and only if $-$ satisfies the basic postulates (***(K-1)***-***(K-6)****) for contraction.*

3.2.7. Definition. *[AGM85] A selection function γ for K is said to be* **transitively relational over** K *if and only if there is a transitive relation $\leq$ over 2^K such that the following identity holds:*
$\gamma(K \perp \alpha) = \{X \in K \perp \alpha | X' \leq X \text{ for all } X' \in K \perp \alpha\}$.

3.2.8. Definition. *[AGM85] A partial meet function $-$ is* **transitively relational** *if and only if it can be determined by some transitively relational selection function.*

3.2.9. Theorem. *[AGM85] Let K be a theory and $-$ a partial meet contraction function over K. Then $-$ is transitively relational if and only if $-$ satisfies* **(K-7)** *and* **(K-8)**.

As special cases of the partial meet construction, there are two other constructions that had already been studied in the literature, maxichoice contractions and full meet contractions. In a maxichoice contraction, the selection function γ selects one single element of $K \perp \alpha$, in case this set is not empty. In [AGM85], it is shown that maxichoice contractions satisfy the following postulate:

(K-F) If $\beta \in K$ and $\beta \notin K - \alpha$, then $\beta \to \alpha \in K - \alpha$ (*fullness*)

In [AGM85] it is actually shown that **(K-F)** together with the basic postulates for contraction exactly characterize maxichoice contractions.

The following results show the undesirable effects of maxichoice operations:

3.2.10. Lemma. *[AM82] If $\alpha \in K$ and $K-\alpha$ is defined by means of a maxichoice contraction operation, then for any formula β, either $\alpha \vee \beta \in K - \alpha$ or $\alpha \vee \neg\beta \in K - \alpha$.*

3.2.11. Corollary. *[AM82] If a revision operation is defined from a maxichoice contraction by means of the Levi identity, then, for any α such that $\neg\alpha \in K$, $K*\alpha$ will be maximal, i.e., for every formula β, either $\beta \in K * \alpha$ or $\neg\beta \in K * \alpha$.*

Suppose I believe p (that Buenos Aires is the capital of Brazil) and have no idea about q (that the King of France is bald). Finding out that $\neg p$ is the case and revising my belief set using a revision based on maxichoice contraction means that I will make a decision as to q or $\neg q$.

In full meet contraction, the whole set $K \perp \alpha$ (if not empty) is selected. Note that full meet contraction is the only one of these three operations that does not require extra-logical information, since no real choice is needed.

In [AGM85], it is shown that full meet contractions are exactly characterized by the basic postulates together with the following postulate:

(K-I) For all α and β, $K-(\alpha \wedge \beta) = (K-\alpha) \cap (K-\beta)$ (*intersection condition*)

The following results show that full meet contraction deletes beliefs that intuitively should be preserved:

3.2.12. LEMMA. *[AM82] If $\alpha \in K$ and $K - \alpha$ is defined by means of a full meet contraction operation, then $\beta \in K - \alpha$ if and only if $\beta \in K$ and $\beta \in Cn(\neg\alpha)$.*

3.2.13. COROLLARY. *[AM82] If a revision operation is defined from full meet contraction by means of the Levi identity, then, for any α such that $\neg\alpha \in K$, $K * \alpha = Cn(\alpha)$.*

Suppose I believe that p (Buenos Aires is the capital of Brazil) and that q (there is no King of France). When I learn $\neg p$ and revise my belief set using a revision operation based on full meet contraction, I give up the belief that there is no King of France.

Another way to construct a contraction operation is based on the notion of *epistemic entrenchment*. A sentence α is less entrenched than sentence β in belief set K if it is easier to give up α than to give up β. According to Gärdenfors, when a belief set K is contracted or revised, the sentences which are given up are the ones with the lowest entrenchment in K [Gär88]. This claim is criticized by Rott in [Rot99]. Rott shows that according to the new input, a sentence which does not have the lowest entrenchment may be given up (cf. example 5).

Given a belief set K, Gärdenfors proposed five postulates that an epistemic entrenchment order should satisfy ($\alpha \leq \beta$ should be read as "β is at least as entrenched as α in K"):

(EE1) For any α, β and γ, if $\alpha \leq \beta$ and $\beta \leq \gamma$, then $\alpha \leq \gamma$. (*transitivity*)
(EE2) For any α and β, if $\alpha \vdash \beta$, then $\alpha \leq \beta$. (*dominance*)
(EE3) For all α and β in K, $\alpha \leq \alpha \wedge \beta$ or $\beta \leq \alpha \wedge \beta$. (*conjunctiveness*)
(EE4) When $K \neq L$, $\alpha \notin K$ if and only if $\alpha \leq \beta$ for all β. (*minimality*)
(EE5) If $\beta \leq \alpha$ for all β, then $\vdash \alpha$. (*maximality*)

The following two conditions from [GM88] show how to determine an epistemic entrenchment ordering given a contraction operator and vice versa:

(C$\leq$) $\alpha \leq \beta$ if and only if $\alpha \notin K - (\alpha \wedge \beta)$ or $\vdash \alpha \wedge \beta$.

(C$-$) $\beta \in K - \alpha$ if and only if $\beta \in K$ and either $\alpha < \alpha \vee \beta$ or $\vdash \alpha$.

The conditions are used in the following theorems, which show that epistemic entrenchment orderings obtained with condition **(C$\leq$)** and contraction operations obtained via condition **(C$-$)** satisfy the desired properties:

3.2.14. THEOREM. *[GM88] If an ordering $\leq$ satisfies* **(EE1)**-**(EE5)**, *then the contraction function determined by* **(C$-$)** *satisfies* **(K-1)**-**(K-8)** *as well as condition* **(C$\leq$)**.

3.2.15. THEOREM. *[GM88] If a contraction operator $-$ satisfies* **(K-1)**-**(K-8)**, *then the ordering $\leq$ that is determined by* **(C$\leq$)** *satisfies* **(EE1)**-**(EE5)** *as well as condition* **(C$-$)**.

Example 5:[Rot99] Let $K = Cn(\{\neg p, q\})$ and let $\varphi \leq \psi$ if and only if every formula of $\{\bot, \neg p \wedge q, \neg p, p \rightarrow q, \top\}$ which implies φ also implies ψ. The ordering $\leq$ satisfies **(EE1)-(EE5)**. Using **(C$-$)** to define a contraction operation $-$ and letting $*$ be its associated revision operator, we see that although q is strictly more entrenched than $q \rightarrow \neg p$, $q \in K * \neg p$ and $q \rightarrow \neg p \notin K * \neg p$.

3.3 Shortcomings

As we have seen, the AGM theory provides a very elegant framework in which to talk about belief revision. However, it presents some problems if what we are looking for is a realistic approach, one that could eventually be implemented. Besides the fact already mentioned that the agents modeled are highly idealized, there are some more concrete facts restricting the application of the theory.

- Using belief sets to represent belief states means that we in general have to deal with infinite sets, which are obviously not adequate for computational implementations.

- Since belief sets are closed under logical consequence, there exists only one inconsistent belief set, that contains the whole language. This means that whenever new information is added to the belief set, consistency has to be checked with regard to the whole belief set, to avoid the risk of losing all the information. This is a very expensive requirement.

- All the AGM operations (and the representation of belief sets) make extensive use of the consequence operation. This makes the change operations, from the computational point of view, extremely expensive.

- The partial meet constructions make use of extra-logical information, in the form of a selection function that selects the "best" elements of a set. The AGM paradigm does not say how such information can be obtained for a revised belief set, i.e., how to model iterated belief revision. Several models for iterated belief change have been proposed extending or revising the AGM theory ([Bou96, DP97, AB99]).

- In revision, incoming information always has higher priority than previous beliefs. But once a belief has been incorporated into the belief set, it immediately loses its special status.

Why are these ideal agents interesting for us? Studying their behavior gives us an aim and a way to evaluate more realistic theories. The AGM paradigm tells us what is to be expected from perfect reasoners. Given this ideal case, we should try to do as well as we can to match this behavior. This includes using heuristics in order to compensate for limited resources [RW91].

In the next section we will present some of the alternatives suggested in the literature in order to maintain the elegance of AGM theory but addressing some problems not addressed by it.

3.4 Alternatives and Refinements

In this section we present some work that was developed to deal with the problems of the AGM theory and which will be used in the next chapters.

3.4.1 Belief Bases

The use of logically closed sets to represent beliefs has received many criticisms, specially from authors concerned with the computability of the theory of belief change. Besides the fact that belief sets are too large to be represented, they make no distinction between basic beliefs and those which were inferred from them. Nebel [Neb90] proposed that instead of always considering a belief set as a whole, one should consider a *belief base*, a finite set containing the central beliefs, and take its logical closure when necessary.

Consider the following example:

> **Example 6:** An agent receives the information that p is the case. Since this is consistent with his previous beliefs, he decides to expand his belief set with p. Since belief sets are logically closed, for every formula α of the language, the agent's belief set K after the expansion with p contains $p \vee \alpha$. Let q be such that p does not bring any new information about q and such that neither q nor $\neg q$ are in the belief set, i.e., the agent is ignorant about q. Suppose now that the agent learns that actually $\neg p$ holds. Since neither $p \vee q$ nor $p \vee \neg q$ imply p, one of these beliefs may be retained in the revision of K by $\neg p$. And then, since $\neg p$ and either $p \vee q$ or $p \vee \neg q$ are in the new belief set, either the agent comes to believe that q holds or he comes to believe that $\neg q$ holds!
>
> This inference seems to be unjustified because, intuitively, the belief that $p \vee q$ is a merely derived belief and has no independent standing. It is in the belief set only because p is. When p is given up, $p \vee q$ should be as well. This problem can be solved with the use of belief bases instead of belief sets.

However, the term "belief base" has been used by two different communities with different meanings. Most authors interested in implementing AGM belief revision [Neb90, Wil94, Dix94] use belief bases to represent the belief set of an agent. A belief base in these approaches is a finite axiomatization of the theory (belief set) representing the agent's commitments. The fact that a formula of the belief set is an element of the belief base does not substantively distinguish it from other formulas that are not part of the belief base but follow logically from it. There are just two possibilities for a formula: either it follows from the belief base, and so the agent is committed to it, or it does not. Authors following this approach usually try to avoid redundancies in the belief base. In databases, for example, this may be a desirable property, since it makes alterations of the data easier. On the other hand, adding new information becomes much more complex, since one has to check for redundancies.

One example of this line of research is the implementation of AGM-style operators given in [Dix94]. In this implementation, a finite set of beliefs is kept as a base for the real set of the agent's beliefs that is given by the closure of the finite representation. In order to minimize the space used, a belief is deleted from the base if it follows from other beliefs. The base is then a minimal set that generates the beliefs of the agent. The same belief set can be generated by different belief bases and in the implementation, the choice between these bases does not reflect any intuitive notion of which beliefs are basic and which are not; except for efficiency considerations, there is no reason for choosing one base rather than another. Moreover, always keeping an inclusion-minimal base which can generate the belief set can cause beliefs to be recomputed every time that they are accessed, which, even though saving space in the belief base, leads to unnecessary waste of computing efforts. Another problem of keeping always a minimal base is that inconsistencies are not allowed: since all beliefs derived from others are deleted, if an inconsistency is present, then everything else is deleted. (This agrees with the use of logically closed sets: if there is any inconsistency, all information is lost.)

The approach that we follow in this work is in line with another use of the term "belief base", defended for example by [Fuh91, Rot96, Han99b]. According to this line, a belief base is a set of formulas representing those beliefs of an agent that have independent standing. Beliefs that follow from the ones in the belief base but are not part of it have a different status, being merely derived beliefs. As an example, consider the bookstore database containing John Smith's date of birth, as introduced in the beginning of Chapter 1. This can be considered an explicit belief, while John's age would be a merely derived belief. The set of formulas to which the agent is committed, is formed by taking the logical closure of the agent's belief base. The distinction between the belief base and its logical closure is similar to the distinction between explicit and implicit beliefs found in the literature. However, the way this distinction is made varies a lot from author to author. The paper by Fagin and Halpern [FH88], for example, presents three different logics for representing beliefs, including a Logic of General Awareness where the set of an agent's implicit beliefs is logically closed, while the explicit beliefs are those elements of this set which the agent is aware of. Harman [Har86] uses the terms in a different way: he calls explicit those beliefs which are somehow explicitly represented and implicit those beliefs which can be inferred from the explicit ones, not necessarily by logical means.

While in the first notion of belief base the stress is on keeping a non-redundant base, and thus requiring a non-trivial operation of expansion, in the second notion what matters is the justification of an agent for believing the formulas of the base. In this case, adding new information to the belief base via expansion is a trivial operation. The complexity arises when one has to give up beliefs (contraction) or use the beliefs for reasoning, since they may well be inconsistent.[5]

Nebel [Neb98] calls the revisions in the first approach *base revision schemes* and those in the second *base-generated revisions*. The result of a base revision scheme does not need to be a belief base, it may be a belief set. In the case of base-generated revisions, the construction of the revision operators is applied on a belief base and gives as a result another belief base. Gärdenfors and Rott [GR95] call this categorical matching. This is an important condition for iterating revision operators.

Note that the different meanings attached to the term "belief base" do not give rise to different extensions of the term: both communities use subsets of the belief sets which generate the complete belief set by logical closure. The difference is in the reasons for choosing one particular subset rather than another. While one community attaches special epistemic status to the beliefs in the belief base, the other does not and is only concerned with the minimality of the base.

Both notions of belief bases have substantial advantages in terms of computability [Neb98], and their increased expressive power as compared to belief sets can be used to represent important features of actual belief systems [Han92a]. On the other hand, belief sets have important advantages in terms of simplicity.

One of the most important difference between belief sets and belief bases is that the latter allow for a simple and direct representation of the fact that there can be different inconsistent belief states. Consider, for instance, the following two belief bases:

$$B_1 = \{p, \neg p, q_1, q_2, q_3, q_4, q_5\}$$
$$B_2 = \{p, \neg p, \neg q_1, \neg q_2, \neg q_3, \neg q_4, \neg q_5\}$$

These are distinct belief bases. We can suppose q_1 but not $\neg q_1$ to be endorsed according to B_1, whereas the opposite holds for B_2. However, on the belief set level these distinctions are lost since $Cn(B_1) = Cn(B_2) = L$. More generally speaking, belief bases but not belief sets allow us to distinguish between different inconsistency-containing belief states.

[5] In [Rot96], two ways of using belief bases are distinguished. Either the belief base is a finite axiomatization of a belief set, which can be obtained by applying Cn on the base (the horizontal perspective), or the belief base is an arbitrary set of formulas, possibly inconsistent, and the commitments of the agent are obtained by applying a more sophisticated inference operator on the base (the vertical perspective). Our approach is closer to the vertical perspective.

In what follows, we will present several constructions for belief base change operations together with their axiomatizations. Unlike AGM operations, which were first studied through rationality postulates, these operations on bases were first defined by the constructions and only then axioms characterizing them were found.

Base expansion is simply set union. Given a belief base B and a formula α, $B + \alpha = B \cup \{\alpha\}$. We will use the same symbol for expansions of belief sets and belief bases and its meaning will be clear from the context.

An operator of partial meet base contraction is obtained by applying the same construction as the one seen in Definition 3.2.4 to a belief base, instead of a belief set:

3.4.1. DEFINITION. *[AGM85, Han91] The* **partial meet base contraction operator** *on B based on a selection function γ is the operator $\dot{-}_\gamma$ such that for all sentences α:*
$B \dot{-}_\gamma \alpha = \bigcap \gamma(B \perp \alpha)$.

Hansson has given the following axiomatic characterization of partial meet base contraction:

3.4.2. THEOREM. *[Han92b] An operator $\dot{-}$ is an operator of partial meet base contraction on B if and only if:*

- *If $\alpha \notin Cn(\emptyset)$, then $\alpha \notin Cn(B \dot{-} \alpha)$ (success)*
- *$B \dot{-} \alpha \subseteq B$ (inclusion)*
- *If $\beta \in B \setminus (B \dot{-} \alpha)$, then there is some B' such that $B \dot{-} \alpha \subseteq B' \subseteq B$, $\alpha \notin Cn(B')$ and $\alpha \in Cn(B' \cup \{\beta\})$ (relevance)*
- *If for all subsets B' of B, $\alpha \in Cn(B')$ if and only if $\beta \in Cn(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)*

In Chapter 5, we will present a generalization of the representation results mentioned in this section, where instead of the classical consequence operation Cn an arbitrary inference operation is used. For each of the theorems we show which are the properties of the inference operation that have to be assumed. Proofs for these generals results can be found in Appendix C.

The classic AGM operation of revision (first developed for belief sets) can be reduced to contraction by $\neg\alpha$ followed by expansion by α [AGM85]. In [Han92b] this was called "internal revision" and an alternative procedure for belief bases, "external revision" was proposed. It consists in first expanding the belief base by α and after that contracting by $\neg\alpha$. External revision cannot be meaningfully applicable to belief sets since, if a belief set that contains $\neg\alpha$ is expanded by α, then the outcome will be equal to the whole language, so that all distinctions are lost.[6]

In an operation of internal partial meet revision by α, the belief base is first partial meet contracted by $\neg\alpha$ and then expanded by α:

3.4.3. DEFINITION. *[AGM85, Han92b] The* **internal partial meet base revision** *of B based on a selection function γ is the operator $\mp_\gamma$ such that for all sentences α:*

$B \mp_\gamma \alpha = \bigcap \gamma(B \perp \neg\alpha) \cup \{\alpha\}$

The following theorem characterizes this operation:

3.4.4. THEOREM. *[Han92b]*

An operator $\mp$ is an operator of internal partial meet revision of B if and only if:

- *If $\neg\alpha \notin Cn(\emptyset)$, then $\neg\alpha \notin Cn(B \mp \alpha)$ (non-contradiction)*
- *$B \mp \alpha \subseteq B \cup \{\alpha\}$ (inclusion)*
- *If $\beta \in B \setminus B \mp \alpha$, then there is some B' such that $B \mp \alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin Cn(B')$ but $\neg\alpha \in Cn(B \cup \{\beta\})$ (relevance)*
- *$\alpha \in B \mp \alpha$ (success)*
- *If for all $B' \subseteq B$, $\neg\alpha \in Cn(B')$ if and only if $\neg\beta \in Cn(B')$, then $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$ (uniformity)*

In an operation of external partial meet revision by α, the belief base is first expanded with α and then partial meet contraction by $\neg\alpha$ takes place:

3.4.5. DEFINITION. *[Han92b] The* **external partial meet revision** *on B determined by a selection function γ is the operator $\pm_\gamma$ such that for all sentences α:*

$B \pm_\gamma \alpha = \bigcap \gamma((B \cup \{\alpha\}) \perp \neg\alpha)$

The following theorem characterizes this operation:

3.4.6. THEOREM. *An operator $\pm$ is an operator of external partial meet revision on B if and only if:*

- *If $\neg\alpha \notin Cn(\emptyset)$, then $\neg\alpha \notin Cn(B \pm \alpha)$ (non-contradiction)*

[6] This could be avoided by allowing different selection mechanisms that do not depend exclusively on the present belief set, but also store somehow its history, so that the inconsistent set would be contracted in different ways depending on the previous belief set and input that led to it.

- $B \pm \alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If* $\beta \in B \setminus B \pm \alpha$*, then there is some* B' *such that* $B \pm \alpha \subseteq B' \subseteq B \cup \{\alpha\}$ *such that* $\neg\alpha \notin Cn(B')$ *and* $\neg\alpha \in Cn(B' \cup \{\beta\})$ *(relevance)*
- $\alpha \in B \pm \alpha$ *(success)*
- *If* α *and* β *are elements of* B *and it holds for all* $B' \subseteq B$ *that* $\neg\alpha \in Cn(B')$ *if and only if* $\neg\beta \in Cn(B')$*, then* $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ *(weak uniformity)*
- $B + \alpha \pm \alpha = B \pm \alpha$ *(pre-expansion)*

The following result shows that usually internal and external revision do not coincide:

3.4.7. Observation. *[Han92b] 1. Uniformity does not hold in general for external partial meet revision.*
2. Pre-expansion does not hold in general for internal partial meet revision.

The following example extracted from [Han99b] shows that internal and external revision may give different results[7]:

> **Example 7:** Let $B = \{p \to r, p \to \neg r\}$.
>
> (a) internal revision: $B \bot \neg p = \{\{p \to r\}, \{p \to \neg r\}\}$, hence,
> for $\gamma_1(B \bot \neg p) = \{\{p \to r\}\}$, we have $B \mp_{\gamma_1} p = \{p \to r, p\}$.
>
> (b) external revision: $B + p \bot \neg p = \{\{p \to r, p\}, \{p \to \neg r, p\}\}$, hence,
> for $\gamma_2(B + p \bot \neg p) = \{\{p \to \neg r, p\}\}$, we have $B \pm_{\gamma_2} p = \{p \to \neg r, p\}$.

The operations of external and internal revision model two different kinds of belief change. It is not always clear which operation is more adequate for formalizing one particular example. Levi (as will be discussed in Section 2.4) defends expansion into inconsistency followed by contraction, i.e., external revision [Lev91] as a natural operation. Hansson offers the following examples to show that sometimes one kind of operation is to be preferred over the other [Han99b]:

> **Example 8:** A man has died in a remote place in which only two other persons, Adam and Bob, were present. Initially, the public prosecutor believes that neither Adam nor Bob has killed him. Thus her belief base contains $\neg a$ and $\neg b$. For simplicity, we may assume

[7] Hansson uses here *global selection functions*, which take as arguments a belief base and a formula. In this example, the two selection functions are clearly not based "on the same ordering". It would be interesting to find out conditions on global selection functions so that internal and external revision coincide.

> that her belief base is $\{\neg a, \neg b\}$. Suppose now that she receives the information that the man has been murdered $(a \vee b)$ and that Adam has previously been convicted of murder several times (c). She revises her beliefs with $((a \vee b) \wedge c)$. Her new belief base after the revision is $\{\neg b, (a \vee b) \wedge c\}$. If instead of Adam, Bob had been convicted of murder (d), the new belief base would have been $\{\neg a, (a \vee b) \wedge d\}$. (page 208)

This example cannot be modeled as internal partial meet revision, since $\{\neg a, \neg b\} \perp (\neg((a \vee b) \wedge c)) = \{\neg a, \neg b\} \perp (\neg((a \vee b) \wedge d))$, and thus, it is impossible to decide for giving up $\neg a$ or $\neg b$ according to the extra evidence c or d. In external belief revision the information is added before the contraction and can play a role in the selection process.

> **Example 9:** I believe that Brian is a Catholic priest (a). I also believe that if Brian is married, then he is not a catholic priest $(b \to \neg a)$. If I find out that Brian is married (b), then I give up my belief that he is a priest in order to accommodate the new information. But suppose I had already expanded my belief base with the new information before noticing it was inconsistent. Then it may be the case that after revision I still believe that Brian is a priest but I gave up the belief that if he is married then he is not a priest. (page 207)

In this case, external belief revision cannot be used, since it satisfies pre-expansion and the fact that b already was in the belief base cannot make any difference.

The main difference seems to be, as noted by Hansson, whether the new piece of information is immediately accepted or not. In cases where the agent is sure about accepting it, external revision seems more adequate. If the agent has to think about it, then internal revision takes place.

Hansson [Han94] introduced another construction for contraction operators, called *kernel contraction*, which is a generalization of the operation of safe contraction defined by Alchourrón and Makinson in [AM85]. The idea behind kernel contraction is that, if we remove from the belief base B at least one element of each α-kernel (minimal subset of B that implies α), then we obtain a belief base that does not imply α [Han94]. To perform these removals of elements, we use an incision function, i.e., a function that selects at least one sentence from each kernel.

3.4.8. Definition. *[Han94] The* **kernel** *operation* $\perp\!\!\!\perp$ *is the operation such that for every set B of formulas and every formula α, $X \in B \perp\!\!\!\perp \alpha$ if and only if:*

1. $X \subseteq B$

2. $\alpha \in Cn(X)$

3. *for all* Y, *if* $Y \subset X$ *then* $\alpha \notin Cn(Y)$

The elements of $B \perp\!\!\!\perp \alpha$ *are called* α*-kernels.*

3.4.9. DEFINITION. *[Han94] An* **incision function** *for* B *is any function* σ *such that for any formula* α*:*

1. $\sigma(B \perp\!\!\!\perp \alpha) \subseteq \bigcup(B \perp\!\!\!\perp \alpha)$, *and*

2. *If* $\emptyset \neq X \in B \perp\!\!\!\perp \alpha$, *then* $X \cap \sigma(B \perp\!\!\!\perp \alpha) \neq \emptyset$.

3.4.10. DEFINITION. *[Han94] Let* σ *be an incision function. The* **kernel contraction** *on* B *determined by* σ *is the operation* $\dot{-}_\sigma$ *such that for all sentences* α*:*

$B \dot{-}_\sigma \alpha = B \setminus \sigma(B \perp\!\!\!\perp \alpha)$

An axiomatic characterization of kernel contraction was obtained in [Han94]:

3.4.11. THEOREM. *The operator* $\dot{-}$ *is an operation of kernel contraction on* B *determined by some incision function if and only if:*

- *If* $\alpha \notin Cn(\emptyset)$, *then* $\alpha \notin Cn(B \dot{-} \alpha)$ *(success)*
- $B \dot{-} \alpha \subseteq B$ *(inclusion)*
- *If* $\beta \in B \setminus B \dot{-} \alpha$, *then there is some* $B' \subseteq B$ *such that* $\alpha \notin Cn(B')$ *and* $\alpha \in Cn(B' \cup \{\beta\})$ *(core-retainment)*
- *If for all subsets* B' *of* B, $\alpha \in Cn(B')$ *if and only if* $\beta \in Cn(B')$, *then* $B \dot{-} \alpha = B \dot{-} \beta$ *(uniformity)*

From Theorems 3.4.2 and 3.4.11, since *relevance* implies *core-retainment*, it follows immediately that every base partial meet contraction operation can be defined as a kernel operation. The converse, however, does not hold:

3.4.12. OBSERVATION. *1. If* $\dot{-}$ *is a partial meet contraction operator, then it is a kernel contraction operator.*

2. It does not hold in general that if $\dot{-}$ *is a kernel contraction operator, then it is a partial meet contraction operator.*

Proof: 1. Follows directly from Theorems 3.4.2 and 3.4.11.

2. To see that not all kernel contraction operators can be defined as partial meet contraction operators, consider the following example: Let $B = \{p, q, p \vee q \rightarrow r\}$. We have that $B \perp r = \{\{p, q\}, \{p \vee q \rightarrow r\}\}$, and thus, there are only three possible values for $\gamma(B \perp r)$: $\{\{p, q\}\}$, that would give $B \dot{-}_\gamma r = \{p, q\}$; $\{\{p \vee q \rightarrow$

$r\}\}$, that would give $B\dot{-}_{\gamma}r = \{p \vee q \rightarrow r\}$, and $\{\{p,q\},\{p \vee q \rightarrow r\}\}$, that would give $B\dot{-}_{\gamma}r = \emptyset$. On the other hand, since $B \perp\!\!\!\perp r = \{\{p, p \vee q \rightarrow r\},\{q, p \vee q \rightarrow r\}\}$, we may have $\sigma(B \perp\!\!\!\perp r) = \{q, p \vee q \rightarrow r\}$, and so $B\dot{-}_{\sigma}r = \{p\}$.

□

Internal and external revisions can also be obtained by combining expansion and kernel contraction. In Section 5.2.3, we give formal definitions and axiomatic characterizations of internal and external kernel revision in a more general form.

Even though the use of belief bases presents already the advantage that belief states are finitely represented and thus, more amenable to computational treatment, as we have seen, the operations of belief change defined for belief bases still make use of the operation of logical closure.[8] Kernel operations are more efficient than partial meet operations, since they look for minimal subsets instead of maximal, but they still rely on checking derivability for several subsets of the belief base.

3.4.2 Non-Prioritized Belief Revision

As we have seen, one of the shortcomings of AGM theory is that incoming information always has the highest priority. Some formalisms for non-prioritized belief-revision ([Han97b, Han99a]) have been developed in which this is not the case, i.e., incoming information may be rejected. In this section we will briefly summarize some of these approaches.

The operation of AGM revision is applied only after the agent decided to accept the new piece of information. Since the agent is a perfect reasoner, if he chooses to reject the new piece of information, then nothing has to be done to the belief set. According to [Lev91], Gärdenfors (in [Gär88]) never addresses the question of why an agent is justified in adding a certain proposition to his belief set.[9]

Following this line of thought, the simplest idea for an operator of non-prioritized belief revision is to divide it in two parts: first analyze the input to check whether it is acceptable or not and then, in case it is acceptable, perform AGM revision. This is the idea behind screened revision [Mak97].

In Galliers' proposal for a theory of autonomous belief revision [Gal92], the agents can decide whether to accept the information or not. Galliers defines the result of a revision operation on a belief base as a set of belief bases, including the original (non-revised) belief base, ordered according to some criteria. The

[8] Alternative approaches based on epistemic entrenchments (E-bases [Rot91]; "ensconcements" [Wil94]) rely on a well behaved pre-order of the beliefs in the base to construct operations of belief change. This pre-order has to satisfy some logical constraints, as we have seen in the end of Chapter 3. Given the pre-order, the operations become less complex, but constructing and maintaining this pre-order can be computationally very costly.

[9] Gärdenfors avoids the question by identifying the input with the requirement that it should be accepted ([Gär88], page 49).

maximal (according to these criteria) set is then preferred.

Two additional operations of change on belief bases were introduced in [Han91] and [Han97a]: consolidation and semi-revision. Consolidation consists in making an inconsistent belief base consistent. Semi-revision is an operation that may, depending on the input sentence, either accept or reject new information.

Consolidation can be modeled as contraction by falsum. Therefore, the two contraction operators introduced in the previous section provide us with two consolidation operators.

The idea behind kernel consolidation is that, if we remove from the belief base at least one element of each inconsistent kernel (inclusion-minimal subset of the base that implies $\perp$), then we obtain a consistent belief base.

3.4.13. DEFINITION. *[Han97a] Let σ be an incision function. Then the* **kernel consolidation operation** *for B determined by σ is the operation $!_\sigma$ such that:*
$B!_\sigma = B \setminus \sigma(B \perp\!\!\!\perp \perp)$.

3.4.14. THEOREM. *[Han97a] An operation $!$ is an operation of kernel consolidation for B if and only if:*

- $\perp \notin Cn(B!)$ *(consistency)*
- $B! \subseteq B$ *(inclusion)*
- *If $\alpha \in B \setminus (B!)$, then there is some X such that $X \subseteq B$, $\perp \notin Cn(X)$ and $\perp \in Cn(X \cup \{\alpha\})$ (core-retainment)*

Consolidation can also be constructed from partial meet contraction.

3.4.15. DEFINITION. *[Han91] The* **partial meet consolidation operator** *for B based on a selection function γ is the operator $!_\gamma$ such that:*
$B!_\gamma = \bigcap \gamma(B \perp \perp)$

3.4.16. THEOREM. *[Han91] An operation $!$ is an operation of partial meet consolidation on B based on some selection function if and only if:*

- $\perp \notin Cn(B!)$ *(consistency)*
- $B! \subseteq B$ *(inclusion)*
- *If $\alpha \in B \setminus (B!)$, then there is some X such that $B! \subseteq X \subseteq B$, $\perp \notin Cn(X)$ and $\perp \in Cn(X \cup \{\alpha\})$ (relevance)*

Semi-revision differs from revision in being non-prioritized, i.e., the incoming information may be either accepted or rejected. In analogy to AGM revision, that can be defined in terms of contraction and expansion, semi-revision can be defined in terms of consolidation and expansion via the following identity [Han97a]:

$B?\alpha = (B + \alpha)!$

Hence, semi-revision consists of two steps: first the belief α is added to the base, and then the resulting base is consolidated. The two consolidation operators introduced above give rise to two semi-revision operators.

As an example, consider the following situation:

> **Example 10:** Suppose I arrive at my friend Paul's house for a visit and see that the lights are on (q). I believe that when the lights are on, Paul is at home ($q \rightarrow p$). Suppose now that I ring the bell and nobody answers. I may conclude that Paul is not at home ($\neg p$). I have expanded into an inconsistent belief base. I must give up some belief in order to regain consistency. Since I trust my senses, I give up the belief that $q \rightarrow p$. The whole process is an operation of semi-revision.

3.4.17. Definition. *[Han97a] The* **kernel semi-revision** *of B based on an incision function σ is the operator $?_\sigma$ such that for all sentences α:*
$B?_\sigma\alpha = (B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp \bot)$

3.4.18. Theorem. *[Han97a] An operator $?$ is an operator of kernel semi-revision if and only if for all sets B of sentences:*

- $\bot \notin Cn(B?\alpha)$ *(consistency)*
- $B?\alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If $\beta \in B \setminus B?\alpha$, then there is some $B' \subseteq B \cup \{\alpha\}$ such that $\bot \notin Cn(B')$ and $\bot \in Cn(B' \cup \{\beta\})$ (core-retainment)*
- $(B + \alpha)?\alpha = B?\alpha$ *(pre-expansion)*
- *If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)*

Recall the example of Paul being not at home and the lights being on. We can model it by kernel semi-revision as follows:

> **Example 10 (continued):** My initial belief base is $B = \{q, q \rightarrow p\}$. I come to believe that Paul is not at home: $B + \neg p = \{q, q \rightarrow p, \neg p\}$. I have to give up one of the beliefs in $B \perp\!\!\!\perp \bot = \{\{q, q \rightarrow p, \neg p\}\}$. I choose to give up $\sigma(B \perp\!\!\!\perp \bot) = \{q \rightarrow p\}$. My resulting belief base is $B?\neg p = B \setminus \sigma(B \perp\!\!\!\perp \bot) = \{q, \neg p\}$.

Semi-revision can also be obtained from partial meet consolidation:

3.4.19. Definition. *[Han97a] The* **partial meet semi-revision** *of B based on a selection function γ is the operator $?_\gamma$ such that for all sentences α:*

$B?_\gamma\alpha = \bigcap \gamma((B \cup \{\alpha\})\bot\bot)$

3.4.20. Theorem. *[Han97a] An operator $?$ is an operator of partial meet semi-revision if and only if for all sets B of sentences:*

- $\bot \notin Cn(B?\alpha)$ *(consistency)*
- $B?\alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If $\beta \in B \setminus B?\alpha$, then there is some B' such that $B?\alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\bot \notin Cn(B')$ and $\bot \in Cn(B' \cup \{\beta\})$ (relevance)*
- $(B + \alpha)?\alpha = B?\alpha$ *(pre-expansion)*
- *If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)*

The following example illustrates the fact that kernel and partial meet semi-revision do not always coincide:

> **Example 11:** Consider the following belief base: $B = \{p, q, p \vee q \to r\}$. To semi-revise it by $\neg r$, we have to expand B with $\neg r$ and then consolidate the resulting belief base. Using partial meet consolidation, since $(B + \neg r)\bot\bot = \{\{p, q, p \vee q \to r\}, \{p, q, \neg r\}, \{p \vee q \to r, \neg r\}\}$, we get one of the following results (depending on the selection function used): $\{p, q, p \vee q \to r\}, \{p, q, \neg r\}, \{p \vee q \to r, \neg r\}, \{p, q\}, \{p \vee q \to r\}, \{\neg r\}$, or the empty set. On the other hand, using kernel consolidation, since $(B + \neg r) \perp\!\!\!\perp \bot = \{\{p, p \vee q \to r, \neg r\}, \{q, p \vee q \to r, \neg r\}\}$, we may have an incision function that chooses to give up p and $p \vee q \to r$, so that we have $\{q, \neg r\}$ as the result of the semi-revision.

In the next chapter we will present our framework for belief change. Our approach is also non-prioritized, but it departs more from the standard AGM theory than the discussed alternative approaches do, since it drops the assumption of ideally rational agents.

Chapter 4

Resource-bounded Belief Change

In this chapter we propose a formal framework to reason about belief revision for non-ideal agents. The framework takes into account limitations of memory and deductive abilities.

For the sake of clarity, we start with a simplified model, which accounts for introspective agents, i.e., agents that can immediately recognize all their beliefs. We will introduce a structure to represent an agent's belief state that distinguishes different types of beliefs according to whether or not they are explicitly represented, whether they are currently active and whether they are fully accepted or provisional. We also define a set of basic operations that change the status of beliefs and show how these operations can be used to model agents with different capacities. The basic operations can be combined to form more complex operations. This is illustrated by showing how to define AGM-style operations using belief states equipped with the basic operations. We discuss how Harman's principles can be interpreted in our framework.

We then present the full model, which introduces the distinction between beliefs the agent is aware of and those he is not. We augment the set of basic operations to deal with this new distinction.

Proofs for all Lemmas and Propositions can be found in Appendix B.

4.1 Belief States

In this section we present our model for belief states. We start by introducing some distinctions between different kinds of beliefs. The example below motivates the distinctions. Consider the following situation:

> **Example 12:** Mary is going out, and her mother tells her that she should take an umbrella. Besides beliefs about other subjects, she holds the belief that if she is going to be outside for a long time,

> then she should take the umbrella. She also believes that she will be outside the whole day. If her mother had not mentioned the umbrella, she would not have thought of it. Upon it being brought to her notice, she concludes she should indeed take the umbrella.

Harman [Har86] proposes that some of the agent's beliefs are explicitly represented. His definition of implicit beliefs is rather vague. Beliefs that can be easily inferred from explicit ones according to classical logic belong to this category, but not all implicit beliefs need to be so derivable. For Harman, when someone explicitly believes in P, he may implicitly believe that he is justified in believing P. An agent may hold also implicit beliefs like the one that elephants do not wear pajamas in the wild mentioned in Section 2.1.

We will not restrict ourselves to a particular notion of inference but rather consider an inference function *Inf* that will depend on the agent being modeled.[1] Harman's argument that there are more implicit beliefs than those which logically follow from the explicit ones is based on his assumption that the logic in question is classical. Intuitively, the notion of inference we would like to use is neither strictly weaker nor strictly stronger than classical logic, but incomparable. Our inference function *Inf* is general enough to allow for inferences that would not be derivable using classical logic. A possible assumption about *Inf* is that inclusion holds, i.e., that for any X, $X \subseteq \mathit{Inf}(X)$. We want *Inf* to give us the inferences the agent can make in one step, i.e., by one application of some rule that the agent knows. Thus we do not want *Inf* to be idempotent.

Following Harman, we will assume that there are beliefs that are explicitly represented, from which others may be inferred. We will identify the set of the agent's implicit beliefs with the set of beliefs that can be inferred from the explicit beliefs, in accordance with the agent's abilities. The distinction between explicit and implicit beliefs is roughly analogous to that between Levi's doxastic performance and doxastic commitments [Lev91] presented in Section 2.4, with the difference that an agent's doxastic commitment is closed under classical logic while for us the set of implicit beliefs is closed under the agent's own logic. The set of explicit beliefs can be seen as a belief base in the second sense discussed in section 3.4.1, i.e., it contains those beliefs with independent standing.

Since we are dealing with resource-bounded agents, it does not make sense to state that an agent can infer everything that logically follows from his beliefs. How much one can infer depends highly on the available resources, like time and memory.

Let $\mathbf{E}$ be the set of the agent's explicit beliefs, and $\mathbf{I}$ the set of his implicit beliefs. The set $\mathbf{I}$ is given by $\mathbf{I} = \mathit{Inf}^*(\mathbf{E}) = \bigcup_{n \geq 0} \mathit{Inf}^n(\mathbf{E})$, where *Inf* is a function that returns the set of formulas that the agent is able to infer from a given set of formulas in one step, and $\mathit{Inf}^0(X) = X$. The set $\mathbf{I}$ represents the set of beliefs

[1] The agent may be allowed to revise his inference function, but we will not deal with this complication in this work.

the agent would be able to infer from **E** if he were given unlimited time. Note that $Inf^*(Inf(X)) = Inf^*(X)$ by definition. The operation Inf^* satisfies inclusion even if Inf does not.

As we have seen in Section 2.2, Cherniak defines a hierarchy of rationality concepts [Che86], on top of which appear ideal agents, with belief states that are deductively closed. On the lowest level of the hierarchy appear agents that are not able to perform any inference. These agents cannot be called rational. In general, resource-bounded agents lie somewhere in the middle of the hierarchy. Cherniak claims that a resource-bounded agent would not be called rational if he tried to make all possible inferences from his beliefs, since this would exhaust his resources without being useful (this is analogous to Harman's Principle of Clutter Avoidance). Cherniak also notes that inference does not necessarily mean the same thing for all agents: not all agents accept the laws of logic and different agents have different limitations. He speaks of *feasible inferences*. Our framework is general enough to represent agents at the bottom of the hierarchy, by taking the agent's inference function to be the identity function, i.e., $Inf(X) = X$. Perfect reasoners can be captured by taking Inf to be Cn.

Another claim that appears in [Che86] is that only a small part of an agent's beliefs can be activated or thought of at a given time. This relies on the distinction between long-term and short-term memory. We will call the information that is currently available for use *active beliefs*. These may be information that still has to be checked, such as recently acquired beliefs, intermediate conclusions in an argument, beliefs related to the current topic, etc. Some elements of the set of active beliefs might not yet be really believed (at least not completely), they still might have to be checked. Every piece of information has first to become active in order to become accepted, rejected or revised. Not all of one's beliefs are active at the same time, as the size of the set of beliefs that can be active is often restricted.

Our belief states consist of two (usually overlapping) sets, the set of explicit beliefs (**E**) and the set of active beliefs (**A**), plus an inference function that determines the set of implicit beliefs ($\mathbf{I}=Inf^*(\mathbf{E})$). In Figure 4.1 we see a representation of an agent's belief state. All changes in belief states take place in the set of active beliefs, possibly affecting the set of explicit beliefs as well.[2]

How do beliefs become active? There are many possible ways in which a belief can be activated. A belief may be active because it was recently acquired, because it is relevant to the current line of reasoning or because it has just been inferred. Usually beliefs are activated due to new input. Two different methods for activating the beliefs which are relevant for a certain input are described in Chapters 5 and 6.

[2]How the operations of belief change affect the set of explicit beliefs can be seen in Definition 4.2.1.

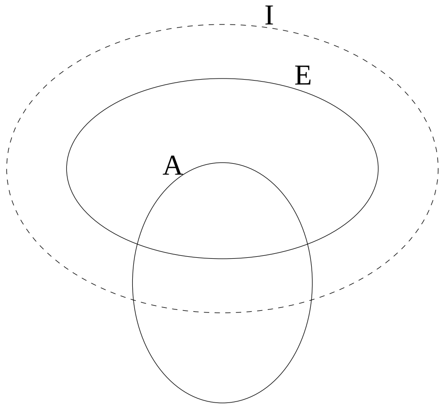

Figure 4.1: Structure of an agent's beliefs

At this point it may be useful to return to our small example to illustrate the difference between explicit and active beliefs.

Mary's belief that if she is going to be outside for a long time, then she should take the umbrella is part of her explicit beliefs and so is her belief that she will be outside the whole day. These beliefs only become active when her mother mentions the umbrella. When Mary thinks of it, she infers that she should take the umbrella. This example shows an argument against representing belief states as logically closed sets. Mary did not hold the belief that she should take the umbrella until the time at which the inference was made. It also shows that not all beliefs are active at the same time. Using Levi's term in a loose sense, Mary's performance catches up with her commitments only after the inference is made. Before performing the inference, Mary's implicit beliefs already contained the belief that she should take the umbrella. This belief was also active, since Mary had heard it from her mother, but it was not yet explicit, since Mary was not yet sure whether she would accept it.

Any new belief, either coming from the "outside" (new input) or from the "inside" (inference), has to survive inquiry before being incorporated into the current beliefs. Since we allow for both inconsistent beliefs and agents which are not ideal reasoners, an inference may well be unsound. That is why inferences should be at first only provisionally accepted. The depth of the inquiry is determined by the agent and his interest in the subject. Harman defines some kinds of cognitive goals that usually guide inquiry [Har86]: the interest in not being inconsistent, interest in the immediate environment, interest in facilitating reasoning (if the agent believes that knowing whether α is true or not would help him to obtain something he desires, he will be interested in α). For instance, consider the following example: if an agent hears that it is raining outside and he intends to go out, before going out with a raincoat and umbrella, he will probably first have a look through the window in order to be sure. But if he has no intention of going out, he might simply accept the information that it is raining and go on reading his newspaper. The agent behaves more skeptically with respect to propositions that have a direct implication to his intentions and plans or about information that comes from unreliable sources.

Harman distinguishes fully accepted beliefs from what he calls working hypotheses, the former being those working hypotheses that managed to survive inquiry. We will call working hypotheses *provisional beliefs*. In a sense, they are not real beliefs, as they are still under investigation, the agent has not yet decided whether to accept them or not. In our framework, the provisional beliefs are roughly the active beliefs that have not yet been accepted, i.e., the beliefs in $\mathbf{A} \backslash \mathbf{E}$. After we introduce the formal framework we will be able to formalize exactly the set of provisional beliefs. This agrees with Levi's claim [Lev91] that an agent should distinguish in his doxastic state between beliefs which are settled and conjectures.

An interesting question is how a provisional belief can be granted membership in the set of accepted beliefs. This is an important part of our framework which has not yet been developed. One suggestion to answer this question is presented in Appendix A.

Back to our example:

> **Example 12 (continued):** Let p stand for "Mary should take an umbrella" and q for "Mary will be outside for a long time". Before talking to her mother, Mary's explicit beliefs contain, among others, the beliefs q and $q \to p$. The implicit beliefs contain, among others, p. The set of active beliefs is empty (actually it would probably contain some remains of other reasoning, but this is not relevant for this argument). When the mother says that Mary should take an umbrella, p becomes a provisional active, but not explicit, belief. Mary does not necessarily believe everything her mother says immediately, so that she has to think about it. This is as if she were asking herself whether she should take the umbrella. The beliefs q and $q \to p$ become active, since they are relevant for deciding whether to accept p. When Mary eventually decides to accept p, this belief is made explicit and the set of active beliefs may get new elements according to new input.

At this point it may be useful to introduce another small example to illustrate the difference between explicit and active beliefs.

> **Example 13:** Consider a PROLOG program. What we call explicit beliefs is the program itself, that is, the facts and rules that are explicitly given. The inference function is the immediate consequence

operator, that gives what can be derived after one application of resolution, and the set of implicit beliefs is the fix-point of this operator. The active beliefs depend on the queries made.

For the program:

```
p :- q, r.
q.
r :- s.
s.
```

we have:

- Explicit beliefs: {`p :- q, r; q; r :- s; s`}

- Implicit beliefs: {`p :- q, r; q; r :- s; s`}∪{`r; p`}

For the query r, the set of active beliefs is first only the query, then $\{r; r:-s\}$, then $\{r; r:-s; s\}$, that is, the set contains open queries as well as the clauses of the program that are used to solve them. Note that the three sets of explicit, implicit and active beliefs are different from each other.

Intuitively it seems that implicit beliefs can also be active. In one way or other, all of an agent's active beliefs have to be explicitly represented, even if only temporarily. But they do not need to be represented in the same way as the explicit beliefs are. Recall the example of a PROLOG program (Example 13). The explicit beliefs are the facts and rules that constitute the program. The implicit beliefs are those facts that *can* be inferred from the program, while the active beliefs are the queries ($\mathbf{A}\backslash\mathbf{I}$) together with those facts and rules which have been either derived ($\mathbf{A}\cap\mathbf{I}$) or used ($\mathbf{A}\cap\mathbf{E}$) at a certain point.

A belief state Σ can be represented by $\langle \mathbf{E}, \mathit{Inf}, \mathbf{A}\rangle$, where $\mathbf{E}$ is the set of the agent's explicit beliefs, *Inf* is the agents inference function and $\mathbf{A}$ is the set of the agent's active beliefs.

Our view of the structure of an agent's beliefs (Figure 4.1) differs from the ones found in the literature, like [FH88] that start from the set of implicit beliefs and eliminate formulas to get to the set of the explicit beliefs, and [Har86] that considers that implicit beliefs can be derived from the explicit ones in a different way than by inference.

We drop the common requirement that the agent's explicit (and implicit) beliefs have to be consistent ([Gär88]), since we believe that agents can have inconsistent beliefs without believing everything (we can keep some beliefs that we even know to be inconsistent).

Even though for most applications the inference function of an agent does not vary, our framework allows for agents that learn new inference rules. If an agent's knowledge is described by first order formulas and the inference function is application of Modus Ponens, the agent may learn that the rule of generalization is valid, which will immediately change his set of implicit beliefs.

As a consequence of the introduction of these distinctions between different kinds of belief, we can represent more kinds of epistemic attitudes than the traditional AGM theory. In the AGM model, an agent may have one of three different epistemic attitudes concerning a sentence α (K represents the agent's belief state):

(i) α is accepted ($\alpha \in K$)
(ii) α is rejected ($\neg\alpha \in K$)
(iii) α is undetermined ($\alpha \notin K$ and $\neg\alpha \notin K$)

Our model allows for a more refined description of an agent's epistemic attitudes ($\langle \mathbf{E}, \mathit{Inf}, \mathbf{A} \rangle$ is the agent's belief state and $\mathbf{I}$ stands for $\mathit{Inf}^*(\mathbf{E})$):

(i) α is accepted ($\alpha \in \mathbf{E}$);
(ii) α is rejected ($\neg\alpha \in \mathbf{E}$);
(iii) α is neither accepted nor rejected but follows from the agent's beliefs ($\alpha \in \mathbf{I} \setminus \mathbf{E}$);
(iv) α is neither accepted nor rejected but can be refuted by the agent ($\neg\alpha \in \mathbf{I} \setminus \mathbf{E}$);
(v) α is under consideration ($\alpha \in \mathbf{A} \setminus \mathbf{I}$ or $\neg\alpha \in \mathbf{A} \setminus \mathbf{I}$); or
(vi) none of the above, i.e., the agent is completely ignorant about α.

Orilia has independently developed a framework for belief revision which also distinguishes active and provisional beliefs (these last are called candidate beliefs in his paper) [Ori99]. His interest is mainly avoiding the logical paradoxes such as the Liar Paradox.

4.2 Basic Operations

In this section we define operations for changing belief states as defined in section 4.1.

Traditionally, revision is seen as a sequence of a contraction and an expansion. Hansson [Han92b] has shown that revision can also be defined by expansion followed by contraction. But this is not a decomposition into simpler steps, since contraction is (computationally) as complicated as revision. We want to decompose revision and contraction into simple operations that show what happens with an agent's belief state in each step, instead of only analyzing the initial and final states.

Beliefs that are active can be forgotten, or stored as explicit (but inactive) beliefs. Since the set of active beliefs is assumed to be very limited in size, there must be a mechanism that, in cases of overflow, selects which beliefs will be discarded or stored. It may be interesting to have the active beliefs ordered by interest, so that things with very low interest will be forgotten first. This ordering could also incorporate recency, beliefs that were recently recalled are more interesting than those that were not used lately.

The first operation we define is similar to AGM expansion in the sense that it consists of simply adding new information to a set without checking for consistency. But the operation takes the limited size of the set into account.[3] When trying to add something to a set that is already at its maximum size, some elements of the set have to be given up. This can be seen as a kind of "forgetting".

If X is a set with maximum capacity m and α is an element we want to add to X, then:

$$X \cup^* \{\alpha\} = \begin{cases} X \cup \{\alpha\} & \textit{if } |X| \leq m \\ X' \cup \{\alpha\}, \textit{ where } X' \subset X \textit{ and } |X'| < m & \textit{otherwise} \end{cases}$$

Note that this operation reduces to a simple union as long as the set is not "full". Since the size m of the set is given as a parameter, the operation is more accurately denoted as $\cup^*_m$. When the set is already at its maximum size, something has to be discarded. There must be a selection mechanism to choose which elements are going to be discarded. If the set X is ordered (for example by the last time the beliefs were recalled), we can stipulate that the minimal elements of the set are the first to be dismissed, i.e., we want to ensure that if an element is dismissed, then there is no other element which is retained that is less than the dismissed one in the order:

$$\forall y(y \in X \setminus X' \to \neg\exists x(x \in X' \wedge x < y)).$$

Even under the assumption that a total order is given, the operation $\cup^*$ as defined above is not completely determined. Usually we would like it to preserve as much as possible from the old beliefs, but there are situations where it might be desirable to delete more than what is strictly necessary. One example is an artificial agent that may "clean" its working memory (the set of active beliefs) upon receiving information about an unrelated topic.

We define now six operations that can be applied to belief states in order to change the status of beliefs.

In our structure, even though the set $\mathbf{E}$ has a limited size, we assume that this size is big enough to be disregarded. The important restriction problem is the size of the set $\mathbf{A}$. Elements dismissed from $\mathbf{A}$ that do not belong to $\mathbf{E}$ or $\mathbf{I}$ simply disappear, are forgotten. The elements of $\mathbf{A}\cap\mathbf{I}$ that are dismissed from $\mathbf{A}$ remain in $\mathbf{E}$ or $\mathbf{I}$.

4.2.1. DEFINITION. **(basic operations)** *Let $\langle \mathbf{E}, Inf, \mathbf{A}\rangle$ be a belief state and α a formula. We define the following operations on $\langle \mathbf{E}, Inf, \mathbf{A}\rangle$ (we will omit the second argument Inf since the operations defined do not affect it):*

[3]When we talk about the size of a set of formulas, we mean something like its complexity. The sets $\{p, q\}$ and $\{p \wedge q\}$ should have the same size. We could, for example, count the occurrence of atoms.

1. *Observation* ($\circ_o$)*: adds an external input to the set of active beliefs.*

   $\langle \mathbf{E}, \mathbf{A} \rangle \circ_o \alpha = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle$

2. *Retrieval* ($\circ_r$)*: retrieves an explicit belief into the set of active beliefs.*

   $$\langle \mathbf{E}, \mathbf{A} \rangle \circ_r \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle, & \textit{if } \alpha \in \mathbf{E} \\ \langle \mathbf{E}, \mathbf{A} \rangle & \textit{otherwise} \end{cases}$$

3. *Acceptance* ($\circ_a$)*: makes an active belief explicit.*[4]

   $$\langle \mathbf{E}, \mathbf{A} \rangle \circ_a \alpha = \begin{cases} \langle \mathbf{E} \cup^* \{\alpha\}, \mathbf{A} \setminus \{\alpha\} \rangle, & \textit{if } \alpha \in \mathbf{A} \\ \langle \mathbf{E}, \mathbf{A} \rangle & \textit{otherwise} \end{cases}$$

4. *Inference* ($\circ_i$)*: infers something from active beliefs.*

   $$\langle \mathbf{E}, \mathbf{A} \rangle \circ_i \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle, & \textit{if } \alpha \in \mathit{Inf}(\mathbf{A}) \\ \langle \mathbf{E}, \mathbf{A} \rangle & \textit{otherwise} \end{cases}$$

5. *Doubting* ($\circ_d$)*: a belief that was accepted is questioned, becoming provisional.*

   $$\langle \mathbf{E}, \mathbf{A} \rangle \circ_d \alpha = \begin{cases} \langle \mathbf{E} \setminus \{\alpha\}, \mathbf{A} \rangle, & \textit{if } \alpha \in \mathbf{A} \cap \mathbf{E} \\ \langle \mathbf{E}, \mathbf{A} \rangle & \textit{otherwise} \end{cases}$$

6. *Rejection* ($\circ_c$)*: rejects an active belief.*

   $$\langle \mathbf{E}, \mathbf{A} \rangle \circ_c \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A} \setminus \{\alpha\} \rangle, & \textit{if } \alpha \in \mathbf{A} \\ \langle \mathbf{E}, \mathbf{A} \rangle & \textit{otherwise} \end{cases}$$

The operations defined above describe how beliefs are incorporated into the structure representing an agent's beliefs and how they move from one set to the other within the structure. Elimination of beliefs may happen as a side effect of the expansion operation $\cup^*$, being thus an "unconscious" operation. However, depending on the kind of agent being modeled, it may be useful to have an operation to perform "conscious" forgetting, for example for robots. Such an operation can easily be modeled by combining the operations of doubting and rejection.

The six operations defined above can be combined to model more complex operations. As an example of such a composition, consider what happens when an agent gets new information via observation. The belief will first come into the set of active beliefs through the operation $\circ_o$ and then the agent may accept it ($\circ_a$). Another example is the operation of conscious forgetting mentioned above.

In what follows, we will use the six operations defined also as operations that take a belief state and a finite enumerated set of formulas and return a belief state, i.e., if Σ is a belief state and $X = \{\chi_1, \chi_2, ..., \chi_n\}$ is a set of formulas, then

[4] Acceptance could also be defined without deleting the accepted belief from $\mathbf{A}$, which seems to be more intuitive for human agents. The choice made here reflects our interest in artificial agents.

$\Sigma \circ_? X = \Sigma \circ_? \chi_1 \circ_? \chi_2 \circ_? \ldots \circ_? \chi_n$, where $\circ_?$ is one of the six basic operations $\circ_o$,$\circ_r$, $\circ_a$, $\circ_i$, $\circ_d$ or $\circ_c$. Since the basic operations make use of $\cup^*$, beliefs may be deleted due to overflow, and different enumerations of the same set may yield different outcomes. In the rest of this chapter, when no particular enumeration is mentioned, the result is valid for all enumerations of the set considered.

It is not difficult to see that the set of operations $\circ_o$, $\circ_r$, $\circ_a$, $\circ_d$, and $\circ_c$ is complete with respect to all possible changes that a belief state may undergo, i.e.:

4.2.2. PROPOSITION. *Given two belief states $\Sigma_1 = \langle E_1, A_1\rangle$ and $\Sigma_2 = \langle E_2, A_2\rangle$, there is a sequence of basic operations that takes Σ_1 into Σ_2.*

It is interesting to note that in this simplified model, observing or inferring a formula α have the same effect, i.e., $\langle \mathbf{E},\mathit{Inf},\mathbf{A}\rangle \circ_o \alpha = \langle \mathbf{E},\mathit{Inf},\mathbf{A}\cup^*\{\alpha\}\rangle = \langle \mathbf{E},\mathit{Inf},\mathbf{A}\rangle \circ_i \alpha$, provided that $\alpha \in \mathit{Inf}(\mathbf{E})$. However, observation depends on the outside world and a formula may be available for inference without being available for observation. In Section 4.5, where we introduce the full model, we will be able to differentiate between observing a formula that can be inferred and one that cannot.

If we want to use the model described above to model an ideal agent, we can simulate AGM operations. This is presented in the next section.

4.3 Embedding AGM Theory

In this section, we show how the operations of expansion and contraction in the AGM sense can be seen as a special case of applications of the operations defined in section 4.2, when the storage space is unlimited and the number of operations performed may be infinite.

We will concentrate on expansion and contraction, since AGM-revision can be defined in terms of the other two operations.

A belief set is a set K such that $Cn(K) = K$. There is no distinction between implicit and explicit beliefs or active and inactive beliefs.

To see how this can be embedded in our structure, we first observe that a belief set in the AGM theory corresponds to a belief state where the sets **I** and **E** are the same, the sets **E** and **A** have no size limit, and *Inf* is Cn. Thus, we define a map function f from belief sets into belief states:

$$f(K) = \langle K, Cn, \emptyset\rangle$$

Since the sets considered in this section do not have any size limit, the operation $\cup^*$ becomes equivalent to the simple set union operation and the basic operations can be easily extended to the form $\Sigma \circ_? X$, where X is a (possibly infinite) set of sentences and $\circ_?$ is one of the six basic operations $\circ_o$,$\circ_r$, $\circ_a$, $\circ_i$, $\circ_d$ or $\circ_c$.

4.3.1. DEFINITION. *Let $\Sigma = \langle X, Cn, \emptyset \rangle$ be a belief state. The logical closure of Σ is given by:*
$Cl(\Sigma) = \Sigma \circ_i Cn(X) \setminus X \circ_a Cn(X) \setminus X.$

It follows directly that if $X = Cn(X)$, then $\langle X, Cn, \emptyset \rangle = Cl(\langle X, Cn, \emptyset \rangle)$.

Since there is no size limit on the set of active beliefs, it is easy to see that:

4.3.2. LEMMA. *Let $\Sigma = \langle K, Cn, \emptyset \rangle$ be a belief state.*
Then $Cl(\Sigma) = \langle Cn(K), Cn, \emptyset \rangle$.

Given a belief set K, the AGM-expansion of K with α is defined by:

$K + \alpha = Cn(K \cup \{\alpha\})$

4.3.3. DEFINITION. *The expansion of a belief state $\Sigma = \langle K, Cn, \emptyset \rangle$ by α is given by:*
$\Sigma \dot{+} \alpha = Cl(\Sigma \circ_o \alpha \circ_a \alpha)$

We have then that by mapping the resulting belief set after an AGM expansion into a belief state, we get the same result as predicted by our theory (see Figure 4.2):

$$\begin{tikzcd} K \arrow[r, "+"] \arrow[d, "f"'] & K+\alpha \arrow[d, "f"] \\ f(K) \arrow[r, "\dot{+}"] & f(K+\alpha) \end{tikzcd}$$

Figure 4.2: Commuting diagram

4.3.4. LEMMA. $f(K) \dot{+} \alpha = f(K + \alpha)$

4.3.5. DEFINITION. *Let γ be a selection function for K. The partial meet contraction of a belief state $\Sigma = \langle K, Cn, \emptyset \rangle$ by α is given by:*
$\Sigma \dot{-}_\gamma \alpha = \Sigma \circ_r \Delta \circ_d \Delta \circ_c \Delta,$
where $\Delta = K \setminus (\bigcap \gamma(K \perp \alpha))$.

4.3.6. LEMMA. *If $-_\gamma$ is the AGM partial meet contraction operator based on γ, then $f(K)\dot{-}_\gamma\{\alpha\} = f(K -_\gamma \alpha)$.*

Since the revision of a belief set can be defined by combining contraction and expansion according to the Levi identity, we are able to conclude the following:

4.3.7. PROPOSITION. *The standard AGM theory of belief revision can be embedded in the framework of belief states equipped with the operations presented in Definition 4.2.1.*

In order to obtain a contraction operation that satisfies the AGM postulates, the operation of rejection has to be repeated until enough sentences have been removed so that α is not a consequence of the first argument of the resulting belief state. Such a definition, as the one in [Gär88], is only appropriate for ideal agents.

As we have seen, it is possible to use our framework to define operations for belief change in the tradition of the AGM theory. But the operations of expansion ($\dot{+}$) and contraction ($\dot{-}$) defined will usually involve an infinite number of basic operations. If we want to define versions of these operations for non-ideal agents, we have to restrict the use of the set of active beliefs.

Instead of removing from the belief state all the beliefs that imply the belief being contracted, a contraction operation for a non-ideal agent should remove only those beliefs that are active and imply the contracted one. This would prevent the agent from immediately reinferring a belief that was given up, satisfying Harman's Get Back Principle (cf. Section 2.1). On the other hand, this would not prevent the agent from later on (when the active beliefs change) reinferring a belief that was given up. In Chapter 5, we present such a contraction operation.

In the case of Karen's aptitude test discussed in Section 2.1, suppose she was already thinking of something else when she got to know that the results were wrong. When she hears the news some of her beliefs become active, like the belief that she has indeed aptitude for science and music. Since this belief depended on the belief that the results of the test were correct, Karen will give it up when she finds out that it is not justified anymore. The set of active beliefs behaves then according to the foundations theory. But there were other beliefs that originally depended on the result of the test, like the belief that she did not work hard enough in music. Since this belief is not active anymore it will be maintained. The set of explicit (and inactive) beliefs behaves more according to the coherence theory, but it can contain inconsistent beliefs. It seems intuitive that if Karen is asked why she thinks she did not work hard enough she will reconstruct some kind of justification, maybe different from the original one. She may find out that she dedicated much more time to history than it was necessary, leaving music aside.

The main point is that the operations of contraction and revision are restricted to the (small) set of active beliefs, which makes them feasible. There is also no consistency checking involving all of an agent's beliefs, only the active beliefs are kept consistent.

The most important problem to be solved is how to select which information should be in the set of active beliefs. Recently acquired information should be active, as well as open queries, but also relevant information should be retrieved from the set of explicit beliefs to be used in the reasoning process. In Chapter 5, one way of retrieving relevant information is presented that considers a formula to be relevant if it helps to prove or refute something under consideration. In the system RABIT [Gar93], a method of marker passing is used to determine which of the explicit beliefs should be retrieved and become active. This is the starting point for the method presented in Chapter 6.

4.4 Harman's Principles

In this section we give an interpretation for the principles in [Har86] that were presented in Section 2.1 and consider how well the current proposal can be integrated with Harman's theory. Let $\langle \mathbf{E}, \mathit{Inf}, \mathbf{A} \rangle$ be a belief state.

1. *Clutter Avoidance:* This principle has as its main implication that the agent should not try to close his beliefs under logical implication, since not all consequences of the agent's explicit beliefs are useful. Clutter Avoidance does not apply to the set of implicit beliefs, that represents what the agent could (but not necessarily wants to) infer. Usually, $\mathbf{E} \neq \mathit{Inf}(\mathbf{E})$.

2. *Recognized Implication Principle:* The agent can only recognize an implication if the premises are accepted and active. Moreover, in order to be accepted, an inference also has to be feasible, i.e., it has to be obtained by a small number of applications of Inf. The agent has reasons to accept a new inferred belief α if $\alpha \in \mathit{Inf}(\mathbf{A} \cap \mathbf{E})$.

3. *Recognized Inconsistency Principle:* The agent is only aware of inconsistencies in his set of active beliefs. If an inconsistency is found, i.e., if the set of active beliefs becomes inconsistent, then there is a reason to correct it. The set $\mathbf{E} \setminus \mathbf{A}$ may be inconsistent, but this will not affect the reasoning.

4. *Principle of Positive Undermining:* An accepted belief can move to the set of provisional beliefs and go through inquiry again if there is evidence against it. Our theory does not say anything about what should count as evidence for or against a belief. We can imagine that a consistent set of accepted beliefs implying α could be seen as evidence for α, but there is more to evidence than this. To describe this, belief states would probably have to be enriched with a structure reflecting justifications. This is left for further work.

5. *Principle of Conservatism:* When changing his beliefs, the agent should perform only the necessary changes. Beliefs that are irrelevant for the change the agent is performing should remain untouched. Changes always involve elements of the set of active beliefs (which may also be elements of the set of implicit beliefs). The only exception is in the case where the capacity of the agent's memory is already exhausted and some beliefs have to be given up (forgotten).

6. *Interest Condition:* Here our theory does not have much to say. This principle implies that an agent's reasoning should be goal-oriented, i.e., that the agent should not make arbitrary inferences but instead pursue a goal. His interest should guide which inferences are worth making.

7. *Get Back Principle:* The agent should not give up a belief that can be reinferred from his active beliefs. This means that when giving up a belief α, enough beliefs have to be given up so that $\alpha \notin \mathit{Inf}(\mathbf{A})$. But it may be the case that $\alpha \in \mathit{Inf}(\mathbf{E})$.

Harman's Recognized Implication Principle touches a problem that is not addressed in this simplified model. An agent may receive information that is implied by his beliefs and fail to notice it. After some time, the agent eventually realizes that the new piece of information was already implicitly believed. At this point we can say that the agent recognized the implication. In this simplified model we cannot distinguish between a piece of information that is active (because it has just been observed) and happens to follow from the agent's explicit belief, without the agent realizing it, and those beliefs that the agent recognizes as being implied by his explicit beliefs. Consider what happens immediately after Mary's mother saying that she should take an umbrella (Example 12). This becomes part of Mary's active (provisional) beliefs and it follows from her explicit beliefs. But only after some reasoning does Mary recognize the fact that the new piece of information was already an implicit belief. In the next section we will present the full model of belief states and show that it accounts for this distinction.

4.5 Refining the Model

As we have seen at the end of last section, a provisional belief may already be believed (explicitly or implicitly) without the agent noticing it. We would like to further refine the set $\mathbf{A}$ to mirror the difference between beliefs that the agent is "aware" of believing (in the sense that he believes that he believes them) and beliefs that he is not aware of.

As can be seen in Figure 4.3, the set of provisional beliefs (**prov**) includes beliefs that were not previously believed as well as beliefs that are already (implicitly or explicitly) believed but of which the agent is not aware.

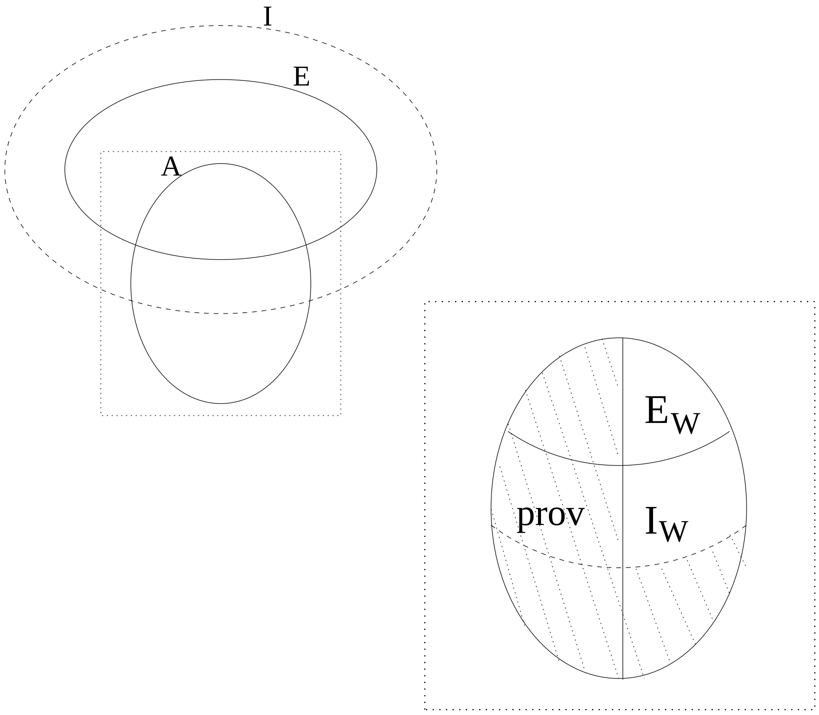

Figure 4.3: The provisional beliefs

The set **W** (see Figure 4.4) is a subset of **A** and contains those beliefs that the agent is aware of. Beliefs in **E**∩**W** are explicit and the agent is aware that they are explicitly believed. Analogously, the agent is aware that he implicitly holds the beliefs in **I**∩**W**. And the beliefs in **W****I** are those which the agent is aware that they are only provisional beliefs, i.e., the agent knows that these beliefs do not follow from his explicit beliefs. To illustrate the difference between the sets **I**∩**W** and **E**∩**W**, recall once more the PROLOG program (Example 13). After the query p, at a certain point r will be in **I**∩**W**, since it was inferred and is available for use. But it will not be in **E**∩**W**, since it is not part of the program.

When Mary hears that she should take an umbrella, this becomes a provisional belief and was already part of Mary's implicit beliefs. This new belief becomes part of $(\mathbf{I} \cap \mathbf{A}) \setminus \mathbf{W}$, i.e, it goes to region 5 in Figure 4.5. After some reasoning, Mary becomes aware that she was already committed to the belief that she should take the umbrella, and the belief moves to region 6 (**I**∩**W**).

A belief state Σ can be completely determined by the following parameters:

- The set **E** of explicit fully accepted beliefs;
- The inference function *Inf*, that takes as argument a set of beliefs and

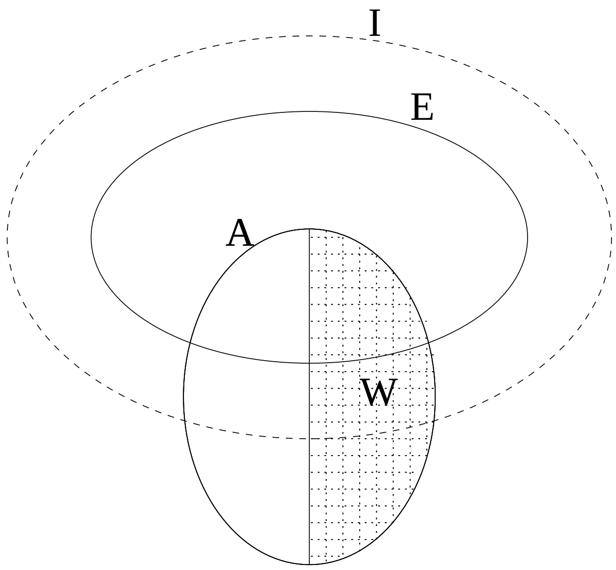

Figure 4.4: The active beliefs

gives as result the set of beliefs that can be inferred by the agent from the argument set in one step;

- The set **A** of active beliefs; and
- The set **W** of active beliefs that the agent is aware of (**W** must be a subset of **A**).

From now on, we will refer to a belief state as $\Sigma = \langle \mathbf{E}, \mathit{Inf}, \mathbf{A}, \mathbf{W} \rangle$.

The set of implicit beliefs is given by $\mathbf{I}$=$\mathit{Inf}^*(\mathbf{E})$, the set of explicit active beliefs of which the agent is aware is given by $\mathbf{E}_W$=$\mathbf{E}\cap\mathbf{W}$, the set of implicit active beliefs of which the agent is aware is given by $\mathbf{I}_W$=$\mathbf{I}\cap\mathbf{W}$, and the set of provisional beliefs is given by $\mathbf{A} \setminus \mathbf{I}_W$.

In Figure 4.5, we see the different regions of an agent's belief state. The set **E** of explicit beliefs is given by the union of regions 2, 3, and 4; the set **I** of implicit beliefs is the union of regions 1, 2, 3, 4, 5, and 6; the set **A** of active beliefs is given by regions 3, 4, 5, 6, 7, and 8; the set of beliefs the agent is aware of (**W**) is the union of regions 4, 6, and 8; the set $\mathbf{E}_W$ of explicit beliefs of which the agent is aware is given by region 4 and the provisional beliefs are given by the regions 3, 5, 7 and 8.

We can now define a new set of basic operations for belief states (the numbers refer to the regions in Figure 4.5):

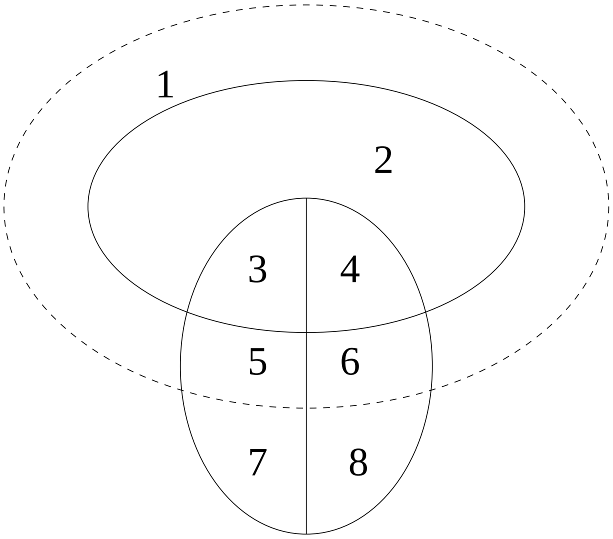

Figure 4.5: Numbered regions

4.5.1. DEFINITION. *The following operations apply to a belief state* $\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle$ *(as in definition 4.2.1, we omit the argument Inf, since it is not affected by the operations):*

1. *Observation* ($\circ_o$) *– The agent receives new information from "outside", either via observation or communication. The new information goes first to the set* **prov** *of provisional beliefs. It may already be an explicit belief, going then to region 3, an implicit belief, going to region 5 or something that was not previously believed, in which case it goes to region 7.*

   $\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_o \alpha = \langle \mathbf{E},\ \mathbf{A}',\ \mathbf{W}'\rangle,$

   where $\mathbf{A}' = \mathbf{A} \cup^* \{\alpha\}$ *and*
   $\mathbf{W}' = \mathbf{W} \cap \mathbf{A}'$.

2. *Retrieval* ($\circ_r$) *– An explicit belief that is inactive becomes active. It moves from region 2 to 4.*

   $$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_r \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A}', \mathbf{W}'\rangle, & \textit{if } \alpha \in \mathbf{E} \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

   where $\mathbf{A}' = \mathbf{A} \cup^* \{\alpha\}$ *and*
   $\mathbf{W}' = (\mathbf{W} \cup^* \{\alpha\}) \cap \mathbf{A}'$.

3. *Acceptance* ($\circ_a$) *– An active belief that was in regions 3, 5, 6, 7 or 8 is fully*

accepted and moves into $\mathbf{E}$ *(region 2).*[5]

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_a \alpha = \begin{cases} \langle \mathbf{E}', \mathbf{A}', \mathbf{W}'\rangle, & \textit{if } \alpha \in \mathbf{A} \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

where $\mathbf{E}' = \mathbf{E} \cup^* \{\alpha\}$,
$\mathbf{A}' = \mathbf{A} \setminus \{\alpha\}$, *and*
$\mathbf{W}' = \mathbf{W} \setminus \{\alpha\}$.

4. *Recognition* ($\circ_g$) – *A belief that was in* $\mathbf{prov} \cap \mathbf{E}$ *(region 3) is discovered to be already explicitly believed and moves into* $\mathbf{E}_W$ *(region 4), a belief that was in* $\mathbf{prov} \cap \mathbf{I}$ *(region 5) is seen to be already implicitly believed and moves into* $\mathbf{I}_W$ *(region 6) or a belief that was in* $\mathbf{prov} \setminus (\mathbf{I} \cup \mathbf{W})$ *(region 7) is recognized as provisional and moves into* $\mathbf{W} \setminus \mathbf{I}$ *(region 8).*

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_g \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A}, \mathbf{W}'\rangle, & \textit{if } \alpha \in \mathbf{A} \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

where $\mathbf{W}' = \mathbf{W} \cup^* \{\alpha\}$.

5. *Inference* ($\circ_i$) – *The new information is inferred from the active beliefs, that is, it comes from the set* $\mathit{Inf}(\mathbf{A})$ *and goes to* $\mathbf{I}_W$ *(region 6).*

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_i \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A}', \mathbf{W}'\rangle, & \textit{if } \alpha \in \mathit{Inf}(\mathbf{A}) \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

where $\mathbf{A}' = \mathbf{A} \cup^* \{\alpha\}$ *and*
$\mathbf{W}' = (\mathbf{W} \cup^* \{\alpha\}) \cap \mathbf{A}'$.

6. *Doubting (explicit)* ($\circ_{d_e}$) – *A belief that was fully accepted is questioned, moving from region 4 to region 8.*

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_{d_e} \alpha = \begin{cases} \langle \mathbf{E}', \mathbf{A}, \mathbf{W}\rangle, & \textit{if } \alpha \in \mathbf{E}_W \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

where $\mathbf{E}' = \mathbf{E} \setminus \{\alpha\}$.

7. *Doubting (implicit)* ($\circ_{d_i}$) – *A belief that was recognized as implicit is questioned, moving from region 6 to region 5.*

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_{d_i} \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A}, \mathbf{W}'\rangle, & \textit{if } \alpha \in \mathbf{W} \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

where $\mathbf{W}' = (\mathbf{W} \setminus \{\alpha\})$.

8. *Rejection* ($\circ_c$) – *A belief that was in* $\mathbf{prov}$ *(regions 3, 5, 7 or 8) is not accepted because it did not survive inquiry and is dismissed.*

$$\langle \mathbf{E},\ \mathbf{A},\ \mathbf{W}\rangle \circ_n \alpha = \begin{cases} \langle \mathbf{E}, \mathbf{A} \setminus \{\alpha\}, \mathbf{W} \setminus \{\alpha\}\rangle, & \textit{if } \alpha \in \mathbf{A} \\ \langle \mathbf{E}, \mathbf{A}, \mathbf{W}\rangle & \textit{otherwise} \end{cases}$$

[5] As in Definition 4.2.1, acceptance could also be defined without deleting the accepted belief from $\mathbf{A}$.

The operation of implicit doubting ($\circ_{d_i}$) may at first appear to be irrational. After all, the doubted belief continues to follow from the explicit beliefs, even if a rejection operation follows the doubting. But the following example shows that the operation is appropriate for resource-bounded agents:

> **Example 14:** Suppose an agent recognizes that α follows from his explicit beliefs, i.e., $\alpha \in \mathbf{I}_A$. Much later, the α is still there, but the agent cannot recall its derivation, since he does not have enough time to search all of his explicit beliefs to see which ones imply α. So he doubts α and rejects it. From his point of view, he does not believe α anymore, but α is still in $\mathbf{I}$.

4.5.2. OBSERVATION. *Let $\Sigma = \langle \mathbf{E},\ \mathbf{A},\ \mathbf{W} \rangle$ be a belief state and $\mathbf{W} \subseteq \mathbf{A}$. Let $\langle \mathbf{E}'$, $\mathbf{A}'$, $\mathbf{W}' \rangle$ be the result of applying one of the basic operations to Σ. Then $\mathbf{W}' \subseteq \mathbf{A}'$.*

Proof: Trivial given the conditions for the application of the operations. □

We will call *active belief change* those revisions that the agent performs consciously, in the sense that the changes are all related to the set of active beliefs. A belief cannot move from region 1 to region 2 in Figure 4.5 without ever being active.

The eight operations are sufficient (but not necessary) to describe any active belief change given our structure. Actually, seven of these eight operations are sufficient, since an inference ($\circ_i$) can be simulated by an observation ($\circ_o$) followed by recognition ($\circ_r$). But note that in opposition to the simplified model, the result of observation and inference is not the same. An inferred sentence goes directly into $\mathbf{I}_W$, while an observed sentence must be first recognized as an implicit belief in order to enter $\mathbf{I}_W$.

The proposition below shows that the set of operations $\circ_o$, $\circ_r$, $\circ_a$, $\circ_g$, $\circ_{d_e}$, $\circ_{d_i}$, and $\circ_c$ is complete with respect to all the changes that a belief state may undergo:

4.5.3. PROPOSITION. *Given any two belief states $\Sigma_1 = \langle E_1, A_1, W_1 \rangle$ and $\Sigma_2 = \langle E_2, A_2, W_2 \rangle$, there is a sequence of basic operations that takes Σ_1 into Σ_2.*

We can show that for a certain kind of agent, the full model and the simplified one coincide:

4.5.4. PROPOSITION. *The simplified model can be embedded in the refined one under the assumption that the agent is introspective, i.e., that $\mathbf{W}{=}\mathbf{A}$.*

The introspection assumption involves a great deal of idealization for human beings but in some cases it may be reasonable for artificial agents. Consider for instance a small database with not too much inferential power and an efficient search mechanism. Such an artificial agent can recognize explicit and implicit beliefs so fast that it does not make sense to distinguish $\mathbf{A}$ from $\mathbf{W}$.

Let f' be a function mapping simplified belief states into refined belief states such that:

$f'(\langle \mathbf{E}, \mathbf{A}\rangle) = \langle \mathbf{E}, \mathbf{A}, \mathbf{A}\rangle$.

Using the mapping f' together with f defined in Section 4.3 gives us an embedding of the AGM framework into the refined model.

Let h be the composition of f and f', that is, $h(X) = f'(f(X)) = \langle X, Cn, \emptyset, \emptyset\rangle$ for every set of formulas X.

Let $\Sigma = \langle K, Cn, \emptyset, \emptyset\rangle$ and let the closure of Σ be given by $Cl(\langle X, Cn, \emptyset, \emptyset\rangle) = \Sigma \circ_i Cn(X) \setminus X \circ_a Cn(X) \setminus X$. It is not difficult to see that $Cl(\langle X, Cn, \emptyset, \emptyset\rangle) = \langle Cn(X), Cn, \emptyset, \emptyset\rangle)$. Let $\Sigma \ddot{+} \alpha = Cl(\Sigma \circ_o \alpha \circ_g \alpha \circ_a \alpha)$.

4.5.5. Lemma. *$h(K)\ddot{+}\alpha = h(K + \alpha)$.*

We can define the partial meet contraction of a belief state $\Sigma = \langle K, Cn, \emptyset, \emptyset\rangle$ as $\Sigma \ddot{-}_\gamma \alpha = \Sigma \circ_r \Delta \circ_{d_e} \Delta \circ_c \Delta$, where $\Delta = K \setminus (\bigcap \gamma(K \perp \alpha))$ and γ is a selection function.

4.5.6. Lemma. *$h(K)\ddot{-}_\gamma \alpha = h(K -_\gamma \alpha)$.*

In Chapter 5 we will see how belief base revision operators can be embedded into our framework.

We can look once more to Mary's example to see that it can be formalized in the full model.

> **Example 12 (continued):** Let p stand for "Mary should take an umbrella" and q for "Mary will be outside for a long time". Before talking to her mother, Mary's explicit beliefs contain, among others, the beliefs q and $q \to p$. The implicit beliefs contain, among others, p. When the mother says that Mary should take an umbrella, p goes to region 5 in Figure 4.5 via an operation of observation. Next, q and $q \to p$ are retrieved into the set of active beliefs and Mary recognizes that p was already an implicit belief. The formula p moves then to region 6.

The need to distinguish $\mathbf{I}_W$ and $\mathbf{E}_W$ from $\mathbf{I}\cap\mathbf{A}$ and $\mathbf{E}\cap\mathbf{A}$ arises in modeling both natural and artificial agents. However, the distinction between $\mathbf{I}_W$ and $\mathbf{E}_W$ seems to be less motivated for natural agents. It seems unnatural that someone sees that a formula follows from his beliefs, accepts it, but decides not to store the information. One example may be a logician trying to prove a theorem. During the process, the logician finds out many formulas that follow from the initial assumptions and may believe they are valid, but he may also see that they are not relevant for the proof and not write them down to remember.

In the next two chapters we turn to the problem of how to decide how beliefs become active.

Chapter 5

Local Change

In this chapter we are going to deal with one of the problems left open in the framework presented in Chapter 4, namely how to decide which beliefs should be retrieved into the set of active beliefs. We define the relevant beliefs for an operation of belief change and define local inference operators that only consider the relevant part of a belief base. This operator is used to define local versions of the operations for belief change. Representation theorems are given for the local operators. At the end of this chapter we will show how the local operators can be embedded in the framework presented in Chapter 4.

We will use the terms "belief bases" and "set of explicit beliefs" indiscriminately. The local inference operators do not correspond exactly to the agent's inference function *Inf* introduced in Chapter 4. Application of one of the local inference operators on the belief base does not have as a result the whole set of the agent's implicit beliefs, but only the relevant part of it. The use of the local operators will become clear after the formalization has been illustrated in Sections 5.1 and 5.2. In Section 5.1.2, we will make the connection between the local inference operators and the agent's inference function *Inf*.

We will consider the five base operations described in section 3.4.1: expansion, contraction, revision, consolidation, and semi-revision. One of these five operations, namely expansion, cannot easily be restricted to a part of the belief base. To expand a belief base B by a sentence α means to perform the simple operation $B + \alpha = B \cup \{\alpha\}$. There is no way to restrict this addition of α to only a part of B, unless we introduce an extra-logical division of B into compartments. As will be seen below, so much can be achieved with the division of B that can be derived from logic alone that we do not wish at this stage to add this further complication.

The other four operations can, however, readily be localized as follows:

1. *Local Contraction*: A belief is removed from a certain part of the belief base. I may, for instance, give up my belief that dinosaurs died out due to

a cosmic collision, without thinking about the consequences of this change for my beliefs about non-paleontological matters.

2. *Local Consolidation*: Inconsistencies are removed from some part of the belief base. The rest of the agent's beliefs may well be inconsistent. For instance, I can make my beliefs about biological evolution consistent, while retaining global inconsistency between biological and religious beliefs.

3. *Local Revision*: A new belief is added to the belief base in such a way that a certain part of the resulting base is made (kept) consistent. If I see, for example, that it is a sunny day in Amsterdam, then this contradicts my belief that it is always raining in Holland, and leads to revision. This can be done without checking whether my beliefs about Brazilian politics are consistent with the new belief.

4. *Local Semi-revision*: A new belief is either accepted or rejected, in such a way that a certain part of the resulting belief base is made (kept) consistent. If I hear that it is snowing in Rio de Janeiro, then this contradicts my beliefs about the climate there. I may either accept or reject this new information. In both cases, my beliefs about Latin grammar (and most other subjects) will be unaffected by the operation.

Informal examples like the following elaboration of Example 10 motivate our search for plausible constructions of local operators:

> **Example 15:** When at home I hear on the radio that my friend Carol has been murdered yesterday night and that there were no traces of doors or windows having been forced. I talked to her yesterday on the phone and she was at home with her flat-mates Ann and Bill. I know that no one else, except for Ann, Bill and Carol had the keys to their apartment. I conclude that Ann or Bill must have done it. But I have known Ann for quite some time and cannot believe that she would be able to murder anyone. I believe that she did not do it. For similar reasons, I believe that Bill did not do it. This is clearly inconsistent with my belief that one of them did it. So I decide to visit my friend Paul to ask what he thinks. In front of his place I see the lights are on. I know that if the lights are on, then Paul is home. I get out of the car and Paul's neighbor, who understands that I come to visit Paul tells me that he is not home. This is all very confusing, but I am sure of one thing: I do not believe I am asleep!
>
> This illustrates the fact that inconsistencies are local, that is, the fact that I have inconsistent beliefs does not cause me to believe in everything.

> I have expanded my belief base with the information given by the neighbor and reached a local inconsistency. I am interested now in whether Paul is at home or not. For a moment, I forget about the murder and think of the reasons that I have to believe that Paul is at home and that he is not. In order to eliminate the local inconsistency, I have to give up at least one of the beliefs. Suppose I ring the bell and Paul answers the door. Then I reject the neighbor's information that Paul is not home. On the other hand, suppose that he does not answer the door. Since I see the lights on, I give up my previous belief that if the lights are on then he is at home. Perhaps he has forgotten the lights on when he left the house. In both cases, I eliminated the local inconsistency with respect to whether or not Paul is at home, but I still have inconsistent beliefs about the crime. This is an example of local consolidation.
>
> The whole operation, that is, adding the information given by the neighbor and then locally consolidating the beliefs illustrates the operation of local semi-revision.

Our first task is to identify the compartments of a belief base in terms of which local change can be defined. This will be done in Section 5.1. In Section 5.2 we define local operations of belief change. In Section 5.3 we compare our approach to some others in the literature. Some concluding remarks are offered in Section 5.5. All formal proofs are deferred to Appendix C.

5.1 Defining Compartments and Local Inference

In Section 5.1.1 we define the notion of compartments that is used in Section 5.1.2 for constructing a local inference operation. Local inference operations are restricted to relevant compartments of a belief base. The idea of using compartments to isolate inconsistencies is not new, see for example [Jaś69, Lew82].

5.1.1 Defining Compartments

Compartments of belief bases can be seen as representations of compartments of minds or databases. There are two major ways to introduce them. First, they may be introduced as an addition to the logic, so that one and the same belief base can be divided into compartments in different ways.[1] Secondly, they may be derived from the logic. The second method is the more economical, requiring no extra entities, and should be tried out first. We are going to use it here.

The concept we will use is that of the compartment "around" a sentence or set of sentences. Intuitively, the compartment in a belief base B around a sentence

[1] This approach was studied by Parikh in [Par96].

α is the subset of B that is relevant to α. Since one and the same sentence β may be relevant to both α_1 and α_2, compartments thus defined will typically be overlapping.

We will assume that the compartment of B around a set of sentences is the subset of B which is relevant to at least one of these sentences, so we can define the compartment around a set A to be the union of the compartments around each of the sentences of A. Hence, letting $c(A, B)$ denote the A-compartment of B:

(1) $c(A, B) = \bigcup_{\alpha \in A} c(\alpha, B)$.

It follows from (1) that $c(\emptyset, B) = \emptyset$ for all B. Furthermore, due to (1), we can content ourselves with defining c for single sentences. Note that (1) involves a severe idealization. It may be the case that for some formulas α, β and γ, α is not relevant for β or for γ, but is relevant for the set $\{\beta, \gamma\}$. An example due to Dubois (cited in [Her97]) lets α be "I take a bath", β be "I use a hair-dryer" and γ be "I die". Most of the time we will be interested in the compartment around a single sentence, usually the input sentence that triggered an operation of belief change, so we do not have to worry too much about this idealization.

We are only interested in compartments around contingent expressions, since intuitively, no formula should be relevant for tautologies or contradictions. For these, we define:

(2) $c(\alpha, B) = \emptyset$ if $\alpha \in Cn(\emptyset)$ or $\neg\alpha \in Cn(\emptyset)$.

Since compartments are based on the underlying logic, they will depend on the inference operator. A useful tool in the definition of logical compartments is that of a *kernel* [Han94]. We give a straightforward generalization of Definition 3.4.8 in order to base kernels on an arbitrary inference operator. Given a sentence α and an inference operator C, the set X of sentences is an *α-kernel* if and only if it is a minimal set C-implying α. The *kernel operator* $\perp\!\!\!\perp_C$, based on C, is the operator that, given a belief base B and a sentence α, selects all α-kernels that are subsets of B. Hence:

5.1.1. DEFINITION. *Let C be an inference operation on L. Then the kernel operation $\perp\!\!\!\perp_C$ is the operation such that for all subsets B of L and elements α of L, $X \in B \perp\!\!\!\perp_C \alpha$ if and only if:*

1. *$X \subseteq B$*
2. *$\alpha \in C(X)$*
3. *for all Y, if $Y \subset X$ then $\alpha \notin C(Y)$*

The elements of $B \perp\!\!\!\perp_C \alpha$ are called α-kernels.
The symbol $\perp\!\!\!\perp$ (without subscript) denotes the kernel operation associated with Cn, the classical consequence operator on L.

5.1.2. Observation. *Let C be an inference operation on the language L and $\perp\!\!\!\perp_C$ its associated kernel operation. If C satisfies compactness, then $B \perp\!\!\!\perp_C \alpha \neq \emptyset$ for all $B \subseteq L$ and all $\alpha \in C(B)$.*

Note that compactness is sufficient for this property. Monotony is not needed. If B is finite, which is typically the case for belief bases, then even compactness is not required.

5.1.3. Observation.

1. *If $B \subseteq B'$, then for every formula α, $(B \perp\!\!\!\perp_C \alpha) \subseteq (B' \perp\!\!\!\perp_C \alpha)$.*
2. *If C satisfies compactness, then $B \perp\!\!\!\perp_C \alpha = B \perp\!\!\!\perp_C \beta$ if and only if for all subsets B' of B, $\alpha \in C(B')$ iff $\beta \in C(B')$.*
3. *$X \in B \perp\!\!\!\perp_C \alpha$ if and only if $X \subseteq B$ and $X \in X \perp\!\!\!\perp_C \alpha$.*

In what follows, we will focus on compartmentalization functions that are based on the classical consequence operation (Cn). A first attempt to define the compartment for α in B is:

(3) $c_1(\alpha, B) = \bigcup(B \perp\!\!\!\perp \alpha)$, for contingent α.

This definition is unsatisfactory since inconsistent kernels will be included, e.g. $c_1(\{p\}, \{p, q, \neg q\}) = \{p, q, \neg q\}$, since $\{q, \neg q\} \in \{p, q, \neg q\} \perp\!\!\!\perp p$. This problem can be solved by leaving out inconsistent kernels:[2]

(4) $c_2(\alpha, B) = \bigcup((B \perp\!\!\!\perp \alpha) \setminus (B \perp\!\!\!\perp \bot))$

But this is insufficient, since negations are relevant.[3] We would like to have, for example: $c(\{q\}, \{p, p \to \neg q, r, r \to s, s\}) = \{p, p \to \neg q\}$. This leads to a modification:

(5) $c_3(\alpha, B) = \bigcup((B \perp\!\!\!\perp \alpha) \cup (B \perp\!\!\!\perp \neg\alpha) \setminus (B \perp\!\!\!\perp \bot))$

Note that the compartments so defined may well be inconsistent.

We now have a definition of $c(\alpha, B)$ for arbitrary sentences α. Combining (1), (2), and (5) we obtain our final definition:

[2] This has the perhaps somewhat counterintuitive consequence that $c(p, \{p \wedge \neg p\}) = c(p, \{\bot\}) = \emptyset$. Unless we adopt a syntax dependent approach (like the one in Chapter 6), we will have to live with that.

[3] [Ols97], page 7: "The belief that $\neg q$ can plausibly be held to be relevant to the justification of the belief that q. If this is true and if q and $\neg q$ are both in the system, then these two beliefs must belong to the same belief module."

5.1.4. DEFINITION. *The function c is the compartmentalization function based on Cn if and only if, for all $A, B \subseteq L$:*

$$c(A, B) = \bigcup_{\alpha \in A} c(\alpha, B),$$

where $$c(\alpha, B) = \begin{cases} \emptyset \text{ if } \alpha \in Cn(\emptyset) \text{ or } \neg\alpha \in Cn(\emptyset) \\ \bigcup(((B \perp\!\!\!\perp \alpha) \cup (B \perp\!\!\!\perp \neg\alpha)) \setminus (B \perp\!\!\!\perp \bot)) \text{ otherwise.} \end{cases}$$

5.1.5. OBSERVATION.

(1) For all sets A and B of sentences, $c(A, B) = c(A, c(A, B))$.

(2) If $A \subseteq A'$ and $B \subseteq B'$, then $c(A, B) \subseteq c(A', B')$.

Formally, the definition of compartment could be extended to deal with compartmentalization functions based on an arbitrary inference operation C, by adding the subscript C to the kernel operator. In the next section we will use the compartmentalization based on Cn to obtain local versions of arbitrary inference operators. One could also generalize the definitions we give in order to localize C with compartments based on C itself. It is not clear which properties such a compartmentalization would have. In the rest of this chapter we will only deal with compartments based on the classical consequence operator Cn.

5.1.2 Defining Local Inference

Using the definition of logical compartments presented above, we can define local inference as the inference relation that considers only the relevant part of the belief set. In other words, we can use logical compartments as defined above to derive, given a (global) consequence operator, a local inference operator for each logical compartment.

5.1.6. DEFINITION. *Let C be an inference operation on L and let c be the compartmentalization function derived from the classical consequence operation Cn. Then for any set A, the A-localization of C is the inference operation C_A such that for all sets B of sentences: $C_A(B) = C(c(A, B))$.*

A set B is called A-locally C-consistent if and only if $\bot \notin C_A(B)$.

If $A = \{\alpha\}$, i.e., A is a singleton, we will write C_α for $C_{\{\alpha\}}$.

In this chapter, we will use a particular notion of inference to represent the agent's reasoning. In the framework introduced in Chapter 4, we considered a generic inference operator *Inf*. The set of the agent's implicit beliefs was given by $Inf^*(\mathbf{E})$, where $\mathbf{E}$ represented the set of the agent's explicit beliefs (his belief base). In this chapter, the set of the agent's implicit beliefs is given by $Inf^*(B)$, where $\alpha \in Inf^*(B)$ if and only if $\alpha \in Cn_\alpha(B)$. It follows from this that the agent can consistently reason about some subjects even if his belief base contains inconsistent beliefs about some other subject.

Example 15 revisited: Let p stand for the proposition "Paul is at home", q for "The lights are on", a for "Ann is the murderer", b for "Bill is the murderer", and r for "I am asleep" and let Cn be the classical consequence operator.

My belief base B after talking to Paul's neighbor contains: $\{q, q \to p, \neg p, a \lor b, \neg a, \neg b, \neg r\}$. I am interested in whether Paul is at home, that is, the relevant beliefs are $c(p, B) = \{q, q \to p, \neg p\}$. Even though this set is inconsistent we have that $r \notin Cn_r(B) = Cn(c(r, B)) = Cn(\{\neg r\})$.

In the example, we see that the agent can consistently reason about r without having to notice that his beliefs about p are inconsistent.

From Definition 5.1.6 it follows that:

5.1.7. Observation. *Let C be an inference operation that satisfies monotony, compactness, and inclusion ($X \subseteq C(X)$). Then:*

1. $\beta \in C_\alpha(B)$ *iff* $\beta \in C_{\neg\alpha}(B)$
2. *If B is C-consistent, then:*
   - *if* $\alpha \in A$, *then* $\alpha \in C_A(B)$ *iff* $\alpha \in C(B)$
   - $C_B(B) = C(B)$
3. $\alpha \in C_\top(B)$ *iff* $\alpha \in C_\bot(B)$ *iff* $\alpha \in C(\emptyset)$
4. *If all elements of B are contingent, then* $C_L(B) = C(B)$
5. $C_A(B) \subseteq C(B)$.

5.1.8. Observation. *The elements of $B \perp\!\!\!\perp_{C_A} \alpha$ are subsets of $c(A, B)$.*

5.1.9. Observation. *For all sets B of sentences it holds that $(C_A)_A(B) = C_A(B)$.*

A number of authors [Gab85, Mak89, Lin91] have studied the formal properties of inference operations. The following is a list of properties found in the literature. An inference operation C satisfies:

Monotony if and only if $B \subseteq D$ implies $C(B) \subseteq C(D)$;

Compactness if and only if for all $\alpha \in C(B)$ there is some finite $D \subseteq B$ such that $\alpha \in C(D)$;

Idempotency (iteration) if and only if $C(C(B)) = C(B)$;

Weak iteration if and only if $C(C(B)) \subseteq C(B)$;

Inclusion if and only if $B \subseteq C(B)$;

Embedded inclusion if and only if $C(B) \subseteq C(C(B))$;

Supraclassicality if and only if $Cn(B) \subseteq C(B)$;

Deduction property if and only if whenever $q \in C(B \cup \{p\})$, then $p \rightarrow q \in C(B)$;

Reductio ad absurdum if and only if whenever $\perp \in C(B \cup \{\neg p\})$, then $p \in C(B)$;

Falsity if and only if for all p, $p \in C(\{\perp\})$;

Consistency preservation if and only if $\perp \notin Cn(B)$ implies $\perp \notin C(B)$;

Cumulativity if and only if $B \subseteq D \subseteq C(B)$ implies $C(D) = C(B)$;

Distributivity if and only if $C(B) \cap C(D) \subseteq C(Cn(B) \cap Cn(D))$;

Explosiveness if and only if for all α and β, $\beta \in C(\{\alpha, \neg\alpha\})$;

Weak Explosiveness if and only if whenever $\perp \in C(B)$, then for all sentences α, $\alpha \in C(B)$;

Non-contravention if and only if, for all α, if $\neg\alpha \in C(B \cup \{\alpha\})$, then $\neg\alpha \in C(B)$; and

α-local non-contravention if and only if, if $\neg\alpha \in C(B \cup \{\alpha\})$, then $\neg\alpha \in C(B)$.

The following are some of the properties the local inference operator C_A has.

5.1.10. THEOREM. *Let C_A be the A-localization of an inference operation C. Then:*

1. *If C satisfies monotony, then C_A satisfies monotony.*
2. *If C satisfies monotony and compactness, then C_A satisfies compactness.*
3. *If C satisfies monotony, compactness, and weak iteration, then C_A satisfies weak iteration.*
4. *If C satisfies monotony, compactness, weak iteration, and inclusion, then C_A satisfies idempotency (iteration) and cumulativity.*

5. *If C satisfies monotony and consistency preservation, then so does C_A.*

6. *If C satisfies weak explosiveness, than so does C_A.*

7. *If C satisfies monotony and inclusion, then C_A satisfies embedded inclusion.*

8. *If C satisfies monotony, compactness, weak explosiveness, and non-contravention, then C_A satisfies α-local non-contravention for all $\alpha \in A$.*

5.1.11. THEOREM. *For each of the following properties there is some set of sentences A such that Cn_A, the A-localization of the classical truth functional consequence operator Cn, does not satisfy the property: inclusion, supraclassicality, deduction property, reductio ad absurdum, falsity, distributivity, explosiveness, and non-contravention.*

Note that since $C_L = C$ when applied on contingent sets, we cannot gain any new structural properties by localizing an inference operator. Usually some properties are lost in the localization.

The classical consequence operator Cn satisfies inclusion, idempotency and monotony. Nonmonotonic logics leave out monotony, resource logics [Gab96] leave out idempotency. Our framework leaves out inclusion. Inclusion and supraclassicality are clearly not wanted, since the purpose of the operator C_A is to ignore irrelevant data.

That explosiveness does not hold agrees with the intuition that some inconsistencies are "less harmful" than others. If an inconsistency is completely irrelevant to the current reasoning, then it should not trivialize the belief state, causing the agent to believe every sentence in the language. On the other hand, if the inconsistency turns out to be relevant, that is, if given a belief base B and a set A of sentences it is the case that $\perp \in C_A(B)$, then the relation C_A becomes "locally" explosive, giving an indication that something has to be repaired.

5.2 The Generalized Belief Revision Operators

In this section we present some generalized types of belief revision operators that can be obtained using the notion of local inference. It turns out that for this purpose, we do not need the full set of properties of the local inference operator introduced above. The results to be presented here are much more general, since they only make use of some of the properties shown to hold for our local inference operator.

All definitions and results in this section refer to a generic inference operator C that may or may not be obtainable as a localization of some other inference operator. Most of the results to be presented in this section are generalizations of results that have been obtained previously for the special case when C is a

Tarskian consequence operator that satisfies the standard properties of supraclassicality, compactness and the deduction property.

Makinson [Mak87] has already remarked that for AGM partial meet contraction, only the recovery postulate requires the logic to be supraclassical. The other five basic postulates hold for partial meet contraction on the assumption that the logic is compact.

5.2.1 Contraction

An operation of *contraction* based on an inference operator C is an operation that, given a set B of sentences and a sentence α, returns a subset of B that does not imply α according to C, except in the limiting case when $\alpha \in C(\emptyset)$. We are going to consider two types of contraction operators, kernel contraction and partial meet contraction.

The idea behind kernel contraction is that, if we remove from the belief base B at least one element of each α-kernel (minimal subset of B that implies α), then we obtain a belief base that does not imply α [Han94]. To perform these removals of elements, we use a generalization of the definition of incision function (Definition 3.4.9), i.e., a function that selects at least one sentence from each kernel:

5.2.1. DEFINITION. *Let C be an inference operator. An* **incision function** *for B is any function σ such that for any formula α:*

1. $\sigma(B \perp\!\!\!\perp_C \alpha) \subseteq \bigcup(B \perp\!\!\!\perp_C \alpha)$, *and*
2. *If $\emptyset \neq X \in B \perp\!\!\!\perp_C \alpha$, then $X \cap \sigma(B \perp\!\!\!\perp_C \alpha) \neq \emptyset$.*

The following definition is a generalization of the definition given in [Han94]:

5.2.2. DEFINITION. *Let C be an inference operation on L and σ an incision function. The kernel contraction on B determined by C and σ is the operation $\dot{-}_{C,\sigma}$ such that for all sentences α:*
$B \dot{-}_{C,\sigma} \alpha = B \setminus \sigma(B \perp\!\!\!\perp_C \alpha)$

An axiomatic characterization of kernel contraction was obtained in [Han94] for a conventional, Tarskian, supraclassical, and compact consequence operator. Surprisingly enough, essentially the same characterization can be obtained with an inference operator that is only required to satisfy monotony and compactness.

5.2.3. THEOREM. *Let C be an inference operation satisfying monotony and compactness. Then $\dot{-}$ is an operation of kernel contraction on B determined by C and some incision function if and only if for all sentences α:*

- *If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)*

- $B\dot{-}\alpha \subseteq B$ *(inclusion)*
- *If* $\beta \in B \setminus B\dot{-}\alpha$*, then there is some* $B' \subseteq B$ *such that* $\alpha \notin C(B')$ *and* $\alpha \in C(B' \cup \{\beta\})$ *(core-retainment)*
- *If for all subsets* B' *of* B*,* $\alpha \in C(B')$ *if and only if* $\beta \in C(B')$*, then* $B\dot{-}\alpha = B\dot{-}\beta$ *(uniformity)*

In partial meet contraction [AGM85], the starting-point is the set of maximal subsets not implying the sentence to be contracted. A selection function (Definition 3.2.3) is used to choose some of these maximal sets, and their intersection is taken as the outcome of the operation.

A remainder operator $\perp_C$ returns for each set B of sentences and each sentence α the maximal subsets of B that do not imply α according to C. We generalize Definition 3.2.1, due to [AM82] as follows:

5.2.4. Definition. *Let* C *be an inference operation on* L*. The remainder operation* $\perp_C$ *is the operation such that for all subsets* B *and elements* α *of* L*,* $X \in B\perp_C\alpha$ *if and only if:*

1. $X \subseteq B$*,*
2. $\alpha \notin C(X)$*, and*
3. $\alpha \in C(Y)$ *for all* Y *such that* $X \subset Y \subseteq B$*.*

We also generalize Definition 3.2.3:

5.2.5. Definition. *Let* C *be an inference operation on* L*. A* C**-selection function** *for* X *is a function* γ *such that:*

- *If* $X\perp_C\alpha \neq \emptyset$*, then* $\emptyset \neq \gamma(X\perp_C\alpha) \subseteq X\perp_C\alpha$*.*
- *Otherwise,* $\gamma(X\perp_C\alpha) = \{X\}$*.*

If it is clear from the context which inference operation is used, we will refer to γ *simply as a selection function.*

5.2.6. Observation. *[AM82](Upper bound property) Let* C *satisfy monotony and compactness. If* $X \subseteq B$ *and* $\alpha \notin C(X)$*, then there is some* X' *such that* $X \subseteq X' \in B\perp_C\alpha$*.*

5.2.7. Definition. *[AGM85] The partial meet base contraction operator on* B *based on an inference operator* C *and a selection function* γ *is the operator* $\dot{-}_{C,\gamma}$ *such that for all sentences* α*:*

$B\dot{-}_{C,\gamma}\alpha = \bigcap \gamma(B\perp_C\alpha)$.

The axiomatic characterization of partial meet contraction for a conventional consequence operator obtained in [Han92b] turns out to be generalizable to the same general category of inference operations that was referred to in Theorem 5.2.3.

5.2.8. THEOREM. *Let C satisfy monotony and compactness. Then $\dot{-}$ is an operator of partial meet contraction on B based on C if and only if for all sentences α:*

- *If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)*
- *$B \dot{-} \alpha \subseteq B$ (inclusion)*
- *If $\beta \in B \setminus (B \dot{-} \alpha)$, then there is some B' such that $B \dot{-} \alpha \subseteq B' \subseteq B$, $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\beta\})$ (relevance)*
- *If for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)*

As in Observation 3.4.12, it is easy to see that all partial meet contractions can be defined as kernel contractions, but the converse does not hold.

5.2.2 Consolidation

By an operation of *consolidation* based on the inference operation C, we mean an operation that, given a set of formulas B, returns a subset of B that is consistent according to C. Consolidation can be modeled as a contraction by falsum. The two contraction operators introduced in the previous section provide us with two consolidation operators.

The idea behind kernel consolidation is that, if we remove from the belief base at least one element of each inconsistent kernel (inclusion-minimal subset of the base that implies $\perp$), then we obtain a consistent belief base.

5.2.9. DEFINITION. *[Han97a] Let C be an inference operation on the language L and σ an incision function. Then the kernel consolidation operation for B determined by C and σ is the operation $!_{C,\sigma}$ such that:*
$B!_{C,\sigma} = B \setminus \sigma(B \perp\!\!\!\perp_C \perp)$.

The following characterization is a generalization of the result obtained in [Han97a]:

5.2.10. THEOREM. *Let C be an inference operation satisfying monotony, compactness, and $\perp \notin C(\emptyset)$. An operation $!$ is an operation of kernel consolidation for B determined by C and some incision function if and only if:*

- *$\perp \notin C(B!)$ (consistency)*

- $B! \subseteq B$ *(inclusion)*
- *If* $\alpha \in B \setminus (B!)$, *then there is some* X *such that* $X \subseteq B$, $\perp \notin C(X)$ *and* $\perp \in C(X \cup \{\alpha\})$ *(core-retainment)*

Consolidation can also be constructed from partial meet contraction.

5.2.11. DEFINITION. *[Han91] The partial meet consolidation operator for* B *based on an inference operator* C *and a selection function* γ *is the operator* $!_{C,\gamma}$ *such that:*
$B!_{C,\gamma} = \bigcap \gamma(B \perp_C \perp)$

The following theorem is a generalization of a theorem reported in [Han91]:

5.2.12. THEOREM. *Let* C *satisfy monotony, compactness, and* $\perp \notin C(\emptyset)$. *An operation* ! *is an operation of partial meet consolidation based on* C *and some selection function if and only if for all sets* B *of sentences:*

- $\perp \notin C(B!)$ *(consistency)*
- $B! \subseteq B$ *(inclusion)*
- *If* $\alpha \in B \setminus (B!)$, *then there is some* X *such that* $B! \subseteq X \subseteq B$, $\perp \notin C(X)$ *and* $\perp \in C(X \cup \{\alpha\})$ *(relevance)*

5.2.3 External and Internal Revision

Revision can be obtained by combining expansion with either kernel or partial meet contraction to obtain either internal or external revision. This leaves us with four options.

An operation of internal kernel revision by a sentence α consists of first using kernel contraction to contract the belief base by $\neg\alpha$ in order to make room for the new piece of information and then expanding by α. Formally:

5.2.13. DEFINITION. *The internal kernel revision on* B *based on an inference operator* C *and an incision function* σ *is the operator* $\mp_{C,\sigma}$ *such that for all sentences* α*:*
$B \mp_{C,\sigma} \alpha = (B \setminus \sigma(B \perp\!\!\!\perp_C \neg\alpha)) \cup \{\alpha\}$

The following theorem characterizes this operation:

5.2.14. THEOREM. *Let* C *satisfy monotony and compactness. An operator* $\mp$ *is an operator of internal kernel revision based on an inference operator* C *if and only if, for all sets* B *of sentences and all sentences* α *such that* C *satisfies* α*-local non-contravention:*

- *If* $\neg\alpha \notin C(\emptyset)$, *then* $\neg\alpha \notin C(B \mp \alpha)$ *(non-contradiction)*
- $B \mp \alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If* $\beta \in B \setminus B \mp \alpha$, *then there is some* $B' \subseteq B$ *such that* $\neg\alpha \notin C(B')$ *and* $\neg\alpha \in C(B' \cup \{\beta\})$ *(core-retainment)*
- $\alpha \in B \mp \alpha$ *(success)*
- *If for all* $B' \subseteq B$, $\neg\alpha \in C(B')$ *if and only if* $\neg\beta \in C(B')$, *then* $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$ *(uniformity)*

In an operation of internal partial meet revision by α, the belief base is first partial meet contracted by $\neg\alpha$ and then expanded by α:

5.2.15. Definition. *[Han92b] The internal partial meet revision of* B *based on an inference operator* C *and a selection function* γ *is the operator* $\mp_{C,\gamma}$ *such that for all sentences* α*:*
$B \mp_{C,\gamma} \alpha = \bigcap \gamma(B \perp_C \neg\alpha) \cup \{\alpha\}$

The following theorem characterizes this operation:

5.2.16. Theorem. *[Han92b] Let* C *be an inference operator satisfying monotony and compactness. An operator* $\mp$ *is an operator of internal partial meet revision based on* C *if and only if, for all sets* B *of sentences and all sentences* α *such that* C *satisfies* α*-local non-contravention:*

- *If* $\neg\alpha \notin C(\emptyset)$, *then* $\neg\alpha \notin C(B \mp \alpha)$ *(non-contradiction)*
- $B \mp \alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If* $\beta \in B \setminus B \mp \alpha$, *then there is some* B' *such that* $B \mp \alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(B')$ *but* $\neg\alpha \in C(B \cup \{\beta\})$ *(relevance)*
- $\alpha \in B \mp \alpha$ *(success)*
- *If for all* $B' \subseteq B$, $\neg\alpha \in C(B')$ *if and only if* $\neg\beta \in C(B')$, *then* $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$ *(uniformity)*

In an operation of external kernel revision by α, the belief base is first expanded with α and then kernel contraction by $\neg\alpha$ takes place:

5.2.17. Definition. *The external kernel revision of* B *based on an inference operator* C *and an incision function* σ *is the operator* $\pm_{C,\sigma}$ *such that for all sentences* α*:*
$B \pm_{C,\sigma} \alpha = (B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha)$

The following theorem characterizes this operation:

5.2.18. THEOREM. *Let C be an inference operator satisfying monotony and compactness. An operator $\pm$ is an operator of external kernel revision based on an inference operator C if and only if, for all sets B of sentences and all sentences α such that C satisfies α-local non-contravention:*

- *If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \pm \alpha)$ (non-contradiction)*
- *$B \pm \alpha \subseteq B \cup \{\alpha\}$ (inclusion)*
- *If $\beta \in B \setminus B \pm \alpha$, then there is some $B' \subseteq B \cup \{\alpha\}$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$ (core-retainment)*
- *$\alpha \in B \pm \alpha$ (success)*
- *If α and β are elements of B and it holds for all $B' \subseteq B$ that $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ (weak uniformity)*
- *$B + \alpha \pm \alpha = B \pm \alpha$ (pre-expansion)*

In an operation of external partial meet revision by α, the belief base is first expanded with α and then partial meet contraction by $\neg\alpha$ takes place:

5.2.19. DEFINITION. *[Han92b] The external partial meet revision of B based on an inference operator C and a selection function γ is the operator $\pm_{C,\gamma}$ such that for all sentences α:*
$B \pm_{C,\gamma} \alpha = \bigcap \gamma((B \cup \{\alpha\}) \perp_C \neg\alpha)$.

The following theorem characterizes this operation:

5.2.20. THEOREM. *Let C be an inference operator satisfying monotony and compactness. An operator $\pm$ is an operator of external partial meet revision based on an inference operator C if and only if, for all sets B of sentences and sentences α such that C satisfies α-local non-contravention:*

- *If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \pm \alpha)$ (non-contradiction)*
- *$B \pm \alpha \subseteq B \cup \{\alpha\}$ (inclusion)*
- *If $\beta \in B \setminus B \pm \alpha$, then there is some B' such that $B \pm \alpha \subseteq B' \subseteq B \cup \{\alpha\}$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$ (relevance)*
- *$\alpha \in B \pm \alpha$ (success)*
- *If α and β are elements of B and it holds for all $B' \subseteq B$ that $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ (weak uniformity)*
- *$B + \alpha \pm \alpha = B \pm \alpha$ (pre-expansion)*

5.2.4 Semi-Revision

Semi-revision consists of two steps: first the belief α is added to the base, and then the resulting base is consolidated. The two consolidation operators introduced above give rise to two semi-revision operators.

5.2.21. DEFINITION. *[Han97a] The kernel semi-revision of B based on an inference operator C and an incision function σ is the operator $?_{C,\sigma}$ such that for all sentences α:*
$B?_{C,\sigma}\alpha = (B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \bot)$

The following theorem is a generalization of a result reported in [Han97a]:

5.2.22. THEOREM. *Let C be an inference operation satisfying monotony, compactness, and $\bot \notin C(\emptyset)$. An operator $?$ is an operator of kernel semi-revision based on C if and only if for all sets B of sentences and sentences α:*

- $\bot \notin C(B?\alpha)$ *(consistency)*
- $B?\alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If $\beta \in B \setminus B?\alpha$, then there is some $B' \subseteq B \cup \{\alpha\}$ such that $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\beta\})$ (core-retainment)*
- $(B + \alpha)?\alpha = B?\alpha$ *(pre-expansion)*
- *If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)*

5.2.23. DEFINITION. *[Han97a] The partial meet semi-revision of B based on an inference operator C and a selection function γ is the operator $?_{C,\gamma}$ such that for all sentences α:*
$B?_{C,\gamma}\alpha = \bigcap \gamma((B \cup \{\alpha\}) \bot_C \bot)$

The following theorem generalizes the characterization given in [Han97a]:

5.2.24. THEOREM. *Let C be an inference operation satisfying monotony, compactness, and $\bot \notin C(\emptyset)$. An operator $?$ is an operator of partial meet semi-revision based on C if and only if for all sets B of sentences and sentences α:*

- $\bot \notin C(B?\alpha)$ *(consistency)*
- $B?\alpha \subseteq B \cup \{\alpha\}$ *(inclusion)*
- *If $\beta \in B \setminus B?\alpha$, then there is some B' such that $B?\alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\beta\})$ (relevance)*
- $(B + \alpha)?\alpha = B?\alpha$ *(pre-expansion)*
- *If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)*

5.2.5 Levi and Harper Identities

One of the hallmarks of the AGM theory is the close connection that holds between revision and contraction operators through the Levi and Harper identities. The way we have defined contraction and internal revision, the Levi identity holds both for kernel and partial meet operations, i.e. for the former we have $B \mp_{C,\sigma} \alpha = (B \dot{-}_{C,\sigma} \neg\alpha) \cup \{\alpha\}$ and for the latter $B \mp_{C,\gamma} \alpha = (B \dot{-}_{C,\gamma} \neg\alpha) \cup \{\alpha\}$.

The Harper identity can also be shown to hold, under fairly weak conditions:

5.2.25. THEOREM. *Let $\dot{-}_C$ satisfy the inclusion and core-retainment postulates for contraction, and let $*_C$ be the internal revision operator based on $\dot{-}_C$ via the Levi identity. Let C satisfy $\neg\alpha$-local non-contravention. Then:*

$B \dot{-}_C \alpha = B \cap (B * \neg\alpha)$ *(the Harper identity)*

In [Han97a], it was shown that consolidation and semi-revision based on classical consequence are tied together by properties that are in a way analogous to the Levi and Harper identity:

$A?\alpha = (A + \alpha)!$. (the $! \to ?$ identity)
$A! = \cup\{A?\alpha | \alpha \in A\}$. (the $? \to !$ identity)

It can be seen directly from the definitions that these properties also hold for consolidation and semi-revision based on the more general inference operators discussed in this chapter, both for kernel operations (with C and σ held constant) and for partial meet operations (with C and γ held constant).

5.2.6 Local Operators

The axiomatic characterizations obtained for the generalized belief revision operators apply to any inference operator satisfying monotony and compactness (contraction), monotony, compactness and $\perp \notin C(\emptyset)$ (consolidation and semi-revision) or monotony, compactness and α-local non-contravention (external and internal revision). From Theorem 5.1.10 it follows that the A-localization of the classical consequence operator Cn, Cn_A, satisfies monotony, compactness and $\perp \notin C(\emptyset)$. Moreover, for $\alpha \in A$ it also satisfies α-local non-contravention. This means that, in order to obtain characterizations of the local operations aimed at in the beginning of this chapter, all we need to do is to substitute the generic inference operator C by Cn_A. In the case of the revision operations, extra care should be taken that in a revision by α, we must have $\alpha \in A$. When the A-compartment of B is empty, the local contraction and local consolidation operations leave B unchanged, while local internal/external revision and local semi-revision all coincide with expansion.

The following results show that local operations give the desired result, in the sense that they do not at all affect the beliefs outside the relevant compartment:

5.2.26. PROPOSITION. *Let B and A be sets of formulas and, let C_A be a local inference operator. If $\dot{-}_{C_A}$ is a partial meet base contraction on B based on C_A, then for every α it holds that $B \setminus c(A, B) \subseteq B \dot{-}_{C_A} \alpha$.*

5.2.27. PROPOSITION. *Let B and A be sets of formulas and, let C_A be a local inference operator. If $\dot{-}_{C_A}$ is a kernel contraction on B based on C_A, then for every α it holds that $B \setminus c(A, B) \subseteq B \dot{-}_{C_A} \alpha$.*

We now go back once more to our example to illustrate the effects of an operation of local consolidation.

> **Example 15 re-revisited:** Before talking to Paul's neighbor, my belief base B contained: $\{q, q \to p, a \vee b, \neg a, \neg b, \neg r\}$. The neighbor says that Paul is not home, so I locally semi-revise my belief base by $\neg p$. This means first adding $\neg p$ set-theoretically to B and then locally consolidating. Let B' be the belief base that replaces B after the expansion, that is, $B' = B \cup \{\neg p\}$. We have $c(p, B') = \{q, q \to p, \neg p\}$ and $B' \perp_{Cn_p} \perp = \{\{q \to p, q\} \cup X, \{q \to p, \neg p\} \cup X, \{q, \neg p\} \cup X\}$, where $X = \{a \vee b, \neg a, \neg b, \neg r\}$. Local partial meet consolidation gives us then, with the selection function $\gamma(B' \perp_{Cn_p} \perp) = \{\{q \to p, q, a \vee b, \neg a, \neg b, \neg r\}\}$, $B'' = \{q \to p, q, a \vee b, \neg a, \neg b, \neg r\} = B$, that is, the new information is rejected. (With another selection function the information would be accepted).

5.3 Related Work

Although the idea of local change seems to have no precedents in the belief revision theory, our proposal for local inference is closely related to at least three other segments of the logical literature: (1) studies of compartments and frames of mind, (2) paraconsistent logic, and (3) relevance logic.

5.3.1 Compartments and Frames of Mind

The idea of isolating inconsistencies by means of compartments has been studied by several authors.

Jaśkowski [Jaś69] introduced his discursive logics as a way of formalizing "what some participant of a discourse is committed to". If participant x asserts α and participant y asserts β, then α and β are true in his logics, but not $\alpha \wedge \beta$, since there may be no participant in the discourse who is committed to both assertions α and β. Assertions associated to different participants of the discourse cannot be combined.

Lewis [Lew82] also defended the idea of fragments that cannot be combined, but the different (possibly overlapping) fragments were parts of one agent's beliefs.

Fagin and Halpern [FH88] formalized a similar idea. In their Logic of Local Reasoning, an agent may have several "frames of mind". An agent believes α if she believes α in some frame of mind. This idea is formalized via an extended Kripke structure, where instead of one accessibility relation between worlds, there is a relation $\mathcal{C}$ between worlds and sets of worlds. If $\mathcal{C}(s) = \{T_1, ..., T_n\}$, then an agent in state s sometimes takes the set of worlds T_1 as possible, sometimes T_2, etc.

Benferhat et al. [BDP97] use yet another approach to isolate inconsistencies, which is very similar to the approach known in AI as WIDTIO (When In Doubt Throw It Out) [Win90]. They define a free (or sound) inference operator that disregards all formulas belonging to a minimal inconsistent subset. So, $B \vdash_{free} \alpha$ if and only if $Free(B) \vdash \alpha$, where $Free(B)$ is the result of deleting every minimal inconsistent subset from B.

Tennant [Ten84], elaborating on an idea from [Smi59], presented the idea of perfect validity, which is closely related to our idea of local inference. A sequent in a Gentzen system is a structure $X : Y$ where X and Y are sets of sentences of a language L. A sequent $X' : Y'$ is a subsequent of $X : Y$ if and only if $X' \subseteq X$ and $Y' \subseteq Y$. If, in addition, $X' \subset X$ or $Y' \subset Y$, then $X' : Y'$ is a proper subsequent of $X : Y$. A sequent $X : Y$ is valid if and only if in every model where all elements of X hold, at least one element of Y holds. A sequent is said to be perfectly valid if and only if it is valid and has no valid proper subsequent. If $X : Y$ is perfectly valid and $Y \neq \emptyset$, then we have: (i) X is consistent, otherwise $X : \emptyset$ would be a valid subsequent of $X : Y$; (ii) if $Y = \{\alpha\}$, then $X \in L \perp\!\!\!\perp \alpha \setminus L \perp\!\!\!\perp \bot$. We can then translate our notion of local inference in terms of perfect validity: $\beta \in Cn_\alpha(B)$ if and only if there are perfectly valid sequents $X_1 : Y_1$, ..., $X_n : Y_n$ such that each $X_i \subseteq B$, either $Y_i = \{\alpha\}$ or $Y_i = \{\neg\alpha\}$, and $\beta \in Cn(\bigcup X_i)$. More generally, $\beta \in Cn_A(B)$ if and only if there are perfectly valid sequents $X_1 : Y_1$, ..., $X_n : Y_n$ such that each $X_i \subseteq B$, either $Y_i = \{\alpha\}$ or $Y_i = \{\neg\alpha\}$ for some $\alpha \in A$, and $\beta \in Cn(\bigcup X_i)$.

5.3.2 Paraconsistent Logic

A logic is said to be paraconsistent if its consequence relation is not explosive, i.e., if there are formulas α and β such that $\beta \notin C(\{\alpha, \neg\alpha\})$. Technically, when a logic has more than one inference operation (as is the case here, since for every set A there is an associated inference operation C_A), all of them have to be non-explosive for the logic to be called paraconsistent [PRN89]. If we take A to be the whole language, C_A is explosive, so the logic of local reasoning cannot be called paraconsistent in this sense. Nevertheless, it contains paraconsistent fragments, and we believe it can capture some of the intuitions that have been the driving forces behind the development of paraconsistent logic.

In [RS95], an approach is taken where the logic under which belief sets are closed is paraconsistent. The motivation is very close to ours, namely to make it possible to work with inconsistent beliefs without trivializing belief states. These authors use first degree entailment, a modification of (propositional) classical logic where each formula can have as its truth value any subset of $\{\texttt{true}, \texttt{false}\}$. A formula α is entailed by a set of formulas X if and only if any valuation that assigns to every element of X either $\{\texttt{true}\}$ or $\{\texttt{true}, \texttt{false}\}$, also assigns to α either $\{\texttt{true}\}$ or $\{\texttt{true}, \texttt{false}\}$.

5.3.3 Relevance Logic

One of the main characteristics of Relevance Logic [Dun86] is that from $\{\alpha, \neg\alpha \vee \beta\}$ one cannot always derive β. Depending on which set we use to limit the compartments, we can obtain the same property. We have, for example, $q \notin C_p(\{p, \neg p \vee q\})$.

There are in the literature several attempts to define the concepts of relevance and dependence. [Lug96] classifies approaches to relevant inference into two groups: those which impose conditions on the contents of implied formulas and those which impose conditions on the deduction of the implied formulas. In the first group she mentions Parry, whose inferences are valid only if no new variables are introduced into the consequent, Anderson and Belnap, who impose that the antecedent and the consequent share variables, and Epstein who associates themes to the formulas and imposes conditions on the themes of the antecedent and the consequent. In the second group she includes approaches requiring that the antecedent be used in an "essential way" in the derivation of the consequent. Among these she mentions Myhill and again Anderson and Belnap.

The way our relevant compartments are defined, through kernel sets, is related to the second approach, that is, we consider relevant to a formula α the formulas that appear in a minimal derivation of α or its negation.

Lang and Marquis [LM98] present several complexity results for different notions of dependency found in the literature.

In [dCH96], Fariñas del Cerro and Herzig introduce a set of nine postulates that any dependence relation should satisfy, and they also show how a contraction operation satisfying the Gärdenfors postulates can be obtained from a dependence relation. Given a dependence relation $\leadsto$, where $\alpha \leadsto \beta$ should be read as "β depends on α", they define a contraction operation for belief sets as follows:

(Def $\dot{-}$) $\gamma \in K \dot{-} \alpha$ if and only if either $\vdash \gamma$ or $\gamma \leadsto \gamma$ and $\alpha \not\leadsto \gamma$

They show that, given a dependence relation $\leadsto$ satisfying their set of postulates, the contraction operation defined by (Def $\dot{-}$) satisfies the eight Gärdenfors postulates for contraction. Although their motivation for studying dependency is very similar to our interest in relevance, namely isolate the part of the belief state that has to be examined for a belief change operation, the results obtained by Fariñas del Cerro and Herzig are very different from ours. The fact that they use belief sets shows already that they are not looking for realistic models.

5.4 Embedding Local Change

In this section, we show how to model local contraction, in the framework defined in Chapter 4. All the other local operations can also be defined using (non-local) expansion and local contraction. For the sake of clarity, we concentrate on the simplified model (Sections 4.1 and 4.2).

Locally contracting a belief base B by α with respect to a set of formulas A consists in giving up enough beliefs from B such that the part of the new base that is relevant for A does not imply α. Intuitively, the set A should contain the formula α, but the formalization is general enough to allow for the use of any set of formulas. The set A should be seen as a context or topic of reasoning.

Two different constructions for local contraction were presented above, together with sets of postulates that characterize them. We will now show how local partial meet contraction can be decomposed into applications of the basic operations presented in Definition 4.2.1. The idea can be easily extended to the other construction (local kernel contraction) as well as to the other local operations.

In an operation of partial meet contraction a selection function is used to select some of the remainders. The elements of the belief base that are not contained in all of the selected remainders are given up. In local partial meet contraction, the operation is restricted to a compartment of the belief base. If we want to contract a belief base B by the formula α with respect to a set of formulas A, the beliefs to be discarded are those in the A-compartment of B that are not contained in all the selected α-remainders of B.

5.4.1. Definition. *We define the retain set of B given α and A as:*

$\rho_\gamma(B, \alpha, A) = \bigcap \gamma(B \perp_{C_A} \alpha)$*, where γ is a selection function.*

The discard set of B given α and A is defined by:

$\delta_\gamma(B, \alpha, A) = c(A, B) \setminus \bigcap \gamma(B \perp_{C_A} \alpha)$*, where γ is a selection function.*

5.4.2. Definition. *Let γ be a selection function. The local partial meet contraction operator on B with respect to A is the operator $-_A$ such that for all sentences α:*

$B -_A \alpha = B \setminus \delta_\gamma(B, \alpha, A)$.[4]

The operation of local partial meet contraction leaves the irrelevant part of the belief base $(B \setminus c(A, B))$ untouched.

Let f be a function from belief bases into belief states such that for all bases B, $f(B) = \langle B, Cn, \emptyset\rangle$. We will now define an operation of local partial meet contraction on belief states that are in the image of f, i.e., belief states of the form $\langle B, Cn, \emptyset\rangle$.

[4]Since by Proposition 5.2.26 the only beliefs discarded in a local partial meet contraction of B by α are elements of $c(A, B)$, we can use $\delta_\gamma(B, \alpha, A)$ as defined.

5.4.3. DEFINITION. *Let c be a compartmentalization function and γ a selection function. The local partial meet contraction of a belief state $\Sigma = \langle B, Cn, \emptyset\rangle$ by α with respect to A is given by:*
$\Sigma -_A \alpha = \Sigma \circ_r c(A, B) \circ_d \delta_\gamma(B, \alpha, A) \circ_c \delta_\gamma(B, \alpha, A) \circ_a \rho_\gamma(B, \alpha, A)$

This operation consist of retrieving the relevant compartment and deleting the beliefs contained in the discard set. The operation of doubting removes the discard set from the set of explicit beliefs, while the operation of rejection removes the discard set from the set of active beliefs. The operation of acceptance moves the retain set into the set of explicit beliefs. Since these were already part of the set of explicit beliefs, if there is no interest in deleting these beliefs from the set of active beliefs (cf. footnote 4 in Definition 4.2.1), this step may be skipped.

The operation of local partial meet contraction of belief states has the same effect on the set of explicit beliefs as the operation defined in 5.4.2, i.e.:

5.4.4. LEMMA. *If A and B are sets of formulas, α is a formula and there is no maximum size for any set involved, then $f(B -_A \alpha) = f(B) -_A \alpha$.*

Since all other local operations can be obtained from applications of local contraction and expansion, we have that:

5.4.5. PROPOSITION. *The theory of Local Change can be embedded in the framework of belief states with the basic operations.*

We will now show how the basic operations can be combined to form an operation of local semi-revision and apply it to an example.

The operation of local partial meet semi-revision can be defined as a composition of expansion and local partial meet consolidation (contraction by falsum):

5.4.6. DEFINITION. *Let c be a compartmentalization function and γ a selection function. The local partial meet semi-revision of a belief state $\Sigma = \langle B, Cn, \emptyset\rangle$ by α in relation to A is given by:*
$\Sigma ?_A \alpha = \Sigma \circ_o \alpha \circ_r c(A, B) \circ_d \delta_\gamma(B \cup \{\alpha\}, \perp, A) \circ_c \delta_\gamma(B \cup \{\alpha\}, \perp, A) \circ_a \rho_\gamma(B \cup \{\alpha\}, \perp, A)$

We return to the example about Mary discussed in Chapter 4 in order to illustrate this operation.

> **Example 12 (continued):** Suppose Mary believes that she will be outside for a long time (q), that if she stays outside for a long time, then she should take an umbrella ($q \to p$), that the moon is not made of green cheese ($\neg a$), that she loves John (b), and that Buenos Aires is the capital of Brazil (c). Her belief base is $B = \{q, q \to p, \neg a, b, c\}$. Her belief state is given initially by: $\beta_0 = \langle B, Cn, \emptyset\rangle$. When her

mother says that she should take the umbrella (p), the new belief state is given by: $\beta_1 = \beta_0 \circ_o p = \langle B, Cn, \{p\}\rangle$. Then the relevant beliefs are retrieved from the base: $\beta_2 = \beta_1 \circ_r \{q, q \to p\} = \langle B, Cn, \{p, q, q \to p\}\rangle$. Since the set of active beliefs is consistent, nothing has to be given up (note that the rest of B could still contain inconsistencies) and the result of locally consolidating gives the same belief state ($\beta_3 = \beta_2$). The active beliefs are now accepted: $\beta_4 = \beta_3 \circ_a \{p, q, q \to p\} = \langle B \cup \{p\}, Cn, \emptyset\rangle$.

Of course the interesting case occurs when Mary's previous beliefs are inconsistent with what her mother says. Suppose she also believed that she did not have to take an umbrella, i.e., the initial belief base was $B' = \{\neg p, q, q \to p, a, b, c\}$ and the initial belief state $\beta'_0 = \langle B', Cn, \emptyset\rangle$. We get $\beta'_1 = \beta'_0 \circ_o p = \langle B', Cn, \{p\}\rangle$ and $\beta'_2 = \beta'_1 \circ_r \{\neg p, q, q \to p\} = \langle B, Cn, \{p, \neg p, q, q \to p\}\rangle$. Now we have that A, the set of active beliefs, is inconsistent. For local partial meet consolidation we get: $A \perp \bot = \{\{q, q \to p, p\}, \{\neg p, q \to p\}, \{q, \neg p\}\}$. Suppose we have that $\gamma(A \perp \bot) = \{\{q, q \to p, p\}\}$. Then the only belief given up is $\neg p$ and the new belief state is $\beta'_3 = \beta'_2 \circ_d \neg p \circ_c \neg p = \langle B' \setminus \{\neg p\}, Cn, \{q, q \to p, p\}\rangle$ and finally we have $\beta'_4 = \beta'_3 \circ_a \{p, q, q \to p\} = \langle (B' \setminus \{\neg p\}) \cup \{p\}, Cn, \emptyset\rangle$.

In the framework presented in Chapter 4, the problem of deciding which beliefs should be accepted was left open. In the particular case of a partial-meet semi-revision, Definition 5.4.6 shows which sentences are accepted, depending on the selection function γ used.

5.5 Conclusions

In this chapter we have defined a consequence operation that considers only the relevant parts of a belief base and shown that this consequence operation can be used to define local versions of the operations of contraction, consolidation, revision, and semi-revision.

These local operations can be axiomatized with plausible postulates. Furthermore, they give the desired result when applied to examples of the type that motivated our work and that cannot be treated satisfactorily in the AGM model.

The construction of compartments shows which of the explicit beliefs of an agent should be retrieved into the set of active beliefs in order to perform an operation of belief change. However, the construction makes use of kernels and is highly inefficient from the computational point of view. It is not difficult to see that the operations of local change, although intuitively more adequate, are computationally as hard as the original belief change operations. In the next chapter we will propose a more efficient method for isolating the relevant compartment.

Chapter 6

Structured Bases

In this chapter we show how the extra structure of belief bases can be used for implementing local change as defined in Chapter 5.

As we have seen, some of the constructions for operations of belief change for belief bases make use of remainders or kernels, i.e., they check subsets of the belief base in order to find maximal subsets not implying a given formula or minimal subsets implying it. Exploring all subsets of a reasonably sized belief base is a very expensive operation. One way to attack the problem is to try to reduce the size of the set to be explored. Intuitively, not all of an agent's beliefs are relevant for deciding what to do with new information. There should be a way of isolating the subset of a belief base that contains the relevant beliefs for an operation of belief change. In Chapter 5, this approach was explored. A notion of *compartment* was presented for retrieving the relevant part of a belief base and then local operations of belief change were defined which act only on the relevant part of the beliefs. The problem is that the way the compartments were defined used the notion of minimal subsets of the base implying some formula, that is, finding a compartment was as expensive as performing a traditional operation of belief change. However, the representation results obtained in Chapter 5 for the local versions of the belief change operations are very general and do not depend on the particular way the compartment is defined, but only on properties of the local inference operation obtained, namely, monotony, compactness, and local non-contravention. The compartment construction in Definition 5.1.4 can be seen as an example, as one way of constructing a compartment such that the localization of the consequence operator to the compartment satisfies the requirements. If we find another way of retrieving the relevant beliefs from a base such that the associated local inference operation also satisfies the requirements, the results in Chapter 5 go through.

In this chapter we present more efficient ways of retrieving the relevant part of a belief base. The retrieved set can then be "plugged" into the constructions of local operations. Even though the local operations also rely on finding minimal or maximal subsets implying or not implying certain formulas, the set to be explored is reduced to a manageable size.

In Section 6.1, a computer system is presented which has a very efficient mechanism for retrieving relevant beliefs from a belief base with extra structure. The ideas of the system inspired some considerations on more general ways of structuring belief bases and on formalizing the notion of retrieving the most relevant beliefs for a certain input. These considerations are presented in Section 6.2. In Section 6.3, some examples of *relatedness relations* are presented which can be used for structuring the belief base. Computational aspects are examined in section 6.4. In Section 6.5, we compare our approach to the one due to Chopra et al. [CGP99].

We will use the terms "belief base" and "database" indiscriminately to refer to a set of formulas of L.

6.1 RABIT

In this section we briefly discuss the system RABIT (**R**easoning **A**bout **B**eliefs **I**n **T**ime). For details, the reader is referred to [Gar93, GED96].

RABIT is intended to simulate commonsense reasoning in an efficient, psychologically based way. The system consists of four modules, each one representing a different sort of memory: LTM (Long-Term Memory), STM (Short-Term Memory), ITM (Intermediate-Term Memory) and RTM (Relevant-Term Memory). The idea is that all the information (the beliefs) is stored in the LTM component, but the reasoning is performed inside a small subset of the beliefs in LTM, the STM. This reasoning is based on the step-logic formalism [DP86]. The ITM component stores the history of the reasoning process and the RTM works as a kind of context, storing the relevant concepts (in RABIT, any symbol in the language, like predicate names and constants, represents a concept). The idea of working on a small subset of the belief base is psychologically as well as computationally motivated.

RABIT has a very efficient algorithm for retrieving information from the LTM which is relevant for a certain input. The LTM is organized as a bipartite graph where the nodes are either formulas or concepts. Each formula is linked to all the concepts that occur in it. So (example from [GED96]), the formula *penguin(tweety)* is linked to the nodes *penguin* and *tweety*. If one wants to retrieve every formula that mentions Tweety, one only needs to take all nodes adjacent to the concept *tweety* in the graph.

A method of marker-passing is used for determining which formulas of the LTM go to STM. The system selects the relevant concepts to start with (for example, the concepts occurring in a new belief acquired) and "spreads" the activation through the edges of the graph. The activation level decreases every time it passes a new edge. Every belief with an activation level above a certain threshold is copied into the STM. By regulating the different parameters, like the initial activation, the decrease function and the threshold of activation, one can keep the STM small.

The LTM may contain contradictions, that are only solved when they become evident in the STM. Beliefs in the LTM as well as in the STM have their source attached. When a contradiction arises in the STM, a precedence order of the sources determines which belief should be given up from the STM. But nothing is deleted from the LTM. This means that a contradiction may be retrieved over and over again due to the same input. In a more recent version of RABIT, there is an "adaptive behavior" that increases the distance between the formula that "lost" in the contradiction solving procedure and any node, so that it becomes less and less likely that the contradiction will be retrieved again.

The RABIT architecture is intuitively very appealing, but there are some aspects of it that prevent it from being useful for our purposes:

1. The retrieval mechanism is very efficient, but after the relevant part of the LTM has been retrieved, the system starts performing all possible inferences from the beliefs in the STM, i.e., the system is not goal-directed.

2. The LTM only grows, new beliefs are added, but nothing is deleted. Besides the fact that this presents obvious disadvantages from the computational point of view, it does not seem very intuitive that an agent, after concluding that a belief is false, continues to hold it for the same reasons. The belief could certainly be reinferred by another line of reasoning but we do not want a system that always follows the same line of reasoning (doomed to failure).[1]

3. The way in which the LTM is organized is purely syntactical. There are no links between concepts, for example. If we know that Paul is a lecturer, that a lecturer is a member of staff and that members of staff are people, there is no "short path" between beliefs about people in general and beliefs about Paul.

The aspects above show that the reasoning part of RABIT is not adequate for our purposes. Nevertheless, the architecture of the program is going to serve as a basis for the rest of this chapter. RABIT's architecture is based on cognitive models of memory and the fact that these models use the notions of small short-term memory and relevance links between beliefs provides an independent motivation for our model of a rational agent's belief state. In the next two sections, we will explore and generalize those aspects of the RABIT architecture that will allow us to construct more efficient operations for belief change.

[1]In some proposals for belief revision [Rya92, Rot96, CGP99], beliefs are never deleted and new beliefs are simply appended to a list without any consistency check. But in these approaches, in contrast to RABIT, the beliefs are ordered according to priority which is used by the inference mechanism in order to avoid inconsistencies.

6.2 Structuring the Belief Base

As we have seen, if we have an efficient method of retrieving relevant beliefs from a belief base, such that the inference relation obtained by localizing an inference operator C is monotonic and compact, we can use it together with the representation results in Chapter 5 and get efficient, well defined operations of belief change.

In this section we show how to use extra structure of belief bases in order to make the retrieval of the relevant beliefs more efficient.

We begin by assuming that a relatedness relation between formulas of the language is given, with the following intended meaning:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ are directly related.

For the moment we leave open what we mean by "directly related". It may have a psychological interpretation given by statements or concepts that individuals associate, a semantic interpretation such as φ and ψ are assertions about the same (or related) topic, an interpretation in terms of causal connections, etc. In Section 6.3 we give some examples of relatedness relations.

Some intuitively desirable properties of $\mathcal{R}$ are[2]:

1. reflexivity – $\mathcal{R}(\varphi, \varphi)$;
2. symmetry – $\mathcal{R}(\varphi, \psi) \Leftrightarrow \mathcal{R}(\psi, \varphi)$;
3. negation invariance – $\mathcal{R}(\varphi, \neg\varphi)$.

Transitivity is not desirable for two reasons: first, we want to be able to model agents with limited resources, who may be unable to calculate the transitive closure of the relatedness function. Second, we want to be able to talk about degrees of relatedness.

Given a relatedness relation, we can represent a belief base as a (possibly disconnected) graph where each node is a formula and there is an edge between φ and ψ if and only if $\mathcal{R}(\varphi, \psi)$. This graph representation gives us immediately a notion of degree of relatedness: the shorter the path between two formulas of the base, the more related they are. Another notion made clear is that of connectedness: the connected components partition the graph into unrelated "topics" or "subjects". Sentences in the same connected component are somehow related, even if far apart. Formally:

6.2.1. DEFINITION. *Let B be a belief base and $\mathcal{R}$ be a relation between formulas. An* **$\mathcal{R}$-path** *between two formulas φ and ψ in a belief base B is a sequence $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ of formulas such that:*

1. $\varphi_0 = \varphi$ and $\varphi_n = \psi$

[2] We will see later in this chapter that even these properties are not necessary.

2. $\{\varphi_1, ..., \varphi_{n-1}\} \subseteq B$

3. $\mathcal{R}(\varphi_i, \varphi_{i+1})$*,* $0 \leq i < n$*.*

If it is clear from the context to which relation we refer, we will talk simply about a path in B.

We represent the fact that P is a path between φ and ψ by $\varphi \overset{P}{\rightsquigarrow} \psi$.

The **length** *of a path $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ is $l(P) = n$*

Note that the extremities of a path in B are not necessarily elements of B. This is due to the fact that we may want to check whether some sentence outside the belief base is related to elements of the belief base.

6.2.2. DEFINITION. *Let B be a belief base and $\mathcal{R}$ a relation between formulas of the language. We say that two formulas φ and ψ are* **related in** *B by $\mathcal{R}$ if and only if there is an $\mathcal{R}$-path P in B such that $\varphi \overset{P}{\rightsquigarrow} \psi$.*

Given two formulas φ and ψ and a belief base B, we can use the length of the shortest path between them in B as the degree of unrelatedness of the formulas. If the formulas are not related in B, the degree of unrelatedness is set to infinity. Formulas with a shorter path between them in B are more closely related in B.

6.2.3. DEFINITION. *Let B be a belief base, $\mathcal{R}$ a relation between formulas of the language and φ and ψ formulas. The* **degree of unrelatedness** *of φ and ψ in B is given by:*

$$u(\varphi, \psi) = \begin{cases} 0 & \text{if } \varphi = \psi \text{ and } \varphi \in B \\ min\{l(P) | \varphi \overset{P}{\rightsquigarrow} \psi, P \text{ in } B\} & \text{if } \varphi,\ \psi \text{ related in } B \text{ by } \mathcal{R},\ \varphi \neq \psi \\ \infty & \text{otherwise} \end{cases}$$

6.2.4. OBSERVATION. *If the relation $\mathcal{R}$ is symmetric, then the unrelatedness degree u restricted to connected components of B is a distance function in the sense that: (i) $u(x, y) \geq 0$; (ii) $u(x, y) = 0$ iff $x = y$; (iii)$u(x, y) = u(y, x)$; and (iv) $u(x, y) \leq u(x, z) + u(z, y)$.*

In Figure 6.1 we see an example of a belief base structured by a relation $\mathcal{R}$ defined by:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share (at least) one atom.

This is just an example of a relatedness relation. This relation is clearly too simplistic to capture a cognitive notion of relevance. Nevertheless, it has some interesting properties which make it a good starting point for studying relevance. The intuitive interpretation of $\mathcal{R}(\varphi, \psi)$ is: "Given a formula φ, consider every formula ψ in which we believe and that involves at least on topic mentioned in φ".

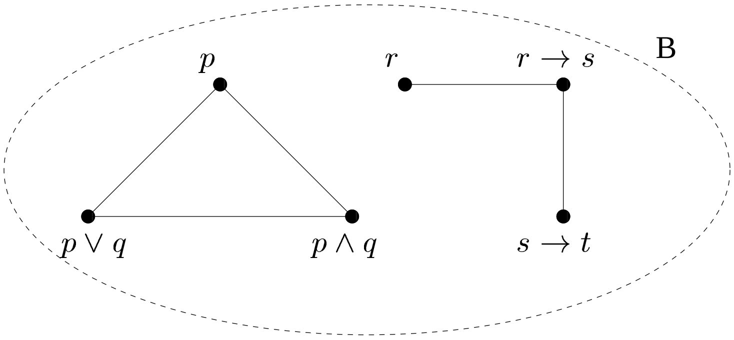

Figure 6.1: Structured belief base

The edges in the figure represent the relation $\mathcal{R}$ between the elements of the belief base. The graph representing the structured base has two disconnected components. The formulas $p \wedge q$ and $r \to s$ are not related in B. The formula $q \wedge r$ is related to all formulas in the belief base, since it is possible to find a path between it and all formulas in the base. Remember that the initial and final nodes of the path do not have to be elements of the base. This means that if one adds the formula $q \wedge r$ to the base, all the other elements become related to each other in B. This can be interpreted in the following way: the two disconnected components of the original base represent beliefs about unrelated subjects. As soon as one introduces a belief mentioning the subjects of both components, all beliefs in the two components become related to each other. There is an implicit assumption here that if one chooses to add a belief $p \wedge q$ instead of two beliefs p and q, then there is some relation between p and q.

We now show, given the structure of a belief base, how to retrieve the set of formulas relevant for a given formula α:

6.2.5. Definition. *The set of formulas of B which are* **relevant for α with degree** *i is given by:*

$\Delta^i(\alpha, B) = \{\varphi \in B | u(\alpha, \varphi) = i\}$ *for* $i \geq 0$

6.2.6. Definition. *The set of formulas of B which are* **relevant for α up to degree** *n is given by:*

$\Delta^{\leq n}(\alpha, B) = \bigcup_{0 \leq i \leq n} \Delta^i(\alpha, B)$ *for* $n \geq 0$

We say that $\Delta^{<\omega}(\alpha, B) = \bigcup_{i \geq 0} \Delta^i(\alpha, B)$ is the set of **relevant formulas for** *α.*

Since the extremities of paths in B do not have to be elements of B, we can retrieve the subset of B containing the formulas which are relevant for α even when $\alpha \notin B$. Adding α to B would only mean that α would also be retrieved, i.e.:

6.2.7. OBSERVATION. *For $\alpha \notin B$, it holds that:*
$\Delta^{\leq n}(\alpha, B \cup \{\alpha\}) \setminus \{\alpha\} = \Delta^{\leq n}(\alpha, B)$.

Proof: For $1 \leq i \leq n$, we have that $\Delta^i(\alpha, B \cup \{\alpha\}) \setminus \{\alpha\} = \Delta^i(\alpha, B)$ and that $\alpha \notin \Delta^i(\alpha, B)$. Since $\alpha \notin B$, $\Delta^0(\alpha, B \cup \{\alpha\}) \setminus \{\alpha\} = \{\alpha\} \setminus \{\alpha\} = \emptyset = \Delta^0(\alpha, B)$. □

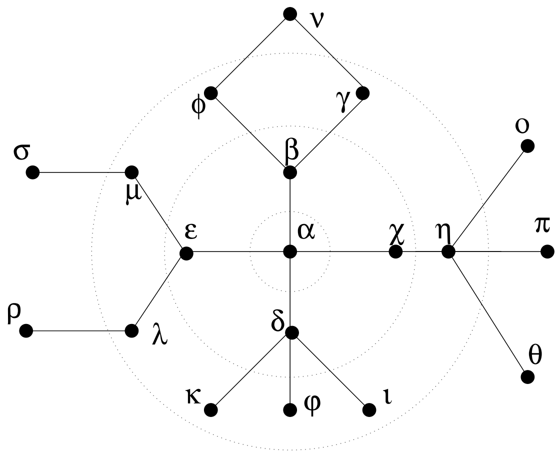

Figure 6.2: Degrees of relevance

In Figure 6.2, we see an example of a structured belief base $B = \{\alpha, \beta, \gamma, \delta, \varepsilon, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, o, \pi, \rho, \sigma, \phi, \varphi, \chi\}$. The dotted circles represent different levels of relevance for α. We have:
$\Delta^0(\alpha, B) = \{\alpha\}$
$\Delta^1(\alpha, B) = \{\beta, \chi, \delta, \varepsilon\}$
$\Delta^2(\alpha, B) = \{\gamma, \eta, \iota, \varphi, \kappa, \lambda, \mu, \phi\}$
$\Delta^3(\alpha, B) = \{\nu, o, \pi, \theta, \rho, \sigma\}$
$\Delta^{\leq\omega}(\alpha, B) = \Delta^0(\alpha, B) \cup \Delta^1(\alpha, B) \cup \Delta^2(\alpha, B) \cup \Delta^3(\alpha, B) = B$

In Section 5.1 we presented a notion of compartment around a sentence. In Definition 5.1.4 the α-compartment of a belief base B was defined as the set of formulas of B that contribute to proving either α or its negation. These were the formulas of B that were relevant for an operation of belief change. As we mentioned before, for the representation results obtained in Chapter 5 the particular construction of the compartments does not matter, only the properties of the inference relation obtained is significant. We now show some properties of the inference operation obtained using $\Delta^{\leq n}$ instead of c (Definition 5.1.4) as a compartmentalization function.

6.2.8. PROPOSITION. *Let B be a finite belief base and α a formula. For every natural number n and any inference operator C, if C is monotonic and compact, then the local inference operations defined as $C^n_\alpha(B) = C(\Delta^{\leq n}(\alpha, B))$ are monotonic and compact. Moreover, if $\perp \notin C(\emptyset)$, then $\perp \notin C^n_\alpha(\emptyset)$ and if C satisfies non-contravention, then C^n_α satisfies α-local non-contravention (if $\neg\alpha \in C^n_\alpha(B \cup \{\alpha\})$, then $\neg\alpha \in C^n_\alpha(B)$).*

Proof: Since all sets considered are finite, compactness follows trivially. For monotony, let B and D be sets of formulas such that $B \subseteq D$. Let $\beta \in C^n_\alpha(B)$, i.e., $\beta \in C(\Delta^{\leq n}(\alpha, B))$. It is easy to see that $\Delta^{\leq n}(\alpha, B) \subseteq \Delta^{\leq n}(\alpha, D)$. Hence, since C is monotonic, $\beta \in C(\Delta^{\leq n}(\alpha, D))$, i.e., $\beta \in C^n_\alpha(D)$.

Suppose that $\perp \in C^n_\alpha(\emptyset)$. Then $\perp \in C(\Delta^{\leq n}(\alpha, \emptyset)) = C(\emptyset)$. Hence, if $\perp \notin C(\emptyset)$, then $\perp \notin C^n_\alpha(\emptyset)$. For α-local non-contravention, let $\neg\alpha \in C^n_\alpha(B \cup \{\alpha\})$. This means that $\neg\alpha \in C(\Delta^{\leq n}(\alpha, B \cup \{\alpha\}))$. If $\alpha \in B$, then it follows trivially that $\neg\alpha \in C^n_\alpha(B)$. Let $\alpha \notin B$. By the non-contravention of C, $\neg\alpha \in C(\Delta^{\leq n}(\alpha, B \cup \{\alpha\}) \setminus \{\alpha\})$. Since by Observation 6.2.7 $\Delta^{\leq n}(\alpha, B \cup \{\alpha\}) \setminus \{\alpha\} = \Delta^{\leq n}(\alpha, B)$, $\neg\alpha \in Cn(\Delta^{\leq n}(\alpha, B))$ and hence, $\neg\alpha \in C^n_\alpha(B)$. □

We are mostly interested in finite belief bases, but if we want to consider infinite bases, it suffices to define C^n_α as:

$$C^n_\alpha(B) = \begin{cases} C(\Delta^{\leq n}(\alpha, B)) & \text{if } B \text{ is finite} \\ \{\phi | \phi \in C^n_\alpha(B') \text{ for some finite } B' \subseteq B\} & \text{otherwise} \end{cases}$$

The proposition above implies that for any n, $\Delta^{\leq n}$ can be used as a compartmentalization function to define local operations characterized by the results presented in Chapter 5. In the case of an ideal agent, an agent with no limits in its capacity of retrieval, all of the agent's relevant beliefs are retrieved, i.e., $\Delta^{\leq \omega}$ is used as a compartmentalization function. Otherwise, we can limit the size of the set retrieved by the choice of n.

It is not difficult to see that we can depart from any relation $\mathcal{R}$ to obtain a local inference operator satisfying the relevant properties. For the proof of Proposition 6.2.8, no properties of $\mathcal{R}$ were used, we only needed conditions on the initial inference operator C. This gives us a very general framework. Depending on the application, one can use the most appropriate notion of relatedness and still obtain a local inference operator satisfying the properties needed for the representation results given in Chapter 5.

6.3 Where Does the Structure Come from?

In the previous section we assumed that a relation of relatedness between elements of the language was given. In this section we will present some ways in which such a relation may be derived.

6.3.1 The Syntactic Approach

Rodrigues [Rod97] claims that asking for a primitive relation between propositional variables or a primitive subject matter assignment may be too strong a requirement. He proposes instead to use directly the relation introduced in the text preceding Figure 6.1, i.e.:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share an atom

He shows that this is actually the smallest relation satisfying the conditions given in [Eps90] for a relatedness relation.[3]

Epstein [Eps90] takes the topic or subject matter of a proposition as a primitive and says that two propositions are related if they share a subject matter. He considers a set of topics $\mathcal{S}$ and supposes that each propositional variable has a non-empty set of topics associated with it. The set of topics associated with a formula is simply the union of the sets of topics associated with each variable appearing in it. Epstein shows that one can also take a relatedness relation $\mathcal{R}$ as a primitive and derive a topic assignment from it by taking the subject matter of a formula φ to be $s(\varphi) = \{\{\varphi, \psi\} | \mathcal{R}(\varphi, \psi)\}$. Clearly, if we take $\mathcal{R}_s$ to be defined as $\mathcal{R}_s(\varphi, \psi)$ if and only if $s(\varphi) \cap s(\psi)$ is non-empty, then we have $\mathcal{R}_s = \mathcal{R}$.

Another way of deriving the relatedness structure from the given data is to consider the knowledge base as a whole and create a graph that has as nodes formulas and atoms. Each formula is then linked to the atoms that occur in it. This is simmilar to the approach followed by the RABIT system. The derived structure is that of a (possibly disconnected) bipartite graph. Atoms can be related to each other only via some formula in the database that contains them. In the same way, formulas are only related to other formulas via some atom they share. This has the advantage that it is very easy to insert a new formula in the graph, one only has to link it to the atoms it contains. On the other hand, since there is no link between atoms, even if the atoms p and q are intuitively directly related, the formulas p and q will only be related if there is any formula in the belief base containing both. One could use tricks, like adding some formula *related(p,q)* to the belief base if the agent believes that p and q are related or adding a tautology like $(p \vee \neg p) \vee (q \vee \neg q)$ to the belief base in order to relate p and q.

Epstein's and RABIT's approaches have the property that the relatedness relation is defined on the whole language and is independent of the particular belief base. This means that adding beliefs to the structure consists of adding the input formula α to the set of nodes and adding an edge between α and

[3] According to [Eps90], any relatedness relation R should satisfy:
R1 – $R(\varphi, \varphi)$
R2 – $R(\varphi, \psi)$ iff $R(\neg\varphi, \psi)$
R3 – $R(\varphi, \psi)$ iff $R(\psi, \varphi)$
R4 – $R(\varphi, \gamma \rightarrow \psi)$ iff $R(\varphi, \gamma)$ or $R(\varphi, \psi)$
R5 – $R(\varphi, \gamma \wedge \psi)$ iff $R(\varphi, \gamma \rightarrow \psi)$.

every node β such that $\mathcal{R}(\alpha, \beta)$. Deleting a belief α from the structure consists of removing α from the set of nodes and removing all the edges leaving from it. After adding or deleting a belief, the result is another structured base. The relation $\mathcal{R}$ does not have to be recalculated.

6.3.2 The Database Approach

In many real applications, there is a part of the database dedicated to the definition of concepts (sometimes called the T-Box) which can be used to generate a relatedness relation.

Let B be a database and A be the subset of B where all the definitions are made. We can say that atoms appearing in the same formula of A are related. We can say that two formulas φ and ψ are related if they share an atom or if there are atoms p in φ and q in ψ such that p and q are related. Formally:

6.3.1. DEFINITION. *Let A be a set of formulas. We define a relation R_A on atoms by:*

$R_A(p, q)$ *if and only if either $p = q$ or p and q occur in the same formula of A.*

The relation R_A can be extended to a relation on the language:

6.3.2. DEFINITION. *Let $\mathcal{R}_A$ be a relation on formulas defined as:*

$\mathcal{R}_A(\varphi, \psi)$ *iff there are atoms p in φ and q in ψ such that $R_A(p, q)$.*

It is easy to see that $\mathcal{R}_A$ is symmetric and reflexive and that for every φ, $\mathcal{R}_A(\varphi, \neg\varphi)$.

The transitive closure of the relation $\mathcal{R}_A$ gives an equivalence relation that determines a partition of the belief base. Each member of the partition is completely independent from the others with respect to the definitions in A.

We can also use the whole belief base B for deriving the relatedness relation, that is, we can have $A = B$.

Note that if $A = \emptyset$ the relation $\mathcal{R}_A$ is the same as in the syntactic approach, i.e., $\mathcal{R}_\emptyset(\varphi, \psi)$ if and only if φ and ψ share an atom.

6.3.3 The Logical Approach

We can define a relatedness relation $\mathcal{R}$ that captures a notion closer to the logical compartments in Chapter 5 as:

$\mathcal{R}(\varphi, \psi)$ if and only if there is a set $A \subseteq B$ such that $A \cup \{\varphi\} \nvdash \bot$ and either $A \nvdash \psi$ and $A \cup \{\varphi\} \vdash \psi$ or $A \nvdash \neg\psi$ and $A \cup \{\varphi\} \vdash \neg\psi$

or:

$\mathcal{R}(\varphi, \psi)$ if and only if $\varphi \in c(\psi, B)$

where c is a compartmentalization function as in Definition 5.1.4.

Note that this relation is not symmetric. The definitions of path, unrelatedness degree and relevant set of formulas (Δ^ω) in section 6.2 can be maintained. The obvious problem of this approach is that it may be very hard to find the related pairs, but if the relation is given, this problem is avoided. If the language is finite, one way in which the relation can be given is as follows: All sets of formulas can be listed.[4] For each pair of formulas φ and ψ, one can select all sets where φ occurs, check whether they imply ψ or $\neg\psi$ and if so, check their subsets to see whether they also imply ψ or $\neg\psi$. If not, then $\mathcal{R}(\varphi, \psi)$. This is of course very costly, but if the same language is going to be used several times, one can pre-compute this relation and use it for all belief bases.

6.3.4 The Hybrid Approach

A fourth alternative is to combine the database or the syntactic approach with the logical one. If the database is such that it contains several (small) independent modules, this can be a good strategy. One can first apply the first notion of relatedness and find a partition of the database. Then one only has to consider the partition members related to the input formula and look for the relevant formulas inside them using the logic formulation of compartments given in Chapter 5. It can be shown that the compartment obtained is the same one that would be obtained checking the whole set.

Let B be a belief base and let $\mathcal{R}$ be the relatedness relation introduced in Figure 6.1, i.e.: $\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share an atom. Since $\mathcal{R}$ is reflexive and symmetric, its transitive closure defines a partition of the elements of B. Let $\{B_1, B_2, ..., B_n\}$ be the elements of the partition. We say that a set B_i is related to a formula α (represented by $\mathcal{R}(\alpha, B_i)$) if and only if $\mathcal{R}(\alpha, \beta)$ for some $\beta \in B_i$. We denote by P_α the set of elements of the partition that are related to α, i.e., $P_\alpha = \{B_i | \mathcal{R}(\alpha, B_i)\}$.

6.3.3. THEOREM. *Let B be a belief base, α be a contingent formula and c be the compartmentalization function as in Definition 5.1.4.Then $c(\alpha, B) = c(\alpha, \bigcup P_\alpha)$.*

Proof: From $\bigcup P_\alpha \subseteq B$ and Observation 5.1.5 it follows that $c(\alpha, \bigcup P_\alpha) \subseteq c(\alpha, B)$. For the other side of the inclusion, let $\beta \in c(\alpha, B)$. This means that there is $X \subseteq B$ such that $\beta \in X$, X is consistent, inclusion minimal and $\alpha \in Cn(X)$ or $\neg\alpha \in Cn(X)$. Suppose by contradiction that $X \not\subseteq \bigcup P_\alpha$. Then, there is $\gamma \in X$ such that for all $B_i \in P_\alpha$, $\gamma \notin B_i$. From this it follows that α is not related to γ, i.e., there is no $\mathcal{R}$ path from α to γ. Since X is consistent it follows that $\alpha \in Cn(X \setminus \{\gamma\})$ or $\neg\alpha \in Cn(X \setminus \{\gamma\})$, contradicting the minimality of

[4] Actually, one only has to consider one representative of each equivalence class.

X. Hence, $\beta \in X \subseteq \bigcup P_\alpha$ and from part 2 of Observation 5.1.5 it follows that $c(\alpha, B) \subseteq c(\alpha, \bigcup P_\alpha)$. □

Example 16: Let $B = \{p, q, p \to q, s, t \vee v, s \to v, w \wedge x, x \to z, z \vee w\}$ and let c be a compartmentalization function as in Definition 5.1.4. The compartment around $p \vee w$ is given by $c(p \vee w, B) = \{p, w \wedge x\}$. The set B can be clearly divided into three unrelated components: $B_1 = \{p, q, p \to q\}$, $B_2 = \{s, t \vee v, s \to v\}$ and $B_3 = \{w \wedge x, x \to z, z \vee w\}$. The formula $p \vee w$ is only related to two of these components, namely B_1 and B_3. We can then disregard B_2 and calculate the compartment by: $c(p \vee w, B) = c(p \vee w, B_1 \cup B_3) = \{p, w \wedge x\}$.

6.4 Computational Aspects

In this section we discuss the complexity of the operation described for retrieving the set of relevant beliefs from a base.

Let us first briefly discuss the notion of an *anytime algorithm*. An anytime algorithm is one that whenever it is interrupted, it has built an approximate solution for a problem, and the longer it runs, the better the approximation gets. There are two kinds of anytime algorithms [RW91]. Interruptible anytime algorithms can be interrupted unexpectedly and they still return an approximation with accurateness which is a function of the time they run. Contract anytime algorithms must be given the available amount of time in advance. If they are interrupted before the expected time, they may not yield any interesting result.

The good thing about the method for retrieving the relevant beliefs is that it is an interruptible anytime method, that is, whenever it is interrupted, it has retrieved the most relevant beliefs, and the longer it runs, the closer it gets to retrieving all the relevant beliefs (the maximal connected subgraph). This is a very desirable property for modeling agents that may not have enough time or memory to find all the related beliefs. In the ideal case, if there is no resource limitation, the method succeeds in retrieving a maximal connected subgraph.

Below we present a sketch of an algorithm that takes as input a formula α and a belief base and returns the set of formulas of the base that are relevant for α. The algorithm can be stopped at any time, always returning the set of most relevant beliefs for α. The algorithm is a modification of the algorithm BFS for breadth first search in [CLR90].

The belief base is represented by a vector of formulas, each one with a list of pointers to the adjacent nodes in the graph. The nodes adjacent to a formula α are given by Adjacent(α). The complexity of the construction of the list of adjacent nodes depends on the relatedness relation used. For the relation used in Figure 6.1, where two formulas are related if and only if they share an atom, we can use extra, invisible nodes corresponding to the atoms of the language. Every formula added to the belief base is linked to all atoms appearing in it. Constructing the list has complexity $O(m \times n)$, where m is the number of occurrences of atoms in the formulas involved (the "size" of the belief base) and n is the number of atoms in the language. If the atoms are organized in some kind of lexicographical order, this complexity becomes $O(m \times \log n)$.

Retrieve(α,B,Relevant):
1. If $\alpha \in B$, then mark(α)
2. $\Delta^1(\alpha, B)$:= Adjacent(α)
3. $i := 1$; stop := false
4. While not stop do
   - 4.1. For all $\beta \in \Delta^i(\alpha, B)$, mark($\beta$)
   - 4.2. i := i+1; $\Delta^i(\alpha, B) = \emptyset$
   - 4.3 For all $\beta \in \Delta^{i-1}(\alpha, B)$,
     $\Delta^i(\alpha, B) := \Delta^i(\alpha, B) \cup \{\varphi \in$Adjacent($\beta$) s.t. not marked($\varphi$)$\}$
   - 4.4 If $\Delta^i(\alpha, B) = \emptyset$, then stop := true
5. Relevant := $\{\beta \in B$ s.t. marked(β)$\}$

At each step, the algorithm looks at the set retrieved at the previous step and gets all the adjacent nodes that have not been visited yet. When all nodes of a connected component have been visited, it halts. Depending on how the information is encoded, the algorithm runs in linear time with respect to the number of formulas of the belief base.

After retrieving the relevant beliefs, traditional belief change constructions can be applied to it, provided the set is small enough (one can stop the algorithm once the relevant set gets bigger than a certain limit).

In a large database, or in more realistic agents, the whole belief base may be connected, i.e., it may be impossible to isolate a small connected component. In such cases, one can think about more sophisticated notions of connectedness, such as finding connected "chunks" in the graph and preferring edges internal to a chunk over the others (see, for example, [vD97]).

6.5 Related Work

In [CGP99], Chopra et al. developed a method for belief revision based on restricted inference. Instead of belief bases, they work with *belief sequences*, a set with a linear order. The order is a temporal one, i.e., $\alpha < \beta$ if and only if β is more recent than α.

They use the notion of the language of a formula [Par96] in order to define a relevance relation:

6.5.1. Definition. *[Par96] Let $Var(\beta)$ be the set of propositional variables that occur in β. The language L_α of α is the smallest set of propositional variables* *used to express* α*, i.e.,* $L_\alpha = Var(\beta)$*, where* $\vdash \alpha \leftrightarrow \beta$ *and for every* β' *such that* $\vdash \alpha \leftrightarrow \beta'$*,* $|Var(\beta)| < |Var(\beta')|$*.*

Parikh has shown that for every formula α there is a unique language L_α [Par96]. Intuitively, the language L_α captures the notion of "what α is about", the topic of α.

6.5.2. Definition. *[CGP99] Two formulas* α *and* β *are* **directly relevant** *for each other if* $L_\alpha \cap L_\beta \neq \emptyset$*. Let* σ *be a belief sequence. Two formulas* α *and* β *are* k**-relevant with relation to** σ *if and only if there are formulas* $\varphi_0, \varphi_1, ..., \varphi_k$ *in* σ *such that:*

1. $\varphi_0 = \alpha$ *and* $\varphi_k = \beta$
2. *For* $i = 0, ..., k-1$*,* φ_i *and* φ_{i+1} *are directly relevant.*

The relatedness relation is basically the same as that used in Figure 6.1.

6.5.3. Definition. *[CGP99] Given a belief sequence* $\sigma = (\beta_1, \beta_2, ...\beta_n)$*, where* $i < j$ *iff* $\beta_i < \beta_j$*, and a formula* α*, a new order for* σ *can be defined as follows:*

$\beta_i \prec_\alpha \beta_j$ *if and only if:*

1. *There is some* r *such that* β_i *is* r*-relevant to* α *and* β_j *is not* r*-relevant to* α*, or*
2. β_i *and* β_j *are equally relevant to* α *but* $j < i$*.*

The new relation $\prec_\alpha$ orders the formulas in σ according to their relevance to α. When two formulas are equally relevant to α, the most recent one has priority.

This order is used to define a local inference operator.

6.5.4. Definition. *[CGP99] Let* σ *be a belief sequence and* γ *a formula. Let* $\delta_1, ...\delta_n$ *be the formulas of* σ *ordered according to* $\prec_\gamma$*.*

Let k *be any natural number. The inference operator* C_k *is defined as:*

$\gamma \in C_k(\sigma)$ *if and only if* $\gamma \in Cn(\Gamma(\sigma, k, \gamma))$*, where* $\Gamma(\sigma, k, \gamma)$ *is constructed as follows:*

$$\Gamma^0 = \emptyset$$
$$\Gamma^{i+1} = \begin{cases} \Gamma^i & \text{if } \neg\delta_{i+1} \in Cn(\Gamma^i) \text{ or if } \delta_{i+i} \text{ is not } k\text{-relevant to } \alpha \\ \Gamma^i \cup \{\delta_{i+1}\} & \text{otherwise} \end{cases}$$
$$\Gamma(\sigma, k, \gamma) = \Gamma^n.$$

The revision operation defined in [CGP99] consists of simply concatenating the new formula to the sequence. The new formula is appended at the end of the current belief sequence, receiving maximal priority. Possible conflicts are solved by the inference operator C_k respecting the temporal and the relevance orderings.

A similar idea was studied by Ryan in [Rya92]. But in his work, Ryan considers only the temporal ordering and does not take relevance into consideration.

This revision method is in line with what Rott calls the "vertical perspective" [Rot96]. The belief sequence may contain inconsistencies and it may not bear any special relationship to the actual beliefs of an agent. The beliefs to which the agent is committed are given by the application of a non-classical inference operation to the belief sequence.

Revision of belief sequences manages to avoid several of the drawbacks of AGM revision discussed at the end of Chapter 3. Belief states are represented by finite sequences of formulas, instead of logically closed theories. One can have different inconsistent belief sequences, i.e, the presence of inconsistency does not lead to a trivial belief state in which the agent believes every formula of the language. The construction of a revised sequence is trivial and can be iterated with no extra machinery. Revision of belief sequences is prioritized, in that incoming information is always accepted with the highest priority. But unlike AGM revision, the new piece of information does not lose its high priority immediately after being accepted.

How does the revision of belief sequences compare to our framework for local change of structured bases, defined in this and the preceding chapters?

The first problem one encounters is that it is not clear how one could define operations of belief change for belief sequences other than expansion and revision. An operation of contraction of belief sequences is not at all trivial. Rott proposes the use of "phantom beliefs" which work as constraints on what can be accepted [Rot96]. To contract a sentence α, one adds a phantom $\neg\alpha$ with the highest priority. This phantom blocks the derivation of α, but cannot be used in positive derivations, i.e., $\neg\alpha$ is not inferred from it.

Moreover, the notion of "compartments" and "local inference" (Γ and C_k) given by Definition 6.5.4 cannot be used for local change, since C_k is non-monotonic (this is proven as Proposition 2 in [CGP99]). Why is it so? Because new beliefs added may overwrite old ones, so that if α was implied by a sequence which is concatenated with $\neg\alpha$, then α is not implied by the new sequence. It is not clear which logical properties the revision of belief sequences has.

One alternative would be to forget the condition that $\neg\delta_{i+1} \in Cn(\Gamma^i)$ in the definition of Γ (Definition 6.5.4). This means that the final Γ could be inconsistent. A prioritized consolidation (using recency in the selection mechanism) could be used to get rid of the inconsistency.

The method given in [CGP99] has the advantage that revising a belief sequence consists in simply concatenating the new belief. Moreover, it keeps the whole history of the changes. But answering queries about the beliefs is more difficult. For applications in which the agent being modeled only receives new information and never contracts its beliefs, it may present an adequate formalization.

From the computational point of view, it presents the problem that the relevance ordering has to be recalculated for every new input. Remember that in our framework, a structured belief base is built representing the relationship between all formulas. When a new formula is added, one only needs to add links between the new formula and the ones related to it. But in the case of belief sequences, even if no formula is added, for every query one needs to calculate the whole relevance structure. There is also a test for consistency in each step of the construction of Γ.

Combining the two methods could be very fruitful. If we add to belief sequences links representing relevance such as those in the structured bases, calculating the relevance order for some input becomes much easier. Studying the properties of a combined framework is left for future work.

6.6 Conclusion

In this chapter we have examined a more efficient way of retrieving the relevant beliefs from a belief base, so that the local operations presented in Chapter 5 remain well characterized. This was achieved by means of adding extra structure to belief bases.

It is possible to ignore differences in the syntactical form of formulas, although we do not do it here, by assuming (i) a logical approach: when we say "p occurs in φ" we actually mean "for all ψ such that $\vdash \varphi \leftrightarrow \psi$, p occurs in ψ"[5]; or (ii) a computational approach: the formulas are all represented in some canonical (normal) form, and equivalent formulas have equal representation.

In the next chapter we are going to apply the ideas for local change using structured bases to the domain of model-based diagnosis. We will show how to obtain diagnoses of a model without examining the whole model. We will also show that some algorithms developed for model-based diagnosis can be used to implement belief change.

[5]This corresponds to "$p \in L_\varphi$" in the notation of Definition 6.5.1.

Chapter 7

Local Diagnosis

In this chapter we are going to present an application of the theory developed in the preceding chapters. We will investigate the problem of finding diagnoses in faulty systems.

Diagnosis is a very active area within the artificial intelligence community. The problem of diagnosis consists in, given an observation of an abnormal behavior, finding the components of the system that may have caused the abnormality [Rei87].

In the area known as model-based diagnosis [HCdK92], a model of the device to be diagnosed is given in some formal language. In this chapter, we will concentrate on model-based diagnosis methods that work by trying to restore the consistency of the system description and the observations.

We will show how a diagnosis problem can be translated into an operation of kernel semi-revision. Kernel semi-revision consists in adding new information to a database and restoring consistency if necessary. To restore consistency, the expanded database is contracted by $\perp$. We will use the operation of kernel consolidation, introduced in [Han97a] and presented in Section 5.2.2.

Then we will show how to use information about the structure of the device being examined in order to obtain more efficient methods of diagnosis. For this, we will use the operation of local kernel semi-revision, presented in Chapter 5, that considers only the relevant part of the database. In Chapter 6, we have presented a simple method for extracting the relevant part of a structured database, which will be used in this chapter.

Beyond just reducing the diagnosis problem to a problem of belief revision, the present chapter aims at opening a cross-fertilization process between two communities. Researchers working on belief revision rely on very elegant and precise logical formalisms, but are very far from implementing a realistic belief revisioner. On the other hand, researchers working in the field of diagnosis have very powerful tools to prune the computational complexity of the problem, allowing them to deal with real-world situations. But several applications lack a clear formalization. By showing that, at least at a high level, the problems are equivalent, we claim that techniques developed by the model-based diagnosis community could be used for implementing belief revision and that the formal framework presented in Chapters 5 and 6 can be used for focussing on the relevant part of a system for diagnosis.

We will show how the traditional algorithm for consistency-based diagnosis given by Reiter can be used for implementing (local) consolidation.

In [Win95], Winslett suggests the use of belief revision techniques for modeling diagnosis, but without analyzing the similarities between the constructions proposed in both fields. She only shows how a particular problem of diagnosis can be formalized as a belief revision problem.

7.1 Reiter Diagnosis

In this section we introduce the standard method for calculating consistency-based diagnosis, due to Reiter [Rei87]. Although Reiter's framework is based on first-order logic, most of the problems studied in the literature do not make use of full first-order logic and can be easily represented in a propositional language. For the sake of simplicity, we will adapt the definitions given in [Rei87] to only mention formulas in the propositional language L that we have used in the preceding chapters.

7.1.1 Basic Definitions

The systems to be diagnosed will be described by a set of propositional formulas. For each component X of the system, we use a propositional variable of the form okX to indicate whether the component is working as it should. If there is no evidence that the system is not working, we can assume that variables of the form okX are true.

7.1.1. Definition. *A* **system** *is a pair (SD,ASS), where:*

1. *SD, the system description, is a finite set of formulas of L and*

2. *ASS, the set of assumables, is a finite set of propositional variables of the form okX.*

An **observation** is a formula of L. We will sometimes represent a system by (SD,ASS,OBS), where OBS is an observation for the system (SD,ASS).

The need for a diagnosis arises when an abnormal behavior is observed, i.e., when SD$\cup$ASS$\cup$OBS is inconsistent. A diagnosis is a minimal set of assumables that must be negated in order to restore consistency.

7.1.2. Definition. *A* **diagnosis** *for (SD,ASS,OBS) is a minimal set* $\Delta \subseteq$*ASS such that:*

$SD \cup OBS \cup ASS \backslash \Delta \cup \{\neg okX | okX \in \Delta\}$ *is consistent.*

A diagnosis for a system does not always exist:

7.1.3. Proposition. *[Rei87] A diagnosis exists for (SD,ASS,OBS) if and only if SD*$\cup$*OBS is consistent.*

Definition 7.1.2 can be simplified as follows:

7.1.4. Proposition. *[Rei87] The set* $\Delta \subseteq$*ASS is a diagnosis for (SD,ASS,OBS) if and only if* Δ *is a minimal set such that* $SD \cup OBS \cup (ASS \backslash \Delta)$ *is consistent.*

Reiter's definition of diagnosis is closely related to formalizations of non-monotonic reasoning such as circumsription [McC80], which tries to minimize abnormality assumptions, or inferences based on expectations [GM94], which try to maximize the use of normality assumptions. In [Poo89], Poole shows how Reiter diagnoses can be translated into a framework originaly designed to deal with default reasoning.

7.1.2 Computing Diagnoses

In this section we will present Reiter's construction for finding diagnoses. Reiter's method for computing diagnosis makes use of the concepts of *conflict sets* and *hitting sets*. A conflict set is a set of assumables that cannot be all true given the observation:

7.1.5. Definition. *[Rei87] A* **conflict set** *for (SD,ASS,OBS) is a set Conf* $= \{okX_1, okX_2, ..., okX_n\} \subseteq$*ASS such that* $SD \cup OBS \cup Conf$ *is inconsistent.*

From Proposition 7.1.4 and Definition 7.1.5 it follows that $\Delta \subseteq$ASS is a diagnosis for (SD,ASS,OBS) if and only if Δ is a minimal set such that ASS$\backslash\Delta$ is not a conflict set for (SD,ASS,OBS).

A hitting set for a collection of sets is a set that intersects all sets of the collection:

7.1.6. Definition. *[Rei87] Let* $\mathcal{C}$ *be a collection of sets. A* **hitting set** *for* $\mathcal{C}$ *is a set* $H \subseteq \bigcup_{S \in \mathcal{C}} S$ *such that for every* $S \in \mathcal{C}$*,* $H \cap S$ *is nonempty. A hitting set for* $\mathcal{C}$ *is minimal if and only if no proper subset of it is a hitting set for* $\mathcal{C}$*.*

The following theorem presents a constructive approach for finding diagnoses:

7.1.7. Theorem. *[Rei87]* $\Delta \subseteq$*ASS is a diagnosis for (SD,ASS,OBS) if and only if* Δ *is a minimal hitting set for the collection of minimal conflict sets for (SD,ASS,OBS).*

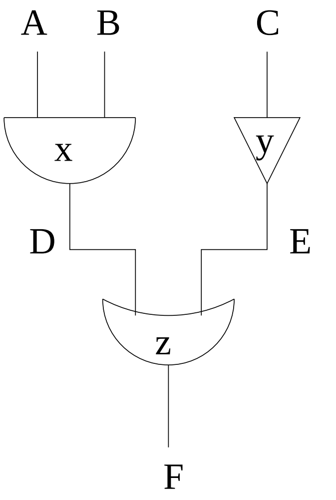

Figure 7.1: Circuit

Example 17: Consider the circuit in Figure 7.1. The system description of this circuit is given by (SD,ASS), where:

$$\text{ASS} = \{okX, okY, okZ\}$$

$$\begin{aligned} \text{SD} = \ & \{(A \wedge B) \wedge okX \rightarrow D, \neg(A \wedge B) \wedge okX \rightarrow \neg D, \\ & C \wedge okY \rightarrow \neg E, \neg C \wedge okY \rightarrow E, \\ & (D \vee E) \wedge okZ \rightarrow F, \neg(D \vee E) \wedge okZ \rightarrow \neg F\} \end{aligned}$$

Suppose we have OBS=$\neg C \wedge \neg F$. This observation is inconsistent with SD $\cup$ ASS. There is only one minimal conflict set for (SD,ASS,OBS): $\{okY, okZ\}$. There are three possible hitting sets: $\{okY\}$,$\{okZ\}$, and $\{okY, okZ\}$. Reiter considers only minimal hitting sets as diagnoses, that is, either Y or Z is not working well.

7.2 Diagnosis via Kernel Semi-Revision

The definitions of the last section bear a striking resemblance to those of the operation of kernel semi-revision presented in Section 3.4.2.

Recall that kernel operations are based on two concepts: kernels and incision functions. The kernels are the minimal subsets of a belief base implying some sentence, while the incision functions are used to decide which elements of the kernels should be given up. Let (SD,ASS,OBS) be a system. The belief base that we are going to semi-revise corresponds to SD$\cup$ASS and the input sentence is OBS. The conflict sets are the assumables in the inconsistent kernels of SD∪ASS∪{OBS}. So, if B=SD∪ASS, the conflict sets are given by $\{X \cap \text{ASS} | X \in (B+\text{OBS}) \perp\!\!\!\perp \bot\}$. Incision functions correspond loosely to hitting sets, the minimal hitting sets being the values of minimal incisions that return only assumables. Note that there is a difference in the status of formulas in SD and those in ASS: formulas in ASS represent expectations and are more easily retracted than those in SD (cf. Definition 7.2.1).

We can model the diagnosis problem as a kernel semi-revision by the observation. Semi-revision can be divided in two steps. First the observation is added to the system description together with the assumables. In case the observation is consistent with the system description together with the assumables, no formula has to be given up. Otherwise, we take the inconsistent kernels and use an incision function to choose which elements of the kernels should be given up.

In the case of diagnosis, we do not wish to give up sentences belonging to the system description or the observation. We prefer to give up the formulas of the form okX, where X is a component of the system. Moreover, we are interested in minimal diagnosis, so the incision should be minimal. For this, we use a special variant of incision function. We modify Definition 3.4.9 so that incisions are minimal and elements of a given set A are prefered over the others:

7.2.1. Definition. *Given a set A, an A-**minimal incision function** is any function σ_A from sets of sets of formulas into sets of formulas such that for any set S of sets of formulas:*

1. $\sigma_A(S) \subseteq \bigcup S$,
2. *If* $\emptyset \neq X \in S$, *then* $X \cap \sigma_A(S) \neq \emptyset$,
3. *If for all* $X \in S$, $X \cap A \neq \emptyset$, *then* $\sigma_A(S) \subseteq A$, *and*
4. $\sigma_A(S)$ *is a minimal set satisfying 1,2, and 3.*

If we take A to be the set of assumables, we obtain an incision function that prefers to select formulas of the form okX over the others.

We can show that for (SD,ASS,OBS), whenever a diagnosis exists, an ASS-minimal incision function will select only elements of ASS:

7.2.2. Proposition. *Let (SD,ASS,OBS) be a system with an observation and σ_{ASS} an ASS-minimal incision function. If a diagnosis exists, then $\sigma_{ASS}((SD \cup ASS \cup OBS) \perp\!\!\!\perp \bot) \subseteq ASS$.*

Proof: A diagnosis exists if and only if SD is consistent with OBS (Proposition 7.1.3). Hence, every inconsistent kernel of SD∪ASS∪OBS must contain an element of ASS. From Definition 7.2.1, it follows that $\sigma_{ASS}((\text{SD} \cup \text{ASS} \cup \text{OBS}) \perp\!\!\!\perp \bot) \subseteq$ASS. □

7.2.3. LEMMA. *The assumables that occur in an inconsistent kernel of the set SD$\cup$ASS$\cup$OBS form a conflict set for (SD,ASS,OBS) and all minimal conflict sets can be obtained in this way, i.e.:*

(i) For every $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$, $X\cap$ASS is a conflict set, and

(ii) For every minimal conflict set Y, there is some $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$ such that $X\cap$ASS$= Y$.

Proof: (i) Let $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$. Then, $X \subseteq (X\cap$ASS)$\cup$SD$\cup$OBS. Since X is inconsistent, so is $(X\cap$ASS)$\cup$SD$\cup$OBS, hence $X\cap$ASS is a conflict set.

(ii) Let Y be a minimal conflict set. Then $Y\cup$SD$\cup$OBS is inconsistent and since $Y \in$ ASS, there is some $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$ such that $X\cap$ASS$\subseteq Y$. Suppose by contradiction that there is some formula α such that $\alpha \in Y$ but $\alpha \notin X\cap$ASS. Since $X\cap$ASS is a conflict set for (SD,ASS,OBS), this contradicts the minimality of Y. Hence, $X\cap$ASS$= Y$. □

Note that not every inconsistent kernel determines a minimal conflict set, since for conflict sets only the elements of ASS matter, i.e., there may be two inconsistent kernels X_1 and X_2 such that $X_1\cap$ASS is a proper subset of $X_2\cap$ASS.

Recall that given an incision function σ, the semi-revision of a set B by a formula α was given by $B?_\sigma\alpha = (B + \alpha) \setminus \sigma((B + \alpha) \perp\!\!\!\perp \bot)$. A diagnosis is given by the elements of ASS that are given up in a kernel semi-revision by the observation.

7.2.4. PROPOSITION. *Let S=(SD,ASS,OBS) be a system and σ_{ASS} an ASS-minimal incison function.*

(SD $\cup$ASS)$\setminus$((SD $\cup$ASS)$?_{\sigma_{ASS}}$OBS) $= \sigma_{ASS}$((SD $\cup$ASS $\cup$OBS) $\perp\!\!\!\perp \bot$) is a diagnosis.

Proof: We have to prove that given a system for which there is a diagnosis and an observation, it holds that:

1. If d is a diagnosis according to Definition 7.1.2, then there is an ASS-minimal incision function σ_{ASS} such that $d= \sigma_{ASS}$((SD $\cup$ASS $\cup$OBS) $\perp\!\!\!\perp \bot$).

2. If σ_{ASS} is an ASS-minimal incision function, then
   σ_{ASS}((SD $\cup$ASS $\cup$OBS) $\perp\!\!\!\perp \bot$)
   is a diagnosis according to Definition 7.1.2.

1. Let σ_{ASS}((SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$)$= d$. We have to show that σ_{ASS} is an ASS-minimal incision function for the relevant domain, i.e., we must show that it satisfies the four conditions of Definition 7.2.1.

(i) $d \subseteq \bigcup$((SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$): If d is a diagnosis according to Definition 7.1.2, then d is a minimal hitting set for the set of all minimal conflicts of (SD,ASS,OBS). From part (ii) of Lemma 7.2.3 we know that for every minimal conflict set Y, there is $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$ such that $X\cap$ASS$= Y$.

(ii) If $\emptyset \neq X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$, then $X \cap d \neq \emptyset$: From part (i) of Lemma 7.2.3, we know that $X\cap$ASS is a conflict set. Since d is a hitting set, $X \cap d \neq \emptyset$.

(iii) If $d \not\subseteq$ASS, then for some $X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot$, $X\cap$ASS$= \emptyset$: Since d is a diagnosis, $d \subseteq$ASS and the condition is trivialy satisfied.

(iv) d is a minimal subset satisfying (i),(ii),(iii): Since d is a diagnosis according to Definition 7.1.2, d is a minimal hitting set.

2. From part (ii) of Lemma 7.2.3, we have that all minimal conflict sets are elements of the set $\{X\cap$ASS$|X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot\}$. We have to show that an ASS-minimal incision function for the inconsistent kernels determines a minimal hitting set for all minimal conflicts. From Definition 7.2.1 and Proposition 7.2.2 it follows that $\sigma_{ASS}($(SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot)$ is a hitting set for the set of minimal conflicts of (SD,ASS,OBS). That it is also a minimal hitting set follows directly from Definition 7.2.1. (Since the non-minimal conflict sets contained in $\{X\cap$ASS$|X \in$ (SD$\cup$ASS$\cup$OBS) $\perp\!\!\!\perp \bot\}$ are supersets of some minimal conflict set and all minimal conflict sets are considered, an ASS-minimal incision function will give the same result as if only minimal conflict sets were considered). $\Box$

Going back to the circuit in Figure 7.1, we see that SD $\cup$ ASS$\cup$ OBS is inconsistent. This means that SD $\cup$ ASS$\cup$ OBS has to be consolidated. There is only one inconsistent kernel:

$$(\text{SD} \cup \text{ASS}\cup\text{OBS}) \perp\!\!\!\perp \bot = \{\{\neg C \wedge okY \rightarrow E, (D \vee E) \wedge okZ \rightarrow F, okY, okZ, \neg C \wedge \neg F\}\}$$

We have two possibilities for ASS-minimal incision functions:
$\sigma_1 = \{okY\}$ and $\sigma_2 = \{okZ\}$
This means that either Y or Z are not working well.

7.3 Using System Structure

Suppose that instead of the circuit depicted in Figure 7.1, we have the circuit in Figure 7.2. Suppose also that we get the same observation, i.e., OBS$= \neg C \wedge \neg F$. Intuitively, only a small part of the circuit (roughly the sub-circuit at figure 7.1) has to be considered in order to arrive to a diagnosis.

In Chapter 5, we have extended the definition of kernel semi-revision to an operation that considers only the relevant part of a database, local kernel semi-revision. In Chapter 6, we have shown how to use structure present in a database in order to find compartments and implement local kernel contraction more efficiently. The key idea of the method described is to use a relation of relatedness between formulas of the belief base. In some applications, as we will see, such a relation is given with the problem. In the case of the circuit shown in Figure 7.2, there is a very natural dependence relation. The output of each of the components depends on the input and on whether the component is working well.

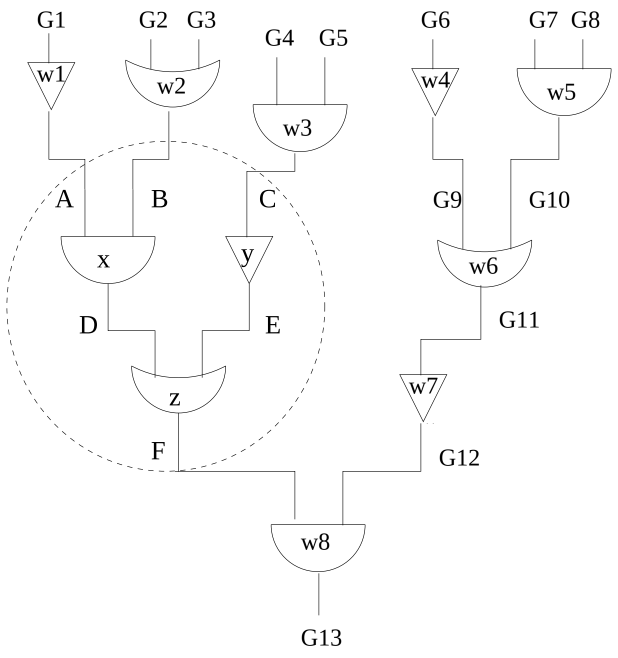

Figure 7.2: Larger circuit

The only observation we have is $\neg C \wedge \neg F$. Since this observation is inconsistent with the system description together with the assumption that all components are working well, there must be some faulty component. Moreover, the faulty component must be in the path between C and F (of course, there may be other faulty components, but we are only searching for the abnormality that explains the observation). We only need to consider the descriptions of components y and z in order to find the diagnosis.

In the next section we will show how to use the framework of Chapter 6 in order to find diagnoses without having to check the entire system description for consistency.

7.4 Local Kernel Diagnosis

As we have seen, diagnosis problems fit very well in the framework for local change that we proposed in the preceding chapters. Besides the fact that the traditional method for finding diagnosis based on the notion of consistency is almost identical to the construction of kernel semi-revision, in most diagnosis problems there is a very natural notion of relatedness that can be used to structure the belief base so that the search for diagnoses becomes more efficient.

In this section we formalize the example in Figure 7.2 in order to show how to derive a concrete relatedness relation from the given database.

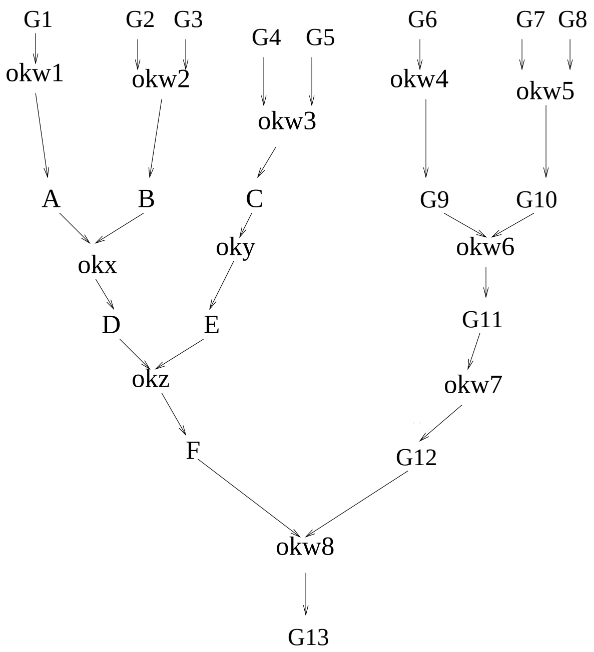

Figure 7.3: Relatedness relation between atoms

We will use a relatedness relation between atoms, as illustrated in Figure 7.3. The relation is not symmetric. We can easily adapt the definitions presented in Chapter 6 to deal with a directed graph.

The basic algorithm is as follows: we start from the propositional variables that occur in the observation and spread the activation in the graph, following the direction of the arcs. The spreading finishes either when the end of the paths are reached or when we run out of resources (time or memory). This is done by the algorithm `Retrieve` below, an adaptation of the algorithm given in Section 6.4.

`Retrieve`(OBS,ASS,Relevant):
1. For all $p \in$ Var(OBS), mark(p)
2. Δ^1(OBS) := Adjacent(Var(OBS))
3. Relevant := Var(OBS)$\cap$ASS
4. i := 1; stop := false
5. While not stop do

5.1. For all $p \in \Delta^i$(OBS), mark(p)
If $p \in$ASS, then Relevant := Relevant $\cup\{p\}$
5.2. i := i+1; Δ^i(OBS)=$\emptyset$
5.3 For all $p \in \Delta^{i-1}$(OBS),
Δ^i(OBS) := Δ^i(OBS) $\cup\{q \in$Adjacent(p) s.t. not marked(q)$\}$
5.4 If Δ^i(OBS)= $\emptyset$, then stop := true

After we have retrieved the relevant assumables, the relevant compartment is taken to be the observation together with all formulas in SD$\cup$ASS which mention the relevant assumables.

`Compartment`(OBS,SD,ASS,Comp):
1. Retrieve(OBS,ASS,Relevant)
2. Comp=OBS
3. For all $p \in$Relevant, Comp:= Comp$\cup\{\alpha \in$SD$\cup$ASS$|p \in$Var$(\alpha)\}$.

As we have seen in Section 6.4, the algorithm for `Retrieve` is an anytime algorithm. The algorithm for `Compartment` is not, at least in principle. But if one keeps the order in which the relevant atoms are retrieved and uses them in this order in line 3 of algorithm `Compartment`, one can be sure that the description of the most relevant components will be retrieved first.

For the circuit in Figure 7.2, we have:

SD=$\{(A \wedge B) \wedge okX \rightarrow D$, $\neg(A \wedge B) \wedge okX \rightarrow \neg D$,
$C \wedge okY \rightarrow \neg E$, $\neg C \wedge okY \rightarrow E$,
$(D \vee E) \wedge okZ \rightarrow F$, $\neg(D \vee E) \wedge okZ \rightarrow \neg F$,
$G1 \wedge okW1 \rightarrow \neg A$, $\neg G1 \wedge okW1 \rightarrow A$,
$(G2 \vee G3) \wedge okW2 \rightarrow B$, $\neg(G2 \vee G3) \wedge okW2 \rightarrow \neg B$,
$(G4 \wedge G5) \wedge okW3 \rightarrow C$, $\neg(G4 \wedge G5) \wedge okW3 \rightarrow \neg C$,
$G6 \wedge okW4 \rightarrow \neg G9$, $\neg G6 \wedge okW4 \rightarrow G9$,
$(G7 \wedge G8) \wedge okW5 \rightarrow G10$, $\neg(G7 \wedge G8) \wedge okW5 \rightarrow \neg G10$,
$(G9 \vee G10) \wedge okW6 \rightarrow G11$, $\neg(G9 \vee G10) \wedge okW6 \rightarrow \neg G11$,
$G11 \wedge okW7 \rightarrow G12$, $\neg G11 \wedge okW7 \rightarrow \neg G12$,
$(F \wedge G12) \wedge okW8 \rightarrow G13$, $\neg(F \wedge G12) \wedge okW8 \rightarrow \neg G13\}$

ASS $=\{okX, okY, okZ, okW1, okW2, okW3, okW4, okW5, okW6,$
$okW7, okW8\}$

If we apply the algorithm `Retrieve`($\neg C \wedge \neg F$,ASS,Relevant) on the graph depicted in Figure 7.3, we get Relevant=$\{$okY,okZ,okW8$\}$. For `Compartment`(OBS, SD, ASS, Comp) we get Comp=$\{\neg C \wedge \neg F, C \wedge okY \rightarrow \neg E, \neg C \wedge okY \rightarrow E, (D \vee E) \wedge okZ \rightarrow F, \neg(D \vee E) \wedge okZ \rightarrow \neg F, (F \wedge G12) \wedge okW8 \rightarrow G13, \neg(F \wedge G12) \wedge okW8 \rightarrow \neg G13, okY, okZ, okW8\}$.

The diagnosis can be searched using only the formulas in Comp. Note that the component w8 was not really relevant for the diagnosis but, nevertheless, we have reduced the set to be semi-revised.

This is a very general method for focusing on a small part of the system description. One can add to it some domain specific heuristics to improve its efficiency. The system IDEA [SC97], used by FIAT repair centers works on dependence graphs that show graphically the relation between the several components of a device.

7.5 Computing Kernel Operations

In this section we will show how Reiter's algorithm for computing diagnosis can be adapted and used for belief revision.

7.5.1 Reiter's Algorithm

The algorithm given in [Rei87] computes all minimal hitting sets for an arbitrary collection of sets. We will use it later for finding the incision functions used in kernel constructions. We present here the version corrected by [GSW89].

The algorithm generates a directed acyclic graph (DAG) with nodes labeled by sets and arcs labeled by elements of the set. The idea is that for each node labeled by a set S, the arcs leaving from it are labeled by the elements of S. Let $H(n)$ denote the set formed by the labels of the path going from the root to node n. Node n has to be labeled by a set S such that $S \cap H(n) = \emptyset$. If no such set can be found, the node is labeled by @. The idea is that every path finishing at a node labeled by @ is a hitting set, since it intersects all possible labels for the nodes.

The algorithm tries to generate as few new node labels as possible. This is due to the fact that for diagnosis (and for belief revision as well), the collection of sets F which can be used as node labels will be given only implicitly. Calculating one element of F involves a call to a theorem prover to find a conflict set (in the case of diagnosis; a kernel in the case of belief revision) and is therefore a very expensive operation.

The algorithm minimizes the number of calls to the theorem prover by pruning the graph while it is being built. When a new node has to be labeled, the algorithm tries to re-use existing labels first. If a node label S is a superset of another label S', then it can be "closed", it does not have to be considered any longer, since any hitting set for F will be a hitting set for $F \setminus \{S\}$.

Let F be a family of sets.

1. Choose one set to label the root node (level 0).
2. For each node n at level i do:

2.a. If n is labeled by a set S, then for every $s \in S$ create an arc departing from n with label s.

2.b. Set $H(n)$ to be the set of arc labels on the path from the root to node n.

2.c. If there is some node n' such that $H(n') = H(n) \cup \{s\}$, then let the s-arc of n point to n'.

2.d. Else, if there is a node n' labeled by @ such that $H(n') \subset (H(n) \cup \{s\})$ then close the s-arc (i.e., do not compute a label or successors for this node).

2.e. Else, if there is some node n' labeled by S' such that $S' \cap (H(n) \cup \{s\}) = \emptyset$, then let the s-arc of n point to a new node labeled by S'.

2.f. Otherwise, let the s-arc point to a new node m and let m be labeled by the first element S' of F such that $S' \cap H(m) = \emptyset$. If no such set exists, then label m by @.

2.g. If there is some node n' labeled by a set S_1 such that $S' \subset S_1$, then relabel node n' by S' and remove all arcs departing from n' which were labeled by elements of $S' \setminus S_1$.

3. Repeat step 2 for level $i + 1$.

The algorithm expands the graph breadth first. Each level is processed by step 2. Steps 2.c and 2.e re-use nodes or labels if possible. Reiter has proven the following theorem:

7.5.1. THEOREM. *[Rei87] Let F be a collection of sets and let D be a graph returned by the algorithm above. The set $\{H(n)|n$ is a node of D labeled by @$\}$ is the collection of minimal hitting sets for F.*

The final algorithm for calculating all diagnoses constructs a DAG as above, except that when it is supposed to generate a new node label, it does so by calling the theorem prover with a smaller set. Let TP be a function such that TP(SD,ASS,OBS) returns a conflict set for (SD,ASS,OBS), i.e, a subset S of ASS such that SD∪S∪OBS is inconsistent. If no conflict set exists, the function returns @. When one needs to compute a label for a node n, label n by TP(SD,ASS\H(n),OBS).

Consider the following example [GSW89]:

> **Example 18:** Let $F = \{\{a, b\}, \{b, c\}, \{a, c\}, \{b, d\}, \{b\}\}$. Figure 7.4 shows part of the graph built by the algorithm. The set $\{a, b\}$ is chosen to label node 0 and two arcs are created with labels a and b. Node 1 is labeled by $\{b, c\}$ and node 2 by $\{a, c\}$ and arcs are created leaving from node 1 labeled by b and c and leaving from node 2 labeled by a and c. Node 3 receives the label @, since there is no set $S \in F$ such that $S \cap H(3) = S \cap \{a, b\} = \emptyset$. Node 4 is labeled by $\{b, d\}$. The arc leaving from node 2 and labeled by a points to node 3, since $H(3) = H(2) \cup \{a\}$ (step 2.c of the algorithm). Node 5 is labeled by @, since there is no set $S \in F$ such that $S \cap H(5) = S \cap \{b, c\} = \emptyset$. Node 6 is closed (step 2.d of the algorithm), since nodes 3 and 5 are labeled by @ and $H(3) \subset H(6)$ and $H(5) \subset H(6)$. When node 7 is labeled by $\{b\} \subset \{a, b\}$, the graph is pruned and the root node 0 is

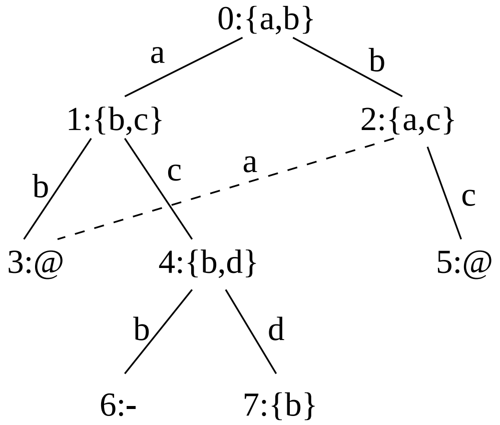

Figure 7.4: Reiter's algorithm – 1

relabeled by $\{b\}$ (step 2.g). The resulting graph is shown in Figure 7.5. The hitting sets are $\{a, b\}$ and $\{b, c\}$.

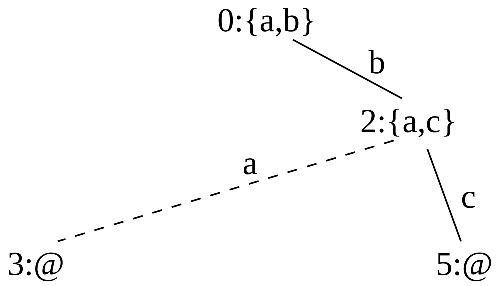

Figure 7.5: Reiter's algorithm – 2

7.5.2 An Algorithm for Kernel Semi-Revision

In order to apply the algorithm for kernel operations, one needs to adapt very few things. Usually, we will not have access to the whole collection of inconsistent kernels. Using a theorem prover in order to find an inconsistent subset of a belief base does not guarantee that the set returned is a minimal one. Nevertheless, even if the set is not minimal, the algorithm returns the collection of values for the minimal incision functions for all inconsistent subsets of the base. In particular, the returned values are values for incision functions for the inconsistent kernels.

Let TP be a function such that TP(B) returns an inconsistent subset of B. We then build a directed acyclic graph using Reiter's algorithm. Whenever a new label for a node n has to be generated, we call TP(B\H(n)).

That the algorithm does what it is expected to do follows directly from the correctness of Reiter's algorithm.

> **Example 19:** Consider the belief base $B = \{\neg a, \neg b, a \vee b, q, q \rightarrow p, \neg p\}$. There are only two inconsistent kernels, $\{\neg a, \neg b, a \vee b\}$ and $\{q, q \rightarrow p, \neg p\}$. However, the theorem prover may find some superset of these sets. Suppose it finds the collection $\{\{\neg a, \neg b, a \vee b, q\}, \{\neg b, a \vee b, q, q \rightarrow p, \neg p\}, \{\neg a, \neg b, a \vee b, q, \neg p\}, \{\neg a, q, q \rightarrow p, \neg p\}, \{\neg a, \neg b, a \vee b\}, \{q, q \rightarrow p, \neg p\}\}$. The algorithm computes the graph shown in Figure 7.6. From part (1) to part (2) in the figure we see that the root node is relabeled by $\{\neg a, \neg b, a \vee b\}$. From part (2) to part (3), since the label of node 5, $\{q, q \rightarrow p, \neg p\}$, is a proper subset of the labels of nodes 1, 2 and 3, these nodes are relabeled.
>
> The values for the incision functions are: $\{\neg a, q \rightarrow p\}$, $\{\neg a, q\}$, $\{\neg a, \neg p\}$, $\{\neg b, q \rightarrow p\}$, $\{\neg b, q\}$, $\{\neg b, \neg p\}$, $\{a \vee b, q \rightarrow p\}$, $\{a \vee b, \neg p\}$, and $\{a \vee b, \neg p\}$.

7.5.3 Applying the Algorithm to Local Operations

The algorithm for kernel operations can be easily combined with the algorithm for finding the relevant compartment presented in Chapter 6. Consider once more the belief base $B = \{\neg a, \neg b, a \vee b, q, q \rightarrow p, \neg p\}$. Suppose we are only interested in solving inconsistencies related to p. If we assume the relatedness relation given by:

$\mathcal{R}(\alpha, \beta)$ if and only if α and β share an atom,

the algorithm Retrieve(p,B,Relevant) given in Section 6.4 returns Relevant=$\{q, q \rightarrow p, \neg p\}$. We can the apply Reiter's algorithm in order to find the possible values for incision functions. It is not difficult to see that the algorithm returns $\{q\}$, $\{q \rightarrow p\}$, and $\{\neg p\}$.

7.6 Related Work

7.6.1 Approximate Diagnosis

In [tTvH96], ten Teije and van Harmelen propose the use of approximations for calculating diagnosis in an efficient way. They use the approximate entailments proposed by Schaerf and Cadoli in [SC95]. Approximate entailments have as a parameter a set S of propositional letters, which are the "relevant" letters. By

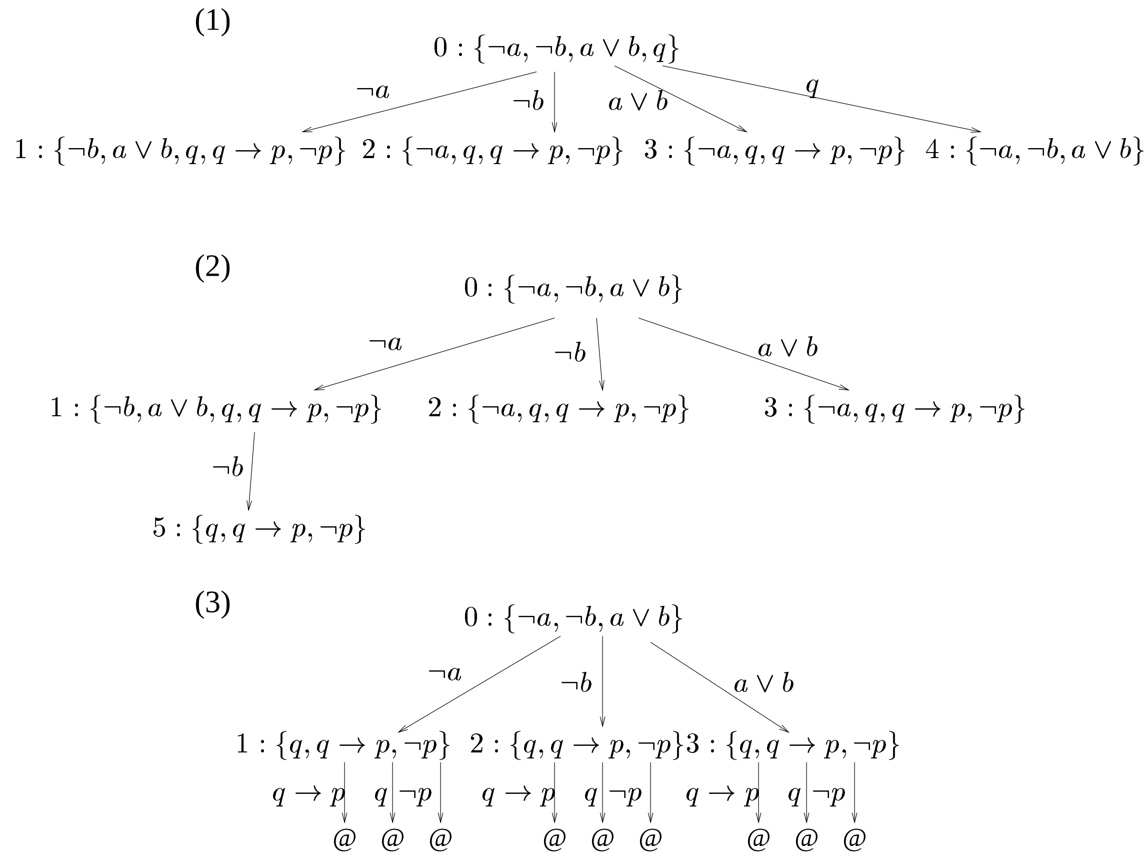

Figure 7.6: Finding incision functions

augmenting the set S, one gets closer to classical entailment. Schaerf and Cadoli define two kinds of approximate entailment, one which is unsound and complete ($\models_1^S$) and one which is sound but incomplete ($\models_3^S$).

Approximate entailments are defined using different sorts of truth assignments. A *1-S-assignment* makes for all $p \notin S$ both p and $\neg p$ false, while a *3-S-assignment* makes at least one of the literals p and $\neg p$ true. For propositional letters in S, both truth assignments behave classically. A set A 1-S-entails a formula α if and only if every 1-S-assignment that satisfies A also satisfies α. The definition of 3-S-entailment is analogous.

Schaerf and Cadoli have shown that when S is augmented, the accuracy of $\models_3^S$ as an approximation of classical entailment improves and the same holds for $\not\models_1^S$ as an approximation of $\not\models$. They also present an incremental algorithm for calculating $\models_1^S$ and $\models_3^S$ and show that in the worst case, i.e., when one needs to increase S until the approximate entailment becomes classical, the complexity is not higher than the complexity of classical entailment. The algorithm can be interrupted as soon as one has a satisfactory approximation.

Ten Teije and van Harmelen propose the use of approximate entailment for diagnosis instead of classical entailment. Varying the set S, one can reduce or increase the number of diagnoses. The idea is very similar to ours: consider only the relevant part of the system and augment the part considered until a satisfactory result is obtained. The main problem, as they point, is to choose a good initial set S. In [tTvH96] and [tTvH97], they give some strategies for selecting the set S.

7.6.2 Assumption-Based Truth Maintenance Systems

Assumption-Based Truth Maintenance Systems (ATMS) were developed by De Kleer [dK86] as a module of a problem solver. This module was supposed to maintain a structure of atoms and causal links so that queries could be answered efficiently. An ATMS consists of a background theory (represented as a set of propositional Horn clauses), assumptions and observations, both represented by positive literals.

The main mechanism of an ATMS is the labeling of atoms. Each atom, including $\perp$, is labeled by the set of minimal sets of assumptions which together with the background theory imply that atom. Using the notation for kernels introduced in Definition 5.1.1, we have that the label of an atom a is given by $A \perp\!\!\!\perp_{C_T} a$, where A is the set of assumptions, T is the background theory and $C_T(X) = C(X \cup T)$. The elements of the label of $\perp$ are called *nogoods* and are used to check the consistency of other labels. The nogoods correspond to the conflict sets in Reiter's construction (Definition 7.1.5). It would be interesting to see how the machinery developed for ATMS can be used for implementing kernel operations and in particular, for calculating the kernels.

7.7 Conclusion

In this chapter we have seen an application of the theory developed in the preceding chapters to the problem of model-based diagnosis. We have shown how the problems of belief revision and diagnosis are related and how one can use the structure of a particular diagnosis problem in order to consider only the relevant components of a system.

We have also shown how Reiter's algorithm for diagnosis can be adapted for implementing belief revision operators. The fact that Reiter's algorithm can be used for belief revision bridges the gap between belief revision theory and implemented systems. Reiter's algorithm is used in several systems and we expect that several computational tools developed for diagnosis systems can be adapted for revision operators.

Chapter 8

Discussion and Future Work

In this thesis, we have presented a theory of belief revision for more realistic agents, i.e., for all of us who do not have access to infinite memory, an infinite amount of time to reach conclusions, and perfect logical ability. This includes all sorts of agents except for the idealized beings which have been studied in most approaches to belief revision.

We have started by presenting a formal framework that includes a model for belief states and simple operations for changing belief states. Our model for belief states distinguishes between different sorts of beliefs, according to whether they are explicitly or implicitly believed, whether they are currently active and whether they are fully believed or only provisional. There are in the literature several proposals which present some of these distinctions, but our proposal differs from the existing ones in the following aspects:

- Unlike Harman's [Har86] and Cherniak's [Che86] informal proposals, we present a formal framework based on sets, where the relationship between the different sorts of beliefs is made clear.

- Unlike formal approaches that distinguish between explicit and implicit beliefs such as the ones due to Fagin and Halpern [FH88] and Levesque [Lev84], we do not require the set of explicit beliefs to be consistent nor assume that the implicit beliefs are the classical consequences of the explicit beliefs.

- As opposed to what is common in Belief Revision, we try to look at belief change operations step-by-step, breaking them into very simple operations on belief states.

We have shown that traditional belief revision operations for idealized agents can be simulated in our framework if we allow for infinite sets and infinite sequences of basic operations.

The basic operations that we describe are used as building blocks for constructing complex operators. The main idea is that when an agent receives new information, this information is not immediately fully accepted, but is instead recorded as a provisional belief. The agent must next decide, in the light of his previous beliefs, whether this provisional belief should be accepted or not. Our framework does not say much about how this decision is taken. In Appendix A, we present an example of one way to implement this decision process. We suggest the use of argumentation theory as presented in [Dun95, Lou98]. Arguments for and against a certain provisional belief are compared and used for taking a decision. We use Loui's framework for resource-bounded argumentation [Lou98], in which the agent does not always succeed in taking all arguments into account, but examines as much as his available resources allow him to.

Another important point in our framework is that the set of active beliefs, i.e., the beliefs that are available for reasoning, is very small when compared to the set of explicit beliefs. This accounts for the intuition that agents cannot think of everything they know at the same time. An agent is usually focussing on a certain topic, or subject matter. One of our basic operations takes care of the retrieval of explicit (but not active) beliefs into the set of active beliefs. But how does the agent decide which beliefs are relevant for a certain operation of belief change? In Chapters 5 and 6 we have presented two different solutions to this problem. The first solution, presented in Chapter 5, uses logic — and only logic — to isolate the relevant part of the agent's beliefs. A belief is considered relevant for a given formula if it contributes to proving or disproving the formula. We define local operations of belief change that affect only the relevant set retrieved. This method has several shortcomings: finding the relevant beliefs is computationally as hard as the traditional belief change operators. Moreover, there is no control over the size of the retrieved set, i.e., it may happen that the set of relevant beliefs is the whole set of explicit beliefs. But the method provides us with some interesting formal results. All the local operations that we have defined are axiomatized, and the representation results show exactly what is needed for each operation to maintain its elegant logical properties. Since these axiomatizations do not depend on the particular notion of relevance we have used, we propose in Chapter 6 a computationally efficient method for finding relevant beliefs. For this method we need extra-logical information about the relatedness relation between beliefs. This information allows for the distinction of different degrees of relevance. We show that relatedness relations can very often be derived from the given set of beliefs or from a particular application, such as the one investigated in Chapter 7. Besides the computational advantages, the method described in Chapter 6 is also very intuitive and is in line with research on cognitive models of memory [And80].

Chapter 7 presents an application of the theory developed in this thesis to the area of model-based diagnosis. Diagnoses of circuits are used as concrete examples and make clear what the abstract theoretical notions actually mean.

A relatedness relation in a circuit may be the causal link between the output and input of a certain component. The method presented in Chapter 6 is used to focus the diagnosis on the relevant part of the circuit. By making the link between belief revision and model-based diagnosis, we also claim that some computational tools developed for diagnosis can be used for implementing belief revision. In particular, we show how Reiter's algorithm for consistency-based diagnosis can be used for implementing kernel semi-revision.

To summarize, the main achievements of the present thesis are:

- We have defined more structured belief states and basic operations that apply to them and showed that under the common assumption of logical omniscience, belief states together with the basic operations can simulate the AGM paradigm.

- (Joint work with Sven Ove Hansson) We have generalized several of the representation results found in the literature for more abstract inference operators. For each belief change operation, we have shown which are the properties that an inference relation must have so that the axiomatizations are correct.

- (Joint work with Sven Ove Hansson) We have defined a notion of local inference which considers only the relevant part of a belief base, and we used this notion to define local versions of existing belief change operations.

- We have presented a method for efficiently retrieving the relevant part of a belief base which can be used together with the local versions of the belief revision operations.

- We have presented diagnosis as a practical application for the local operations of belief change.

- We have shown how algorithms developed for model-based diagnosis can be adapted to implement belief revision.

There are several directions in which the work presented here can be expanded. We list below some ideas for future work:

- It would be interesting to study more thoroughly the decision process for accepting provisional beliefs. An idea would be to follow the proposal presented in Appendix A and investigate existing argumentation systems to see whether they can be used together with our framework for resource-bounded revision.

- Since the representation results presented in Chapter 5 are very general, it would be interesting to see whether they can be used together with Schaerf and Cadoli's approximate entailment [SC95], giving an anytime method for approximate belief revision and maintaining the logical axiomatization.

- The framework presented in Chapter 6 can be used with any notion of relevance or relatedness. Applying specific notions of relevance proposed in the literature could make clear whether our framework is in line with the intuitions behind these proposals.

- Reiter's algorithm for diagnosis can be further adapted in order to implement other constructions for belief revision operators. In particular, it would be interesting to check whether the selection mechanism used in partial meet operations can be embedded into the algorithm.

Appendix A

Full Acceptance Via Argumentation

As we have seen, in most approaches to belief revision, incoming information is given the highest priority, so that if a contradiction arises, some of the previous beliefs have to be given up. In approaches to non-prioritized belief revision [Han97b, Han99a], i.e., revision in which the new piece of information does not have the highest priority, the decision whether to accept new information or not is usually taken with the help of extra-logical means such as selection functions or incision functions, but there is no real recipe of how to construct these functions. In this chapter, we explore a different idea — using argumentation theory for deciding whether new information is acceptable.

We will use the simplified model of belief states introduced in Section 4.1 together with the operations defined in Section 4.2. In this chapter, we turn to a question left open by the model in Chapter 4, namely how to decide whether a provisional belief should be accepted. The model that we will present is very simplified. The aim of this chapter is only to show that argumentation theory can be a useful tool to combine with our framework.

According to Dung [Dun95], a formula is believable "if it can be argued successfully against attacking arguments". Dung also says that reasoning about one's own beliefs is like performing an internal argument [Dun95]. Our concept of provisional beliefs is based on Harman's idea of *tentative hypotheses*. In order to be fully accepted, a tentative hypothesis has to survive the best attempts to refute it [Har86]. In our case, "best attempts" are as good as the agent is capable given his limitations.

This is reflected in the framework for resource-bounded argumentation given in [Lou98]. Loui describes a very general framework where there are a number of parties involved, some of which (the players) are allowed to make locutions, the others being advocates. Each of the players tries to get the current opinion to be in his favor by presenting arguments. A vector represents the resources consumed at each move.

A protocol for disputation has to be defined and depends on the application.

These are the real "rules of the game", which determine what is allowed as a move, who is allowed to make the next move, how the moves affect the current opinion, and what the conditions for termination are. In [Lou98], some protocols are presented, which can be chosen according to the intended application.

In the next section we will present the theory of argumentation based on [Lou98]. In section A.2 we present our proposal for using the theory sketched in section A.1 to enrich the framework presented in Chapter 4.

A.1 Argumentation

In this section, we introduce the basic concepts of argumentation theory that we will need in this chapter. This section is based on [Lou98].

Argumentation has been investigated by researchers in the area of philosophy and artificial intelligence. Recently, it became clear that argumentation can be seen as a kind of non-monotonic reasoning. Arguments are not proofs, but some kind of justification for a claim, usually defeasible. An argumentation process usually follows some protocol. Once the parties involved in the disputation agree on the protocol, the outcome of an argumentation process following the protocol is considered fair.

Disputations are highly non-monotonic. The outcome depends on the particular way in which the argumentation process took place and if the process continues, the outcome may change. Nevertheless, the process is fair (provided the disputants agreed about the protocol) and the outcome is warranted.

An argument is usually a pair formed by a set of formulas and one special formula, the claim. The set of formulas serves as a justification for the claim. Arguments are related to each other in several ways. Arguments can interfere with each other, in case their claims (or subclaims) are inconsistent.

Loui [Lou98] defines a very general framework for argumentation that has to be "filled in" in order to model particular kinds of disputation.

An argumentation process is a sequence of locutions, where each locution is a triple formed by one party, the argument and the resources consumed. The participants of the argumentation process do not necessarily have access to the same information. They may also have different shares of resources at their disposal. In our case, we will use argumentation processes where only two parties are involved, **pro** and **con**. A variable *current.opinion* stores the party which is winning the disputation at a certain point. The parties try to switch the current opinion in their favor by advancing locutions. Since we are modeling an internal argumentation process, where a single agent is involved and plays the roles of **pro** and **con**, we can assume that both parties have access to the same information.

A.2 Using Argumentation for Accepting Beliefs

In this section we present our proposal for using argumentation in order to decide whether a provisional belief should be accepted or not. In our case, the argumentation is an internal process where a single agent plays the role of **pro** and **con**, analyzing the arguments for and against a given provisional belief. Since we are dealing with resource-bounded agents, this internal argumentation will not always succeed in examining all reasons for accepting or rejecting a belief. By defining a protocol for this process, we have to take care that the outcome can be considered fair.

There are two ways in which a sentence can become a provisional belief:

1. New information may be acquired by an operation of observation, i.e., come from the outside world. This new piece of information has to be checked before being fully accepted. In this case, **con** tries to argue against it. If he fails, the provisional belief is accepted, since it has survived the best attempts to refute it. If **con** succeeds, the provisional belief is rejected.

2. A sentence that was previously accepted, an explicit belief, may become provisional if the agent gets evidence against it. In this case, inquiry is reopened ([Har86]) and **pro** tries to argue for the sentence. If he fails, the provisional belief is rejected. If **pro** succeeds, the provisional belief becomes fully accepted again.

In the framework presented in Section 4.1, there are two clearly limited resources: the size of the set of active beliefs and the number of basic operations used in the disputation process. Since in our case a single agent is playing the roles of **pro** and **con**, the set of active beliefs is a shared resource, both **pro** and **con** have access to the whole set.

All the sentences in the arguments presented become active. The elements of the set of active beliefs are ordered according to the time in which they were introduced in the argumentation. When the set gets too big, the oldest elements are "forgotten". If the discarded elements were explicit beliefs that were retrieved, they remain in the set of explicit beliefs but become inactive. If they were only provisional beliefs, then they are irremediably forgotten and dismissed from the whole structure.

An argument for us will be a sequence of elements of the set of explicit beliefs which is a derivation for its claim according to a finite (small) number of applications of inference rules known by the agent. An argument *arg* of player p (=**pro** or **con**) is counterargued when the other player presents an argument against one of the elements of *arg* (its subclaims). An argument *arg* of player p is defeated if it is counterargued by *arg*' and p does not manage to counterargue *arg*' (either because there are no counterarguments or because the resources are exhausted).

When an argument is introduced by one of the players, the beliefs that are part of it are retrieved into the set of active beliefs. When an argument is counterargued, its claim becomes provisional. If an argument is defeated, its claim is rejected.

The protocol we will be using assumes that the resources are equally divided, i.e., if player p_1 has exhausted his share of resources but p_2 has not, then p_2 is still allowed one move. Except for this situation, the players alternate the moves. No repetition of counterargued (sub-)arguments is allowed.

Suppose a sentence α is observed. The current opinion is set to **pro** and **con** tries to find an argument for $\neg\alpha$. If he fails, then α is accepted, otherwise, current opinion is set to **con** and **pro** tries to either counterargue the last argument or present a new argument for α. If **pro** fails, then α is rejected. Otherwise, current opinion is set to **pro** and **con** tries to either counterargue the last argument or present a new argument for $\neg\alpha$. The process continues until resources are exhausted. The player favored by current opinion wins.

A.3 Example

We will now see an example application of the protocol described in Section A.2.

We first have to give some more details about the procedure. The claim to be verified, a provisional belief, remains active during the whole argumentation. It cannot be dismissed due to overflow in the set of active beliefs. The set of active beliefs is ordered by recency, i.e., beliefs that have been used first are the first to be forgotten in case of overflow. However, if an active belief is reused, it becomes more recent and changes place in the order. This agrees with cognitive models of memory, as for example in [And80].

The claim which is being verified and claims of arguments that have been counterargued cannot be used in new arguments.

The size of the set of active beliefs, one of the limited resources, is given by the number of atoms occurring in its formulas. Part of the history of the process is kept in the form of arguments advanced, so that there is no repetition. This should also be limited in size like the set of active beliefs, but in the example we will ignore this fact.

We will use the following logic for the example:

1. atoms $a, b, c, ..., p$ standing for "*a*lbert comes to the party", "*b*etty comes to the party", "*c*harles comes to the party", ..., "*p*atrick comes to the party".
2. formulas $x \to y$ standing for "If x comes to the party, then y comes to the party" ; $x \to \neg y$ standing for "If x comes to the party, then y does not come to the party", etc.
3. inference rules modus ponens ($x, x \to y \Rightarrow y$) and inversion ($x \to y \Rightarrow \neg y \to \neg x$).

Depending on who likes whom and who dislikes whom, we know who is (or is not) going to come to the party given who is (or is not) coming. Moreover, we know of some people that are coming ($a$lbert, $f$erry, $h$arold, $k$ate, and $o$swald). Our initial set of explicit beliefs is:

$E = \{a, a \to b, b \to c, c \to d, d \to g, f, f \to e, e \to \neg c, \neg c \to \neg p, h, h \to i, i \to j, j \to \neg e, k, k \to l, l \to m, m \to n, n \to \neg i, o, o \to p, p \to \neg l, \neg l \to \neg b\}$.

We assume that the maximum size of the set of active beliefs is 20. We want to know whether $i$van is coming to the party:

- Step 1: **con** tries to refute i, presenting an argument for $\neg i$. The formulas in the argument are retrieved from the set of explicit beliefs and stored as active beliefs. Inference is applied four times in order to get to the claim $\neg i$ from the argument.

  - **con** presents argument $\{k, k \to l, l \to m, m \to n, n \to \neg i\}$ for $\neg i$.
  - 9 basic operations: retrieval $\{k, k \to l, l \to m, m \to n, n \to \neg i\}$; inference $\{l, m, n, \neg i\}$
  - $A = \{i, k, k \to l, l, l \to m, m, m \to n, n, n \to \neg i, \neg i\}$; $|A|$=14
  - History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}\}$
  - current.opinion = **con**

- Step 2: **pro** advances an argument against one of the subclaims of the previous argument. The previous argument is counterargued, but not yet defeated, since **con** may counterargue this present argument. The set of active beliefs grows to its maximum size, 20. The oldest active belief besides the claim (k) is dismissed to make space for the new activated beliefs.

  - **pro** presents counterargument $\{o, o \to p, p \to \neg l\}$ against l.
  - 5 basic operations: retrieval $\{o, o \to p, p \to \neg l\}$; inference $\{p, \neg l\}$.
  - $A = \{i, k \to l, l, l \to m, m, m \to n, n, n \to \neg i, \neg i, o, o \to p, p, p \to \neg l, \neg l\}$; $|A|$=20
  - History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}, \{o, o \to p, p \to \neg l\}\}$
  - current.opinion = **pro**

- Step 3: **con** counterargues the previous argument. The oldest elements of the set of active beliefs (except i) are dismissed to make space for the new beliefs retrieved.

  - **con** presents counterargument $\{f, f \to e, e \to \neg c, \neg c \to \neg p\}$ against p.

- 7 basic operations: retrieval $\{f, f \to e, e \to \neg c, \neg c \to \neg p\}$; inference $\{e, \neg c, \neg p\}$.
- $A = \{i, \neg i, o, o \to p, p, p \to \neg l, \neg l, f, f \to e, e, e \to \neg c, \neg c, \neg c \to \neg p, \neg p\}\}$; $|A|$=19
- History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}, \{o, o \to p, p \to \neg l\}, \{f, f \to e, e \to \neg c, \neg c \to \neg p\}\}$
- current.opinion = **con**

• Step 4: **pro** counterargues the previous argument. Again, some elements of the set of active beliefs must be dismissed.

- **pro** presents counterargument $\{a, a \to b, b \to c\}$ against $\neg c$.
- 5 basic operations: retrieval $\{a, a \to b, b \to c\}$; inference $\{b, c\}$
- $A = \{i, \neg l, f, f \to e, e, e \to \neg c, \neg c, \neg c \to \neg p, \neg p, a, a \to b, b, b \to c, c\}$; $|A|$=19
- History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}, \{o, o \to p, p \to \neg l\}, \{f, f \to e, e \to \neg c, \neg c \to \neg p\}, \{a, a \to b, b \to c\}\}$
- current.opinion = **pro**

• Step 5: **con**'s arguments were defeated, since he cannot counterargue the previous arguments advanced by **pro** anymore. **con** advances a new argument against i. Some of the beliefs used in this argument $(f, f \to e, e \to \neg j)$ are already active so they do not need to be retrieved. They only change place in the set of active beliefs.

- **con** presents counterargument $\{f, f \to e, e \to \neg j, \neg j \to \neg i\}$ against i.
- 6 basic operations: retrieval $\{j \to \neg e, i \to j\}$; inferences $\{e \to \neg j, \neg j, \neg j \to \neg i, \neg i\}$
- $A = \{i, b, b \to c, c, f, f \to e, e, j \to \neg e, e \to \neg j, \neg j, i \to j, \neg j \to \neg i, \neg i\}$; $|A|$=19
- History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}, \{o, o \to p, p \to \neg l\}, \{f, f \to e, e \to \neg c, \neg c \to \neg p\}, \{a, a \to b, b \to c\}, \{f, f \to e, e \to \neg j, \neg j \to \neg i\}\}$
- current.opinion = **con**

• Step 6: **pro** cannot counterargue the previous argument, but presents instead a new argument for i. Since **con** does not have any other counterarguments or arguments for $\neg i$, **pro** wins the disputation.

- **pro** presents argument $\{h, h \to i\}$ for i.

- 3 basic operations: retrieval $\{h, h \to i\}$; inference $\{i\}$.
- $A = \{i, c, f, f \to e, e, j \to \neg e, e \to \neg j, \neg j, i \to j, \neg j \to \neg i, \neg i, h, h \to i\}$; $|A|$=19
- History: $\{\{k, k \to l, l \to m, m \to n, n \to \neg i\}, \{o, o \to p, p \to \neg l\}, \{f, f \to e, e \to \neg c, \neg c \to \neg p\}, \{a, a \to b, b \to c\}, \{f, f \to e, e \to \neg j, \neg j \to \neg i\}, \{h, h \to i\}\}$
- current.opinion = **pro**

A.4 Discussion

The aim of this chapter was to give a flavor as to how argumentation theory can be used in order to decide which beliefs should be fully accepted. These ideas enrich the framework presented in Chapter 4.

The protocol and the example presented are clearly too simple minded to encode the real decision process, but they illustrate the internal process of "weighing" the arguments in favor and against a certain claim that takes place when an agent is confronted with information about which he is not sure.

In most argumentation systems there is some notion of priority which is used in two phases: first to help one player to select which argument he is going to advance among the several possibilities, and then to judge whether the argument advanced is good enough to defeat the other player's argument. Another feature of argumentation systems which we have ignored here is the existence of defeasible implications.

In our example, we have a set which collects the arguments that have occurred during the process. Since we are dealing with resource-bounded agents, this set must also be limited in size. The set is used to avoid repetitions of arguments. But if an argument was used so early that it has already been forgotten, there is no way to avoid repetition.

Future work includes examining existing implemented argumentation systems in order to refine the protocol of the argumentation process. One such system is presented in [SL92], together with a mathematical treatment of the relations between arguments.

It would also be interesting to check whether we can have a uniform treatment of priority, i.e., whether the same notion of priority used to select arguments can be used for the revision mechanism.

Appendix B

Proofs related to Chapter 4

Proposition 4.2.2: Given any two belief states $\Sigma_1 = \langle E_1, A_1 \rangle$ and $\Sigma_2 = \langle E_2, A_2 \rangle$, there is a sequence of basic operations that takes Σ_1 into Σ_2.

Proof: Let $\Sigma_1 = \langle E_1, A_1 \rangle$ and $\Sigma_2 = \langle E_2, A_2 \rangle$. Since the size of the set of active beliefs is limited, we have to move the formulas one by one. Let $\{\varphi_1, \varphi_2, ..., \varphi_n\}$ be an enumeration of the elements of $E_1 \setminus E_2$ and $\{\psi_1, \psi_2, ..., \psi_m\}$ an enumeration of the elements of $E_2 \setminus E_1$. We have:

$\Sigma_1 \circ_r \varphi_1 \circ_d \varphi_1 \circ_c \varphi_1 \circ_r \varphi_2 \circ_d \varphi_2 \circ_c \varphi_2 \circ_r ..... \circ_r \varphi_n \circ_d \varphi_n \circ_c \varphi_n = \langle E_1 \cap E_2, A_1 \rangle$.

$\langle E_1 \cap E_2, A_1 \rangle \circ_o \psi_1 \circ_a \psi_1 \circ_o \psi_2 \circ_a \psi_2 \circ_o ... \circ_o \psi_m \circ_a \psi_m = \langle E_2, A_1 \rangle$.

In the same way, by repeating the operations of rejection ($\circ_c$) for the elements of $A_1 \setminus A_2$ and then observation ($\circ_o$) for the elements of $A_2 \setminus A_1$, we obtain Σ_2. $\square$

Lemma 4.3.4: $f(K)\dot{+}\varphi = f(K + \varphi)$

Proof: $f(K)\dot{+}\varphi = Cl(\langle K, Cn, \emptyset \rangle \circ_o \varphi \circ_a \varphi) = Cl(\langle K, Cn, \{\varphi\} \rangle \circ_a \varphi) = Cl(\langle K \cup \{\varphi\}, Cn, \emptyset \rangle) = (\textit{Lemma}\ 4.3.2) \langle Cn(K \cup \{\varphi\}), Cn, \emptyset \rangle = f(K + \varphi)$ $\square$

Lemma 4.3.6: If $-_\gamma$ is the AGM partial meet contraction operator based on γ, then $f(K)\dot{-}_\gamma\{\varphi\} = f(K -_\gamma \varphi)$.

Proof: $f(K)\dot{-}_\gamma\{\varphi\} = \langle K, Cn, \emptyset \rangle \circ_r K \setminus (\bigcap \gamma(K \bot \alpha)) \circ_d K \setminus (\bigcap \gamma(K \bot \alpha)) \circ_c K \setminus (\bigcap \gamma(K \bot \alpha)) = \langle K, Cn, K \setminus (\bigcap \gamma(K \bot \alpha)) \rangle \circ_d K \setminus (\bigcap \gamma(K \bot \alpha)) \circ_c K \setminus (\bigcap \gamma(K \bot \alpha)) = \langle \bigcap \gamma(K \bot \alpha), Cn, K \setminus (\bigcap \gamma(K \bot \alpha)) \rangle \circ_c K \setminus (\bigcap \gamma(K \bot \alpha)) = \langle \bigcap \gamma(K \bot \alpha), Cn, \emptyset \rangle = f(K -_\gamma \varphi)$. $\square$

Proposition 4.5.3: Given any two belief states $\Sigma_1 = \langle E_1, A_1, W_1 \rangle$ and $\Sigma_2 = \langle E_2, A_2, W_2 \rangle$, there is a sequence of basic operations that takes Σ_1 into Σ_2.

Proof: Since the size of the set of active beliefs is limited, we have to move the formulas one by one. Let $\{\varphi_1, \varphi_2, ..., \varphi_n\}$ be an enumeration of the elements of $E_1 \setminus E_2$ and $\{\psi_1, \psi_2, ..., \psi_m\}$ an enumeration of the elements of $E_2 \setminus E_1$. We have:

$\Sigma_1 \circ_r \varphi_1 \circ_{d_e} \varphi_1 \circ_c \varphi_1 \circ_r \varphi_2 \circ_{d_e} \varphi_2 \circ_c \varphi_2 \circ_r \ldots\ldots \circ_r \varphi_n \circ_{d_e} \varphi_n \circ_c \varphi_n = \langle E_1 \cap E_2, A_1, X \rangle$.
$\langle E_1 \cap E_2, A_1, X \rangle \circ_o \psi_1 \circ_a \psi_1 \circ_o \psi_2 \circ_a \psi_2 \circ_o \ldots \circ_o \psi_m \circ_a \psi_m = \langle E_2, A_1, Y \rangle$.
$\langle E_2, A_1, Y \rangle \circ_c (A_1 \setminus A_2) \circ_o (A_2 \setminus A_1) = \langle E_2, A_2, Z \rangle$.
By Observation 4.5.2, $Z \subseteq (A_2)$.
$\langle E_2, A_2, Z \rangle \circ_{d_i} (Z \setminus W_2) \circ_g (W_2 \setminus Z) = \Sigma_2$. □

Proposition 4.5.4: The simplified model can be embedded in the refined one under the assumption that the agent is introspective, i.e., that $\mathbf{W}{=}\mathbf{A}$.

Proof: Let f' be a function mapping simplified belief states into refined belief states such that:
$f'(\langle \mathbf{E}, \mathbf{A} \rangle) = \langle \mathbf{E}, \mathbf{A}, \mathbf{A} \rangle$.

We will show that each of the operations in Definition 4.2.1 can be simulated by a sequence of operations in Definition 4.5.1, i.e., that f' maps belief states modified by basic operations in the simplified model into belief states in the refined model satisfying $\mathbf{W}{=}\mathbf{A}$.

We will use the same symbols for the operations in Definitions 4.2.1 and 4.5.1, it will be clear from the context whether the operations are the ones of the simplified or the refined model.

- Observation: $f'(\langle \mathbf{E}, \mathbf{A} \rangle \circ_o \alpha) = f'(\langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle) = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A} \rangle) \circ_o \alpha \circ_g \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A} \rangle \circ_o \alpha \circ_g \alpha = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \rangle \circ_g \alpha = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\Sigma \circ_o \alpha) = f'(\Sigma) \circ_o \alpha \circ_g \alpha$.

- Retrieval: $f'(\langle \mathbf{E}, \mathbf{A} \rangle \circ_r \alpha) = f'(\langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle) = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A} \rangle) \circ_r \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A} \rangle \circ_r \alpha = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\Sigma \circ_r \alpha) = f'(\Sigma) \circ_r \alpha$.

- Acceptance: $f'(\langle \mathbf{E}, \mathbf{A} \rangle \circ_a \alpha) = f'(\langle \mathbf{E} \cup^* \{\alpha\}, \mathbf{A} \setminus \{\alpha\} \rangle) = \langle \mathbf{E} \cup^* \{\alpha\}, \mathbf{A} \setminus \{\alpha\}, \mathbf{A} \setminus \{\alpha\} \rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A} \rangle) \circ_a \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A} \rangle \circ_a \alpha = \langle \mathbf{E} \cup^* \{\alpha\}, \mathbf{A} \setminus \{\alpha\}, \mathbf{A} \setminus \{\alpha\} \rangle$.

  $f'(\Sigma \circ_a \alpha) = f'(\Sigma) \circ_a \alpha$.

- Inference: $f'(\langle \mathbf{E}, \mathbf{A} \rangle \circ_i \alpha) = f'(\langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\} \rangle) = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A} \rangle) \circ_i \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A} \rangle \circ_i \alpha = \langle \mathbf{E}, \mathbf{A} \cup^* \{\alpha\}, \mathbf{A} \cup^* \{\alpha\} \rangle$.

  $f'(\Sigma \circ_i \alpha) = f'(\Sigma) \circ_i \alpha$.

- Doubting: $f'(\langle \mathbf{E}, \mathbf{A}\rangle \circ_d \alpha) = f'(\langle \mathbf{E} \setminus \{\alpha\}, \mathbf{A}\rangle) =$
  $\langle \mathbf{E} \setminus \{\alpha\}, \mathbf{A}, \mathbf{A}\rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A}\rangle) \circ_{d_e} \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A}\rangle \circ_{d_e} \alpha = \langle \mathbf{E} \setminus \{\alpha\}, \mathbf{A}, \mathbf{A}\rangle$.

  $f'(\Sigma \circ_d \alpha) = f'(\Sigma) \circ_{d_e} \alpha$.

- Rejection: $f'(\langle \mathbf{E}, \mathbf{A}\rangle \circ_c \alpha) = f'(\langle \mathbf{E}, \mathbf{A} \setminus \{\alpha\}\rangle) = \langle \mathbf{E}, \mathbf{A} \setminus \{\alpha\}, \mathbf{A} \setminus \{\alpha\}\rangle$.

  $f'(\langle \mathbf{E}, \mathbf{A}\rangle) \circ_c \alpha = \langle \mathbf{E}, \mathbf{A}, \mathbf{A}\rangle \circ_c \alpha = \langle \mathbf{E}, \mathbf{A} \setminus \{\alpha\}, \mathbf{A} \setminus \{\alpha\}\rangle$.

  $f'(\Sigma \circ_c \alpha) = f'(\Sigma) \circ_c \alpha$.

$\square$

Lemma 4.5.5: $h(K)\ddot{+}\alpha = h(K + \alpha)$.

Proof: $h(K)\ddot{+}\alpha = \langle K, Cn, \emptyset, \emptyset\rangle\ddot{+}\alpha = Cl(\langle K, Cn, \emptyset, \emptyset\rangle \circ_o \alpha \circ_g \alpha \circ_a \alpha) = Cl(\langle K, Cn, \{\alpha\}, \emptyset\rangle \circ_g \alpha \circ_a \alpha) = Cl(\langle K, Cn, \{\alpha\}, \{\alpha\}\rangle \circ_a \alpha =$
$Cl(\langle K \cup \{\alpha\}, Cn, \emptyset, \emptyset\rangle) = \langle Cn(K \cup \{\alpha\}), Cn, \emptyset, \emptyset\rangle = h(Cn(K \cup \{\alpha\})) = h(K + \alpha)$. $\square$

Lemma 4.5.6: $h(K)\ddot{-}_\gamma\alpha = h(K -_\gamma \alpha)$.

Proof: $h(K)\ddot{-}_\gamma\alpha = \langle K, Cn, \emptyset, \emptyset\rangle\ddot{-}_\gamma\alpha = \langle K, Cn, \emptyset, \emptyset\rangle \circ_r \Delta \circ_{d_e} \Delta \circ_c \Delta = \langle K, Cn, \Delta, \Delta\rangle \circ_{d_e} \Delta \circ_c \Delta = \langle K \setminus \Delta, Cn, \Delta, \Delta\rangle \circ_c \Delta = \langle \bigcap \gamma(K \perp \alpha), Cn, \emptyset, \emptyset\rangle = h(\bigcap \gamma(K \perp \alpha)) = h(K -_\gamma \alpha)$. $\square$

Appendix C

Proofs related to Chapter 5

Observation 5.1.2: Let C be an inference operation on the language L and $\perp\!\!\!\perp_C$ its associated kernel operation. If C satisfies compactness, then $B \perp\!\!\!\perp_C \alpha \neq \emptyset$ for all $B \subseteq L$ and all $\alpha \in C(B)$.

Proof: Let compactness be satisfied and $\alpha \in C(B)$. Then there is some finite subset Z of B such that $\alpha \in C(Z)$. Let X be any inclusion-minimal subset of Z with the property $\alpha \in C(X)$. Then $X \in B \perp\!\!\!\perp_C \alpha$. $\square$

Observation 5.1.3:

1. If $B \subseteq B'$, then for every formula α, $(B \perp\!\!\!\perp_C \alpha) \subseteq (B' \perp\!\!\!\perp_C \alpha)$.

2. If C satisfies compactness and monotony, then $B \perp\!\!\!\perp_C \alpha = B \perp\!\!\!\perp_C \beta$ if and only if for all subsets B' of B, $\alpha \in C(B')$ iff $\beta \in C(B')$.

3. $X \in B \perp\!\!\!\perp_C \alpha$ if and only if $X \subseteq B$ and $X \in X \perp\!\!\!\perp_C \alpha$.

Proof: (Part 1:) Let $X \in (B \perp\!\!\!\perp_C \alpha)$. Then X is an inclusion-minimal subset of B such that $\alpha \in C(X)$. Since $B \subseteq B'$, X is also an inclusion-minimal subset of B' such that $\alpha \in C(X)$, i.e., $X \in (B' \perp\!\!\!\perp_C \alpha)$.

(Part 2:) Without loss of generality, suppose that there is some $B' \subseteq B$ such that $\alpha \in C(B')$ and $\beta \notin C(B')$. Then, since C is compact, there is some $B'' \subseteq B'$ such that $B'' \in (B \perp\!\!\!\perp_C \alpha)$. It follows from $\beta \notin C(B')$ by the monotony of C that $\beta \notin C(B'')$ and hence, $B'' \notin (B \perp\!\!\!\perp_C \beta)$. For the other side of the implication, suppose, without loss of generality, that there is $B' \in (B \perp\!\!\!\perp_C \alpha)$ such that $B' \notin (B \perp\!\!\!\perp_C \beta)$. If $\beta \notin C(B')$, B' is a subset of B such that $\alpha \in C(B')$ and $\beta \notin C(B')$. Otherwise, if $\beta \in C(B')$, then since $B' \notin (B \perp\!\!\!\perp_C \beta)$, B' is not minimal and there is $B'' \subset B'$ such that $\beta \in C(B'')$. But since $B' \in (B \perp\!\!\!\perp_C \alpha)$, $\alpha \notin C(B'')$.

(Part 3:) That $X \subseteq B$ and $X \in X \perp\!\!\!\perp_C \alpha$ imply $X \in B \perp\!\!\!\perp_C \alpha$ follows directly from Part 1. For the other direction of the implication, note that from the definition of $\perp\!\!\!\perp_C$, $X \in B \perp\!\!\!\perp_C \alpha$ implies that $X \subseteq B$ and that X is an inclusion-minimal subset of B such that $\alpha \in C(X)$. Then X is an inclusion minimal subset of X such that $\alpha \in C(X)$, i.e., $X \in X \perp\!\!\!\perp_C \alpha$. □

Observation 5.1.5:
(1) For all sets A and B of sentences, $c(A, B) = c(A, c(A, B))$.
(2) If $A \subseteq A'$ and $B \subseteq B'$, then $c(A, B) \subseteq c(A', B')$.

Proof: (1) That $c(A, c(A, X)) \subseteq c(A, X)$ follows directly from Definition 5.1.4. To see that $c(A, X) \subseteq c(A, c(A, X))$, let $\alpha \in c(A, X)$. This means that there is some $\delta \in A$ and some $D \subseteq X$ such that $\alpha \in D$, D is C-consistent, and either $D \in X \perp\!\!\!\perp_C \delta$ or $D \in X \perp\!\!\!\perp_C \neg\delta$. But then we have that $D \subseteq c(A, X)$ and since $c(A, X) \subseteq X$, $D \in c(A, X) \perp\!\!\!\perp_C \delta$ or $D \in c(A, X) \perp\!\!\!\perp_C \neg\delta$ and hence $\alpha \in D \subseteq c(A, c(A, X))$.

(2) It follows directly from Definition 5.1.4 that it is sufficient to show that, for any sentence δ, if $B \subseteq B'$ then $c(\delta, B) \subseteq c(\delta, B')$. Let $\beta \in c(\delta, B)$. It follows from Definition 5.1.4 that there is some X such that $\beta \in X$, X is C-consistent, and either $X \in B \perp\!\!\!\perp_C \delta$ or $X \in B \perp\!\!\!\perp_C \neg\delta$. It follows from $X \in B \perp\!\!\!\perp_C \delta$ and $B \subseteq B'$ that $X \in B' \perp\!\!\!\perp_C \delta$, and similarly from $X \in B \perp\!\!\!\perp_C \neg\delta$ and $B \subseteq B'$ that $X \in B' \perp\!\!\!\perp_C \neg\delta$. We can conclude from this that $\beta \in c(\delta, B')$. □

Observation 5.1.7: Let C be an inference operation that satisfies monotony, compactness, and inclusion ($X \subseteq C(X)$). Then:

1. $\beta \in C_\alpha(B)$ iff $\beta \in C_{\neg\alpha}(B)$
2. If B is C-consistent, then:
    - if $\alpha \in A$, then $\alpha \in C_A(B)$ iff $\alpha \in C(B)$
    - $C_B(B) = C(B)$
3. $\alpha \in C_\top(B)$ iff $\alpha \in C_\perp(B)$ iff $\alpha \in C(\emptyset)$
4. If all elements of B are contingent, then $C_L(B) = C(B)$
5. $C_A(B) \subseteq C(B)$.

Proof:

1. Let $\beta \in C_\alpha(B)$. Then $\beta \in C(c(\alpha, B))$ and from Definition 5.1.4 it follows that $\beta \in C(c(\neg\alpha, B))$. Hence, $\beta \in C_{\neg\alpha}(B)$.
2. Let $\perp \notin C(B)$.
    - Let $\alpha \in A$. If $\alpha \in C_A(B)$, then $\alpha \in C(c(A, B))$ and since $c(A, B) \subseteq B$, it follows from monotony that $\alpha \in C(B)$. For the other side of the implication, let $\alpha \in C(B)$. Then it follows from compactness that

there is an inclusion-minimal subset X of B such that $\alpha \in C(X)$. Since $\perp \notin C(B)$, by monotony, $\perp \notin C(X)$. Hence, $X \in B \perp\!\!\!\perp_C \alpha$ and $X \notin B \perp\!\!\!\perp_C \perp$, i.e., $X \subseteq c(A, B)$ and from monotony it follows that $\alpha \in C(c(A, B)) = C_A(B)$.

- From $\alpha \in C_B(B) = C(c(B, B))$ it follows by monotony that $\alpha \in C(B)$. For the other side of the implication, let $\alpha \in C(B)$. Then by compactness, there is some inclusion-minimal $X \subseteq B$ such that $\alpha \in C(X)$. For every element δ of X it holds that $\delta \in c(\delta, B)$ and since $\delta \in X \subseteq B$, $\delta \in c(B, B)$. Hence, $X \subseteq c(B, B)$ and from monotony it follows that $\alpha \in C(c(B, B)) = C_B(B)$.

3. That $\alpha \in C_\top(B)$ if and only if $\alpha \in C_\perp(B)$ follows from part 1. Let $\alpha \in C_\top(B)$. Then $\alpha \in C(c(\top, B))$ and by Definition 5.1.4, $\alpha \in C(\emptyset)$. In the same way, if $\alpha \in C(\emptyset)$, then by Definition 5.1.4, $\alpha \in C(c(\top, B)) = C_\top(B)$.

4. It suffices to proof that $c(L, B) = B$. That $c(L, B) \subseteq B$ follows directly form the definition of compartments. Let $\alpha \in B$. Since α is contingent, from Definition 5.1.4 it follows that $\alpha \in c(\alpha, B)$ and from Part 2 of Observation 5.1.5 it follows that $\alpha \in c(L, B)$. Hence, $B \subseteq c(L, B)$.

5. Since $c(A, B) \subseteq B$, it follows from the monotony of C that $C_A(B) = C(c(A, B)) \subseteq C(B)$.

□

Observation 5.1.8: The elements of $B \perp\!\!\!\perp_{C_A} \alpha$ are subsets of $c(A, B)$.

Proof: Let $X \in B \perp\!\!\!\perp_{C_A} \alpha$. Then $\alpha \in C_A(X)$, that is, $\alpha \in C(c(A, X))$ and for no proper subset X' of X it holds that $\alpha \in C_A(X')$. From part 1 of Observation 5.1.5, it follows that $\alpha \in C(c(A, c(A, X))) = C_A(c(A, X))$. By part 2 of Observation 5.1.5, since $X \subseteq B$, $c(A, X) \subseteq c(A, B)$. Suppose, for contradiction, that $X \not\subseteq c(A, B)$. Then, since $c(A, X) \subseteq c(A, B)$, $c(A, X)$ must be a proper subset of X, and since $\alpha \in C_A(c(A, X))$, this contradicts the minimality of X. Hence, $X = c(A, X) \subseteq c(A, B)$. □

Observation 5.1.9: For all sets B of sentences it holds that $(C_A)_A(B) = C_A(B)$.

Proof: By definition, $(C_A)_A(B) = C_A(c(A, B)) = C(c(A, c(A, B))) =$ (Observation 5.1.5, part 1) $C(c(A, B)) = C_A(B)$. □

Theorem 5.1.10: Let be C_A the A-localization of an inference operation C. Then:

1. If C satisfies monotony, then C_A satisfies monotony and Sen.

2. If C satisfies monotony and compactness, then C_A satisfies compactness.

3. If C satisfies monotony, compactness, and weak iteration, then C_A satisfies weak iteration.

4. If C satisfies monotony, compactness, weak iteration, and inclusion, then C_A satisfies idempotency (iteration) and cumulativity.

5. If C satisfies monotony and consistency preservation, then so does C_A.

6. If C satisfies weak explosiveness, than so does C_A.

7. If C satisfies monotony and inclusion, then C_A satisfies embedded inclusion.

8. If C satisfies monotony, compactness, weak explosiveness, and non-contravention, then C_A satisfies α-local non-contravention for all $\alpha \in A$.

Proof: *Part 1*: Let $\alpha \in C_A(B)$, i.e., $\alpha \in C(c(A, B))$ and let D be a superset of B, i.e., $B \subseteq D$. It follows from part 2 of observation 5.1.5 that $c(A, B) \subseteq c(A, D)$. Hence, since C is monotonic, $\alpha \in C(c(A, D))$, i.e., $\alpha \in C_A(D)$.

Part 2: Let C satisfy monotony and compactness. For compactness, let $\alpha \in C_A(B)$, i.e., $\alpha \in C(c(A, B))$. From the compactness of C it follows that there is an inclusion-minimal finite set $D' \subseteq c(A, B) \subseteq B$ such that $\alpha \in C(D')$. Since $D' \subseteq c(A, B)$ and D' is finite, there is a consistent inclusion-minimal finite set $D \subseteq B$ such that for every $\varepsilon \in D'$ there is a finite set $X \subseteq D$ and there is an element δ of A such that (i) $X \vdash \delta$ or $X \vdash \neg\delta$, and (ii) $\varepsilon \in X$. By construction, $D' \subseteq D$ and moreover, $c(A, D) = D$. From this and $\alpha \in C(D')$ we conclude, since C is compact, that $\alpha \in C(D)$ and hence $\alpha \in C_A(D)$.

Part 3: Let C satisfy monotony and weak iteration. Let $\alpha \in C_A(C_A(B)) = C(c(A, C_A(B)))$. This means that there is a set $X \subseteq c(A, C_A(B))$ such that $\alpha \in C(X)$. Since $c(A, C_A(B)) \subseteq C_A(B)$, we have $X \subseteq C_A(B) = C(c(A, B))$. It follows from the compactness of C that there is some set $Y \subseteq c(A, B)$ such that $X \subseteq C(Y)$ and thus, by the monotony and weak iteration of C, $\alpha \in C(Y)$. Hence, $\alpha \in C_A(B)$ and $C_A(C_A(B)) \subseteq C_A(B)$.

Part 4: For idempotency, it suffices to show that $C_A(B) \subseteq C_A(C_A(B))$. Let $\alpha \in C_A(B) = C(c(A, B))$. Then there is a set $X \subseteq c(A, B)$ such that $\alpha \in C(X)$. By part 1 of observation 5.1.5, $c(A, B) = c(A, c(A, B))$ and from part 2 of observation 5.1.5 together with the inclusion property of C it follows that $c(A, B) \subseteq c(A, C(c(A, B))) = c(A, C_A(B))$. Hence, $X \subseteq c(A, C_A(B))$ and thus by the monotony of C, $\alpha \in C(c(A, C_A(B))) = C_A(C_A(B))$. Cumulativity follows directly from monotony and idempotency.

Part 5: Monotony follows from part 1. For consistency preservation, let $\perp \in C_A(B)$. By definition this means that $\perp \in C(c(A, B))$. Since $c(A, B) \subseteq B$ and C satisfies monotony, we have $\perp \in C(B)$. Hence, from the consistency preservation of C, $\perp \in Cn(B)$.

Part 6: Let C satisfy weak explosiveness and let $\perp \in C_A(B)$. This means that $\perp \in C(c(A,B))$. It follows from the weak explosiveness of C that for every sentence α, $\alpha \in C(c(A,B)) = C_A(B)$.

Part 7: Let C satisfy monotony and inclusion. By Part 1 of Observation 5.1.5, $C_A(B) = C(c(A,B)) = C(c(A,c(A,B)))$. By the inclusion property of C, $c(A,B) \subseteq C(c(A,B))$ and from Part 2 of Observation 5.1.5 it follows that $c(A,c(A,B)) \subseteq c(A,C(c(A,B)))$. Hence, by the monotony of C, $C_A(B) = C(c(A,c(A,B))) \subseteq C(A,C(c(A,B))) = C_A(C_A(B))$.

Part 8: Let C satisfy monotony, compactness, weak explosiveness and non-contravention. Let $\alpha \in A$. If $\perp \in C(B)$, then by weak explosiveness $\neg\alpha \in C(B)$ and α-local non-contravention holds trivially. Let $\perp \notin C(B)$ and $\neg\alpha \in C_A(B \cup \{\alpha\}) = C(c(A,B \cup \{\alpha\})$. Then by the monotony of C, $\neg\alpha \in C(B \cup \{\alpha\})$ and from the non-contravention of C it follows that $\neg\alpha \in C(B)$. Since $\perp \notin C(B)$ and $\alpha \in A$, it follows from part 2 of Observation 5.1.7 that $\neg\alpha \in C_A(B)$. □

Theorem 5.1.11: For each of the following properties there is some set of sentences A such that Cn_A, the A-localization of the classical truth functional consequence operator Cn, does not satisfy the property: inclusion, supraclassicality, deduction property, reductio ad absurdum, falsity, distributivity, explosiveness, and non-contravention.

Proof: Let p, q, r, and s be logically independent.

1. (Inclusion) Let $B = \{p,q\}$ and $A = \{p\}$. Then $q \notin Cn_A(B) = Cn(\{p\})$.

2. (Supraclassicality) Let $B = \{p\}$, $A = \{q\}$. Then $p \in Cn(B)$ but $p \notin Cn_A(B)$.

3. (Deduction property) Let $B = \{r, r \to (p \to q)\}$ and $A = \{q\}$. Then $c(A, B \cup \{p\}) = B \cup \{p\}$ and hence $q \in Cn_A(B \cup \{p\})$. But $c(A,B) = \emptyset$, thus, $p \to q \notin Cn_A(B)$.

4. (Reductio ad absurdum) Let $B = \{\neg p \to r, \neg r\}$ and $A = \{r\}$. Then $c(A, B \cup \{\neg p\}) = B \cup \{\neg p\}$ and hence, $\perp \in Cn_A(B \cup \{\neg p\})$. On the other hand, $c(A,B) = \{\neg r\}$ from which it follows that $p \notin Cn_A(B) = Cn(\{\neg r\})$.

5. (Falsity) By definition $c(A, \perp) = \emptyset$, hence, $Cn_A(\{\perp\}) = Cn(\emptyset)$.

6. (Distributivity) Let $B = \{p, p \to q, q \to r\}$, $D = \{p, p \to s, s \to \neg r\}$ and $A = \{r\}$. Then we have $c(A,B) = B$, $c(A,D) = D$, and therefore, $p \in Cn_A(B) \cap Cn_A(D)$. On the other hand, $Cn(B) \cap Cn(D) = Cn(\{p \wedge (q \vee s)\})$. Hence, $c(A, Cn(Cn(B) \cap Cn(D))) = \emptyset$, and $p \notin Cn_A(Cn(B) \cap Cn(D))$.

7. (Explosiveness) Let $B = \{p, \neg p\}$ and $A = \{q\}$. Then $r \notin Cn_A(B) = Cn(c(q, \{p, \neg p\})) = Cn(\emptyset)$.

8. (Non-contravention) Let $B = \{p \rightarrow q, p \rightarrow \neg q\}$ and $A = \{q\}$. Then $\neg p \in Cn_A(B \cup \{\alpha\}) = Cn(B \cup \{\alpha\})$, but $\neg p \notin Cn_A(B) = Cn(\emptyset)$.

□

Theorem 5.2.3: Let C be an inference operation satisfying monotony and compactness. Then $\dot{-}$ is an operation of kernel contraction on B determined by C and some incision function if and only if for all sentences α:

- If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)
- $B \dot{-} \alpha \subseteq B$ (inclusion)
- If $\beta \in B \setminus B \dot{-} \alpha$, then there is some $B' \subseteq B$ such that $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\beta\})$ (core-retainment)
- If for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)

Proof: (i) construction $\Rightarrow$ postulates: Let $\dot{-}_{C,\sigma}$ be an operation of kernel contraction defined by an inference operator C and an incision function σ. By the definition of kernel contraction, $\dot{-}_{C,\sigma}$ satisfies *inclusion*. In order to prove that *success* is satisfied, suppose to the contrary that $\alpha \in C(B \dot{-}_{C,\sigma} \alpha)$ for some α such that $\alpha \notin C(\emptyset)$. It follows from monotony and compactness that there must be some $X \in B \perp\!\!\!\perp_C \alpha$ such that $X \subseteq B \dot{-}_{C,\sigma} \alpha$. From $\alpha \notin C(\emptyset)$ it follows that $X \neq \emptyset$ and by the definition of σ, $\sigma(B \perp\!\!\!\perp_C \alpha) \cap X \neq \emptyset$. Then, by the construction, $X \not\subseteq B \dot{-}_{C,\sigma} \alpha$.

For *core-retainment*, let β be an element of $B \setminus B \dot{-}_{C,\sigma} \alpha$. By the construction of $B \dot{-}_{C,\sigma} \alpha$, $\beta \in \sigma(B \perp\!\!\!\perp_C \alpha)$. Since $\sigma(B \perp\!\!\!\perp_C \alpha) \subseteq \bigcup(B \perp\!\!\!\perp_C \alpha)$, there is some set $X \in B \perp\!\!\!\perp_C \alpha$ such that $\beta \in X$. By the definition of kernel sets, $X \subseteq B$, $\alpha \in C(X)$ and for all $Y \subset X$, $\alpha \notin C(Y)$. Let B' be $X \setminus \{\beta\}$.

For *uniformity*, suppose that for all subsets B' of B it holds that $\alpha \in C(B')$ if and only if $\beta \in C(B')$. Then $B \perp\!\!\!\perp_C \alpha = B \perp\!\!\!\perp_C \beta$ and since σ is a function, $\sigma(B \perp\!\!\!\perp_C \alpha) = \sigma(B \perp\!\!\!\perp_C \beta)$. By the definition of kernel contraction, it follows that $B \dot{-}_{C,\sigma} \alpha = B \dot{-}_{C,\sigma} \beta$.

(ii) postulates $\Rightarrow$ construction: Let $\dot{-}$ be an operation that satisfies *inclusion*, *success*, *core-retainment* and *uniformity* and let B be a set of sentences. Let σ be such that for all sentences α, $\sigma(B \perp\!\!\!\perp_C \alpha) = B \setminus B \dot{-} \alpha$. We have to show (a) that $\dot{-}$ is based on σ in the manner of Definition 5.2.2, (b) that σ is well-defined, i.e., that it is a function for the given domain, and (c) that it is an incision function for the given domain.

(a) By *inclusion* and the definition of σ we have $B \dot{-} \alpha = B \setminus \sigma(B \perp\!\!\!\perp_C \alpha)$.

(b) To see that σ is a function, let α and β be such that $B \perp\!\!\!\perp_C \alpha = B \perp\!\!\!\perp_C \beta$. We have to show that $\sigma(B \perp\!\!\!\perp_C \alpha) = \sigma(B \perp\!\!\!\perp_C \beta)$. From $B \perp\!\!\!\perp_C \alpha = B \perp\!\!\!\perp_C \beta$ it follows due to the monotony of C that for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$. By *uniformity*, $B \dot{-} \alpha = B \dot{-} \beta$ and by the definition of σ and *inclusion*, $\sigma(B \perp\!\!\!\perp_C \alpha) = \sigma(B \perp\!\!\!\perp_C \beta)$.

(c) To show that σ is an incision function we need to prove that the two conditions of Definition 5.2.1 are satisfied. To prove the first of them, let $x \in \sigma(B \perp\!\!\!\perp_C \alpha)$. Then $x \in B \setminus B \dot{-} \alpha$ and it follows from *core-retainment* that there is a set $B' \subseteq B$ such that $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{x\})$. Due to monotony and compactness, there is a set $X \subseteq B' \cup \{x\}$ such that $X \in B \perp\!\!\!\perp_C \alpha$ and $x \in X$. We can conclude that $x \in \bigcup(B \perp\!\!\!\perp_C \alpha)$. For the second condition of Definition 5.2.1, let $\emptyset \neq X \in B \perp\!\!\!\perp_C \alpha$ and suppose that $X \cap \sigma(B \perp\!\!\!\perp_C \alpha) = \emptyset$. We know from part (a) that $B \dot{-} \alpha = B \setminus \sigma(B \perp\!\!\!\perp_C \alpha)$. Since $X \subseteq B$, we can conclude that $X \subseteq B \dot{-} \alpha$. Since $\alpha \in C(X)$, monotony yields $\alpha \in C(B \dot{-} \alpha)$. This contradicts *success*, and we may conclude that $X \cap \sigma(B \perp\!\!\!\perp_C \alpha) \neq \emptyset$. □

Theorem 5.2.8: Let C satisfy monotony and compactness. Then $\dot{-}$ is an operator of partial meet contraction on B based on C if and only if for all sentences α:

- If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)
- $B \dot{-} \alpha \subseteq B$ (inclusion)
- If $\beta \in B \setminus (B \dot{-} \alpha)$, then there is some B' such that $B \dot{-} \alpha \subseteq B' \subseteq B$, $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\beta\})$ (relevance)
- If for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)

Proof: (i) construction ⇒ postulates: Let $\dot{-}_{C,\gamma}$ be a partial meet contraction operator based on an operation C and a selection function γ. We need to show that it satisfies the four postulates.

Inclusion follows directly from the definition of partial meet contraction.

Success: Let $\alpha \notin C(\emptyset)$. Since C satisfies monotony and compactness, it follows from the upper bound property (Observation 5.2.6) that $B \perp_C \alpha$ is non-empty. Due to Definition 5.2.5, so is $\gamma(B \perp_C \alpha)$, and due to Definition 5.2.7, there is at least one set X such that $B \dot{-}_{C,\gamma} \alpha \subseteq X \in \gamma(B \perp_C \alpha)$. Since $\alpha \notin C(X)$ and C satisfies monotony, we can conclude from $B \dot{-}_{C,\gamma} \alpha \subseteq X$ that $\alpha \notin C(B \dot{-}_{C,\gamma} \alpha)$.

Relevance: Let $\beta \in B \setminus (B \dot{-}_{C,\gamma} \alpha)$. Then, according to Definition 5.2.7, there is some B' such that $\beta \notin B' \in \gamma(B \perp_C \alpha)$. It follows that $B \dot{-}_{C,\gamma} \alpha \subseteq B' \subseteq B$, $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\beta\})$.

Uniformity: Suppose that for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$. It is easy to see that $B \perp_C \alpha = B \perp_C \beta$. Since γ is a function, it follows that $\gamma(B \perp_C \alpha) = \gamma(B \perp_C \beta)$ and by the definition of partial meet contraction, $B \dot{-}_{C,\gamma} \alpha = B \dot{-}_{C,\gamma} \beta$.

(ii) postulates ⇒ construction: Let $\dot{-}$ be an operation that satisfies the four postulates, and let B be a set of sentences. Let γ be a function such that for all sentences α: $\gamma(B\perp_C\alpha) = \{X \in B\perp_C\alpha | B\dot{-}\alpha \subseteq X\}$ (that γ is a function for the given domain follows from *uniformity*). We have to prove (1) that γ is a selection function for the given domain in the sense of the Definition 5.2.5 and (2) that $B\dot{-}\alpha = \bigcap\gamma(B\perp_C\alpha)$.

(1) We can do this by showing that $\gamma(B\perp_C\alpha)$ is non-empty whenever $B\perp_C\alpha \neq \emptyset$. Suppose $B\perp_C\alpha \neq \emptyset$. It follows form monotony that $\alpha \notin C(\emptyset)$, hence from *success* that $\alpha \notin C(B\dot{-}\alpha)$. It follows from *inclusion* and the upper bound property that there is some X such that $B\dot{-}\alpha \subseteq X \in B\perp_C\alpha$. It follows from our definition of γ that $X \in \gamma(B\perp_C\alpha)$, so that $\gamma(B\perp_C\alpha)$ is non-empty.

(2) To prove that $B\dot{-}\alpha = \bigcap\gamma(B\perp_C\alpha)$, we first observe that by the definition of γ, $B\dot{-}\alpha \subseteq \bigcap\gamma(B\perp_C\alpha)$. It remains to be shown that $\bigcap\gamma(B\perp_C\alpha) \subseteq B\dot{-}\alpha$. In order to do this, we will let $\beta \notin B\dot{-}\alpha$. We have to prove that $\beta \notin \bigcap\gamma(B\perp_C\alpha)$. Since this is trivial when $\beta \notin B$, we may assume that $\beta \in B$.

From $\beta \in B \setminus B\dot{-}\alpha$ it follows by *relevance* that there is some B' such that $B\dot{-}\alpha \subseteq B' \subseteq B$, $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\beta\})$. It follows from the upper bound property that there is some X such that $B' \subseteq X \in B\perp_C\alpha$ and from monotony that $\beta \notin X$. Then by the definition, $X \in \gamma(B\perp_C\alpha)$, from which follows that $\beta \notin \bigcap\gamma(B\perp_C\alpha)$. □

Theorem 5.2.10: Let C be an inference operation satisfying monotony, compactness, and $\perp \notin C(\emptyset)$. An operation ! is an operation of kernel consolidation for B determined by C and some incision function if and only if:

- $\perp \notin C(B!)$ (consistency)
- $B! \subseteq B$ (inclusion)
- If $\alpha \in B \setminus (B!)$, then there is some X such that $X \subseteq B$, $\perp \notin C(X)$ and $\perp \in C(X \cup \{\alpha\})$ (core-retainment)

Proof: (i) construction $\Rightarrow$ postulates: Let $!_{C,\sigma}$ be a kernel consolidation operator based on an operation C that satisfies the conditions given and an incision function σ. We need to show that it satisfies the three postulates.

Inclusion follows directly from the definition of kernel consolidation.

Consistency: Suppose to the contrary that $\perp \in C(B!_{C,\sigma})$. Then according to the compactness of C there is some finite subset X of $B!_{C,\sigma}$ such that $\perp \in C(X)$, hence there is some inclusion-minimal subset Z of X such that $\perp \in C(Z)$, i.e., $Z \in B \perp\!\!\!\perp_C \perp$. Due to $\perp \notin C(\emptyset)$, $Z \neq \emptyset$. Therefore, according to Definitions 5.2.1 and 5.2.9, Z cannot be a subset of $B!_{C,\sigma}$. This contradiction shows that consistency holds.

Core-retainment: Let $\alpha \in B$ and $\alpha \notin B!_{C,\sigma}$. Then $\alpha \in \sigma(B \perp\!\!\!\perp_C \perp)$. According to Definition 5.2.1, $\sigma(B \perp\!\!\!\perp_C \perp) \subseteq \bigcup(B \perp\!\!\!\perp_C \perp)$, hence there is some set D such that $\alpha \in D \in B \perp\!\!\!\perp_C \perp$. Let $X = D \setminus \{\alpha\}$. Then $X \subseteq B$, $\perp \notin C(X)$ and $\perp \in C(X \cup \{\alpha\})$, which shows that core-retainment is satisfied.

(ii) postulates $\Rightarrow$ construction: Let ! be an operation that satisfies the three postulates of the theorem. We are going to show that ! is a kernel consolidation based on C and some incision function σ. For that purpose, let σ be such that: $\sigma(B \perp\!\!\!\perp_C \bot) = B \setminus (B!)$.

We need to verify (1) that σ is an incision function (for the domain covered by the definition), and (2) that the kernel contraction based on C and σ coincides with !.

(1) Clearly, σ is a function. We need to show that it satisfies conditions 1 and 2 of Definition 5.2.1.

For condition 1, we are going to show that $\sigma(B \perp\!\!\!\perp_C \bot) \subseteq \bigcup(B \perp\!\!\!\perp_C \bot)$. Let $\alpha \in \sigma(B \perp\!\!\!\perp_C \bot)$. It follows from *core-retainment* that there is some X such that $X \subseteq B$, $\bot \notin C(X)$ and $\bot \in C(X \cup \{\alpha\})$. Let X' be an inclusion-minimal subset of X such that $\bot \in C(X' \cup \{\alpha\})$. According to monotony, since $X' \subseteq X$ and $\bot \notin C(X)$ we also have $\bot \notin C(X')$. It follows that $X' \cup \{\alpha\} \in B \perp\!\!\!\perp_C \bot$, hence $\alpha \in \bigcup(B \perp\!\!\!\perp_C \bot)$.

For condition 2, let $\emptyset \neq X \in B \perp\!\!\!\perp_C \bot$. We need to show that $X \cap \sigma(B \perp\!\!\!\perp_C \bot)$ is non-empty. By *consistency*, $\bot \notin C(B!)$. Since $\bot \in C(X)$ we may conclude from the monotony of C that $X \not\subseteq B!$, i.e., that there is some ε such that $\varepsilon \in X$ and $\varepsilon \notin B!$. Since $X \subseteq B$ it follows that $\varepsilon \in B \setminus (B!)$, i.e., $\varepsilon \in \sigma(B \perp\!\!\!\perp_C \bot)$. Thus, $\varepsilon \in X \cap \sigma(B \perp\!\!\!\perp_C \bot)$ which is sufficient to show that condition 2 is satisfied.

(2) It follows from *inclusion* and our definition $\sigma(B \perp\!\!\!\perp_C \bot) = B \setminus (B!)$ that $B! = B \setminus \sigma(B \perp\!\!\!\perp_C \bot)$. □

Theorem 5.2.12: Let C satisfy monotony, compactness, and $\bot \notin C(\emptyset)$. An operation ! is an operation of partial meet consolidation based on C and some selection function if and only if for all sets B of sentences:

- $\bot \notin C(B!)$ (consistency)
- $B! \subseteq B$ (inclusion)
- If $\alpha \in B \setminus (B!)$, then there is some X such that $B! \subseteq X \subseteq B$, $\bot \notin C(X)$ and $\bot \in C(X \cup \{\alpha\})$ (relevance)

Proof: (i) construction $\Rightarrow$ postulates: Let $!_{C,\gamma}$ be a partial meet consolidation operator based on an operation C that satisfies the conditions given and a selection function γ. We need to show that it satisfies the three postulates.

Inclusion follows directly from the Definitions 5.2.4 and 5.2.11.

Consistency: Due to $\bot \notin C(\emptyset)$, monotony, compactness, and the upper bound property, $B \bot_C \bot$ is non-empty (Observation 5.2.6). Due to Definition 5.2.5, so is $\gamma(B \bot_C \bot)$, and due to Definition 5.2.11, there is some set X such that $B!_{C,\gamma} \subseteq X \in B \bot_C \bot$. Since $\bot \notin C(X)$ and C satisfies monotony, we can conclude from $B!_{C,\gamma} \subseteq X$ that $\bot \notin C(B!_{C,\gamma})$.

Relevance: Let $\alpha \in B \setminus (B!_{C,\gamma})$. Then, according to Definition 5.2.11, there is some X such that $\alpha \notin X \in \gamma(B\perp_C\perp)$. It follows that $B!_{C,\gamma} \subseteq X \subseteq B$, $\perp \notin C(X)$ and $\perp \in C(X \cup \{\alpha\})$.

(ii) postulates $\Rightarrow$ construction: Let ! be an operation that satisfies the three postulates, and let B be a set of sentences. Let γ be a function such that: $\gamma(B\perp_C\perp) = \{X \in B\perp_C\perp | B! \subseteq X\}$. We have to prove (1) that γ is a selection function for the given domain in the sense of Definition 5.2.5 and (2) that $B! = \bigcap \gamma(B\perp_C\perp)$.

(1) We can do this by showing that $\gamma(B\perp_C\perp)$ is non-empty whenever $B\perp_C\perp$ is non-empty. It follows from *consistency* and *inclusion*, using the upper bound property, that there is some X such that $B! \subseteq X \in B\perp_C\perp$. It follows from the definition of γ that $X \in \gamma(B\perp_C\perp)$, so that $\gamma(B\perp_C\perp)$ is non-empty.

(2) To prove that $B! = \bigcap \gamma(B\perp_C\perp)$, we first observe that by the definition of γ, $B! \subseteq \bigcap \gamma(B\perp_C\perp)$. It remains to be shown that $\bigcap \gamma(B\perp_C\perp) \subseteq B!$. In order to do this, we will assume that $\beta \notin B!$ and prove that $\beta \notin \bigcap \gamma(B\perp_C\perp)$. Since this is trivial whenever $\beta \notin B$, only the case when $\beta \in B \setminus B!$ needs to be proved.

From $\beta \in B \setminus B!$ follows by *relevance* that there is some X such that $B! \subseteq X \subseteq B$, $\perp \notin C(X)$ and $\perp \in C(X \cup \{\beta\})$. It follows from the upper bound property that there is some X' such that $X \subseteq X' \in B\perp_C\perp$ and from monotony that $\beta \notin X'$. Then by the definition, $X' \in \gamma(B\perp_C\perp)$, from which follows that $\beta \notin \bigcap \gamma(B\perp_C\perp)$. $\square$

Theorem 5.2.14: (i) construction $\Rightarrow$ postulates: Let C satisfy monotony and compactness. An operator $\mp$ is an operator of internal kernel revision based on an inference operator C if and only if, for all sets B of sentences and all sentences α such that C satisfies α-local non-contravention:

- If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \mp \alpha)$ (non-contradiction)
- $B \mp \alpha \subseteq B \cup \{\alpha\}$ (inclusion)
- If $\beta \in B \setminus B \mp \alpha$, then there is some $B' \subseteq B$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$ (core-retainment)
- $\alpha \in B \mp \alpha$ (success)
- If for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$ (uniformity)

Proof: Let $\mp_{C,\sigma}$ be an internal kernel revision operator based on an inference C and an incision function σ. We have to show that it satisfies the five postulates.

Inclusion and *success* follow trivially from the construction.

Non-contradiction: Let $\neg\alpha \notin C(\emptyset)$. Then it follows from the proof of success for contraction in Theorem 5.2.3 that $\neg\alpha \notin C(B \setminus \sigma(B \perp\!\!\!\perp_C \neg\alpha))$. From this and α-local non-contravention follows $\neg\alpha \notin C(B \mp_{C,\sigma} \alpha)$.

Core-retainment: If $\neg\alpha \in C(\emptyset)$, then $B \perp\!\!\!\perp_C \neg\alpha = \emptyset$, $B \subseteq B \mp_{C,\sigma} \alpha$ and the postulate holds trivially. Otherwise, let $\beta \in B \setminus (B \mp_{C,\sigma} \alpha)$. By the construction of $B \mp_{C,\sigma} \alpha$, there must be $X \in B \perp\!\!\!\perp_C \neg\alpha$ such that $\beta \in X$. Let $B' = X \setminus \{\beta\}$. Then $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$.

Uniformity: Suppose that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. Then $B \perp\!\!\!\perp_C \neg\alpha = B \perp\!\!\!\perp_C \neg\beta$ and it follows that $\sigma(B \mp_{C,\sigma} \neg\alpha) = \sigma(B \mp_{C,\sigma} \neg\beta)$ and thus $B \cap (B \mp_{C,\sigma} \alpha) = B \cap (B \mp_{C,\sigma} \beta)$.

(ii) postulates $\Rightarrow$ construction:

Let $\mp$ be an operator that satisfies the given conditions. Let B be a set of sentences and let σ be a function such that for every sentence α, $\sigma(B \perp\!\!\!\perp_C \neg\alpha) = B \setminus (B \cap (B \mp \alpha))$. We have to prove that (a) σ is a function, (b) σ is an incision function, and (c) $\mp = \mp_{C,\sigma}$.

(a) Let α and β be such that $B \perp\!\!\!\perp_C \neg\alpha = B \perp\!\!\!\perp_C \neg\beta$. It follows from monotony that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. By *uniformity*, $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$. Thus, $\sigma(B \perp\!\!\!\perp_C \neg\alpha) = \sigma(B \perp\!\!\!\perp_C \neg\beta)$.

(b) Let $\delta \in \sigma(B \perp\!\!\!\perp_C \neg\alpha)$. Then, $\delta \in B \setminus (B \cap (B \mp \alpha)) = B \setminus (B \mp \alpha)$. By *core-retainment*, there is $B' \subseteq B$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\delta\})$. Then there is some inclusion-minimal subset Y of B' such that $\neg\alpha \in C(Y \cup \{\delta\})$. Let $X = Y \cup \{\delta\}$. It then follows from the non-emptiness of $\sigma(B \perp\!\!\!\perp_C \neg\alpha)$ that $\neg\alpha \notin C(\emptyset)$ and that no proper subset of X C-implies $\neg\alpha$. Hence, $\delta \in X \in B \perp\!\!\!\perp_C \neg\alpha$ and so $\delta \in \bigcup(B \perp\!\!\!\perp_C \neg\alpha)$.

Next, let $\emptyset \neq X \in B \perp\!\!\!\perp_C \neg\alpha$. Then $\neg\alpha \notin C(\emptyset)$, and $\neg\alpha \in C(X)$. Suppose that $X \subseteq B \mp \alpha$. Then it follows from monotony that $\neg\alpha \in C(B \mp \alpha)$, contrary to *non-contradiction*. We may conclude that $X \not\subseteq B \mp \alpha$. Then, by *inclusion*, $X \not\subseteq B \cap B \mp \alpha$. Hence, there is some $\delta \in X$ such that $\delta \in B \setminus (B \cap B \mp \alpha) = \sigma(B \perp\!\!\!\perp_C \neg\alpha)$ and we have $X \cap \sigma(B \perp\!\!\!\perp_C \neg\alpha) \neq \emptyset$.

(c) $B \mp_{C,\sigma} \alpha = (B \setminus \sigma(B \perp\!\!\!\perp_C \neg\alpha)) \cup \{\alpha\} = (B \setminus (B \setminus B \mp \alpha)) \cup \{\alpha\} =$ (by *inclusion*) $(B \mp \alpha) \cup \{\alpha\} =$ (by *success*) $B \mp \alpha$. □

Theorem 5.2.16: Let C be an inference operator satisfying monotony and compactness. An operator $\mp$ is an operator of internal partial meet revision based on C if and only if, for all sets B of sentences and all sentences α such that C satisfies α-local non-contravention:

- If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \mp \alpha)$ (non-contradiction)
- $B \mp \alpha \subseteq B \cup \{\alpha\}$ (inclusion)
- If $\beta \in B \setminus B \mp \alpha$, then there is some B' such that $B \mp \alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(B')$ but $\neg\alpha \in C(B \cup \{\beta\})$ (relevance)
- $\alpha \in B \mp \alpha$ (success)
- If for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \mp \alpha) = B \cap (B \mp \beta)$ (uniformity)

Proof: (i) construction $\Rightarrow$ postulates:

Let $\mp_{C,\gamma}$ be an internal partial meet revision operator based on an inference C and a selection function γ. We have to show that it satisfies the five postulates.

That *inclusion* and *success* hold follows directly from the construction of $\mp_{C,\gamma}$.

Non-contradiction: There are two cases: (1) $B\bot_C\neg\alpha = \emptyset$. Then $\neg\alpha \in C(\emptyset)$, so that the postulate holds vacuously. (2) $B\bot_C\neg\alpha \neq \emptyset$. Then $\gamma(B\bot_C\neg\alpha) \neq \emptyset$, and it follows by monotony that $\neg\alpha \notin C(\bigcap\gamma(B\bot_C\neg\alpha))$. It follows by α-local non-contravention that $\neg\alpha \notin C(\bigcap\gamma(B\bot_C\neg\alpha) \cup \{\alpha\}) = B \mp_{C,\gamma} \alpha$.

Relevance: Let $\beta \in B \setminus B \mp_{C,\gamma} \alpha$. Then $\beta \notin \bigcap\gamma(B\bot_C\neg\alpha)$, and there is a set X such that $X \in \gamma(B\bot_C\neg\alpha)$ and $\beta \notin X$. Let $B' = X \cup \{\alpha\}$. Then $B \mp_{C,\gamma} \alpha \subseteq B' \subseteq B \cup \{\alpha\}$. It follows from $\neg\alpha \notin C(X)$, due to non-contravention, that $\neg\alpha \notin C(X \cup \{\alpha\})$, i.e., $\neg\alpha \notin C(B')$. Furthermore, it follows from $\neg\alpha \in C(X \cup \{\beta\})$ by monotony that $\neg\alpha \in C(X \cup \{\alpha\} \cup \{\beta\})$, i.e., $\neg\alpha \in C(B' \cup \{\beta\})$.

Uniformity: Suppose that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. Then, $B\bot_C\neg\alpha = B\bot_C\neg\beta$ and $\gamma(B\bot_C\neg\alpha) = \gamma(B\bot_C\neg\beta)$, which means that $B \cap (B \mp_{C,\gamma} \alpha) = B \cap (B \mp_{C,\gamma} \beta)$

(ii) postulates $\Rightarrow$ construction:

Let $\mp$ be an operator satisfying the postulates above. Let B be a set of sentences and let γ be a function such that for every sentence α, $\gamma(B\bot_C\alpha) = \{X \in B\bot_C\alpha | B \cap (B \mp \neg\alpha) \subseteq X\}$.

We have to show that: (a) γ is a function for the given domain, (b) γ is a selection function for the given domain, and (c) $B \mp_{C,\gamma} \alpha = B \mp \alpha$.

(a) Let α and β be such that $B\bot_C\alpha = B\bot_C\beta$. Then, due to monotony, for all subsets B' of B, $\alpha \in C(B')$ if and only if $\beta \in C(B')$. By *uniformity*, $B \cap (B \mp \neg\alpha) = B \cap (B \mp \neg\beta)$ and hence, $\gamma(B\bot_C\alpha) = \gamma(B\bot_C\beta)$.

(b) It follows directly from our definition of γ that it satisfies condition 1 of Definition 5.2.5. To see that it satisfies condition 2, let $B\bot_C\alpha \neq \emptyset$. Then $\alpha \notin C(\emptyset)$. It follows from *non-contradiction* that $\alpha \notin C(B \mp \neg\alpha)$. It follows from monotony that $\alpha \notin C(B \cap (B \mp \neg\alpha))$. Thus, by the upper bound property, there is an $X \in B\bot_C\alpha$ such that $B \cap (B \mp \neg\alpha) \subseteq X$. By the definition of γ, this means that $\gamma(B\bot_C\alpha)$ is nonempty.

(c) By the definition of γ and *inclusion*, $B \mp \alpha \subseteq (B \cap B \mp \alpha) \cup \{\alpha\} \subseteq \bigcap\gamma(B\bot_C\neg\alpha) \cup \{\alpha\} = B \mp_{C,\gamma} \alpha$. For the other side of the inclusion, let $\beta \notin B \mp \alpha$. If $\beta \notin B$, then $\beta \notin B \mp_{C,\gamma} \alpha$ follows directly. We can assume therefore that $\beta \in B$. It follows from *relevance* that there is some B' such that $B \mp \alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(B')$, and $\neg\alpha \in C(B' \cup \{\beta\})$. Since $\alpha \in B'$, it follows by α-local non-contravention from $\neg\alpha \notin C(B')$ that $\neg\alpha \notin C(B' \setminus \{\alpha\})$. It follows from Observation 5.2.6 that there is some $X \in B\bot_C\neg\alpha$ such that $(B' \setminus \{\alpha\}) \subseteq X$. Hence, $B \cap (B \mp \alpha) \subseteq (B' \setminus \{\alpha\}) \subseteq X$, from which follows that $X \in \gamma(B\bot_C\neg\alpha)$.

It follows from $\neg\alpha \in C(B' \cup \{\beta\})$ and non-contravention that $\neg\alpha \in C((B' \setminus \{\alpha\}) \cup \{\beta\})$, hence by monotony $\neg\alpha \in C(X \cup \{\beta\})$. Since $\neg\alpha \notin C(X)$, it follows that $\beta \notin X$. Hence, $\beta \notin \bigcap\gamma(B\bot_C\neg\alpha)$. It follows that $\beta \notin B \mp_{C,\gamma} \alpha$. $\square$

Theorem 5.2.18: Let C be an inference operator satisfying monotony and compactness. An operator $\pm$ is an operator of external kernel revision based on an inference operator C if and only if, for all sets B of sentences and all sentences α such that C satisfies α-local non-contravention:

- If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \pm \alpha)$ (non-contradiction)
- $B \pm \alpha \subseteq B \cup \{\alpha\}$ (inclusion)
- If $\beta \in B \setminus B \pm \alpha$, then there is some $B' \subseteq B \cup \{\alpha\}$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$ (core-retainment)
- $\alpha \in B \pm \alpha$ (success)
- If α and β are elements of B and it holds for all $B' \subseteq B$ that $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ (weak uniformity)
- $B + \alpha \pm \alpha = B \pm \alpha$ (pre-expansion)

Proof:(i) construction $\Rightarrow$ postulates:

Let $\pm_{C,\sigma}$ be an external kernel revision operator based on an inference C and an incision function σ. We have to show that it satisfies the six postulates.

Inclusion and *pre-expansion* follow directly from the construction of $\pm_{C,\sigma}$.

Non-contradiction: Let $\neg\alpha \notin C(\emptyset)$. Then it follows from the proof of success for contraction in Theorem 5.2.3 that $\neg\alpha \notin C((B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha))$, i.e., $\neg\alpha \notin C(B \pm_{C,\sigma} \alpha)$.

Core-retainment: Let $\beta \in B \setminus B \pm_{C,\sigma} \alpha$. By the construction of $\pm_{C,\sigma}$, we have $\beta \in \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha)$. Hence there is some X such that $\beta \in X \in ((B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha)$. Let $B' = X \setminus \{\beta\}$. Then $B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$.

Success: It follows from α-local non-contravention that for all $X \in (B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha$, $\alpha \notin X$. Thus, $\alpha \notin \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \neg\alpha)$, and hence $\alpha \in B \pm_{C,\sigma} \alpha$.

Weak-uniformity: Let α and β be elements of B such that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. Then $B \perp\!\!\!\perp_C \neg\alpha = B \perp\!\!\!\perp_C \neg\beta$ and hence, $\sigma(B \perp\!\!\!\perp_C \neg\alpha) = \sigma(B \perp\!\!\!\perp_C \neg\beta)$. Since $\alpha, \beta \in B$, this implies that $B \pm_{C,\sigma} \alpha = B \pm_{C,\sigma} \beta$.

(ii) postulates $\Rightarrow$ construction:

Let $\pm$ be an operator that satisfies the given conditions. Let σ be such that for all sets B and all sentences $\alpha \in B$, $\sigma(B \perp\!\!\!\perp_C \neg\alpha) = B \setminus (B \pm \alpha)$. We have to prove that (a) σ is a function for this domain, (b) σ is an incision function for this domain, and (c) $\pm = \pm_{C,\sigma}$.

(a) Let α and β be elements of B such that $B \perp\!\!\!\perp_C \neg\alpha = B \perp\!\!\!\perp_C \neg\beta$. It follows that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. By *weak uniformity*, $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$. Thus, $\sigma(B \perp\!\!\!\perp_C \neg\alpha) = \sigma(B \perp\!\!\!\perp_C \neg\beta)$.

(b) Let $\delta \in \sigma(B \perp\!\!\!\perp_C \neg\alpha)$. Then, $\delta \in B \setminus (B \pm \alpha)$. By *core-retainment*, there is some $B' \subseteq B$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\delta\})$. Then due to monotony, there is a set $X \subseteq B \cup \{\delta\}$ such that $\delta \in X \in B \perp\!\!\!\perp_C \neg\alpha$ and so, $\delta \in \bigcup(B \perp\!\!\!\perp_C \neg\alpha)$.

Now, let $\emptyset \neq X \in B \perp\!\!\!\perp_C \neg\alpha$ for some $B \perp\!\!\!\perp_C \neg\alpha$ that is in the domain of σ as defined above. Then $\neg\alpha \notin C(\emptyset)$. Suppose that $X \subseteq B \pm \alpha$. Then it follows from monotony that $\neg\alpha \in C(B \pm \alpha)$, contrary to *non-contradiction*. Hence, $X \not\subseteq B \pm \alpha$. Since $\alpha \in B$, *inclusion* yields $B \pm \alpha \subseteq B$, hence $B \setminus (B \pm \alpha) \cap X = \sigma(B \perp\!\!\!\perp_C \neg\alpha) \cap X \neq \emptyset$.

(c) $B \pm_{C,\sigma} \alpha = (B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp \neg\alpha) =$ (by the definition of σ above) $(B + \alpha) \setminus ((B + \alpha) \setminus ((B + \alpha) \pm \alpha)) =$ (by *inclusion*) $(B + \alpha) \pm \alpha =$ (by *pre-expansion*) $B \pm \alpha$. □

Theorem 5.2.20: Let C be an inference operator satisfying monotony and compactness. An operator $\pm$ is an operator of external partial meet revision based on an inference operator C if and only if, for all sets B of sentences and sentences α such that C satisfies α-local non-contravention:

- If $\neg\alpha \notin C(\emptyset)$, then $\neg\alpha \notin C(B \pm \alpha)$ (non-contradiction)
- $B \pm \alpha \subseteq B \cup \{\alpha\}$ (inclusion)
- If $\beta \in B \setminus B \pm \alpha$, then there is some B' such that $B \pm \alpha \subseteq B' \subseteq B \cup \{\alpha\}$ such that $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$ (relevance)
- $\alpha \in B \pm \alpha$ (success)
- If α and β are elements of B and it holds for all $B' \subseteq B$ that $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$, then $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ (weak uniformity)
- $B + \alpha \pm \alpha = B \pm \alpha$ (pre-expansion)

Proof(i) construction $\Rightarrow$ postulates:

Let $\pm_{C,\gamma}$ be an external partial meet revision operator based on an inference C and a selection function γ. We have to show that it satisfies the six postulates.

Inclusion and *pre-expansion* follow directly from the construction of $\pm_{C,\gamma}$.

Non-contradiction: Let $\neg\alpha \notin C(\emptyset)$. Then $(B \cup \{\alpha\})\perp_C\neg\alpha \neq \emptyset$. It follows from Definition 5.2.15 that $B \pm_{C,\gamma} \alpha = \bigcap\gamma((B \cup \{\alpha\})\perp_C\neg\alpha)$. It follows from $\neg\alpha \notin C(\emptyset)$ that for all $X \in (B \cup \{\alpha\})\perp_C\neg\alpha$, $\neg\alpha \notin C(X)$, and from monotony that $\neg\alpha \notin B \pm_{C,\gamma} \alpha$.

Relevance: Let $\beta \in B \setminus B \pm_{C,\gamma} \alpha$. By the construction of $\pm_{C,\gamma}$ there is some $X \in \gamma(B \cup \{\alpha\}\perp_C\neg\alpha)$ such that $\beta \notin X$. Then $B \pm_{C,\gamma} \alpha \subseteq X \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(X)$ and $\neg\alpha \in C(X \cup \{\beta\})$.

Success: If $B \cup \{\alpha\}\perp_C\neg\alpha = \emptyset$, then success follows directly from the definition. Otherwise it follows from α-local non-contravention that each element of $(B \cup \{\alpha\})\perp_C\neg\alpha$ contains α, and then so does $B \pm_{C,\gamma} \alpha$.

Weak-uniformity: Let α and β be elements of B such that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. Then $B\perp_C\neg\alpha = B\perp_C\neg\beta$ and hence, $\gamma(B\perp_C\neg\alpha) = \gamma(B\perp_C\neg\beta)$. Since $\alpha, \beta \in B$, this implies that $B \pm_{C,\gamma} \alpha = B \pm_{C,\gamma} \beta$.

(ii) postulates $\Rightarrow$ construction:

Let $\pm$ be an operator that satisfies the given conditions. Let γ be such that for every set of sentences B if $\alpha \in B$, then $\gamma(B\perp_C\neg\alpha) = \{X \in B\perp_C\neg\alpha | B \pm \alpha \subseteq X\}$. We have to prove that (a) γ is a function in the given domain, (b) γ is a selection function in the given domain, and (c) $\pm = \pm_{C,\gamma}$.

(a) Let α and β be elements of B such that $B\perp_C\neg\alpha = B\perp_C\neg\beta$. It follows from monotony that for all $B' \subseteq B$, $\neg\alpha \in C(B')$ if and only if $\neg\beta \in C(B')$. By *weak uniformity*, $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$. By *inclusion*, $B \pm \alpha \subseteq B \cup \{\alpha\} = B$ and $B \pm \beta \subseteq B \cup \{\beta\} = B$. Hence it follows from $B \cap (B \pm \alpha) = B \cap (B \pm \beta)$ that $B \pm \alpha = B \pm \beta$. By the construction of γ, $\gamma(B\perp_C\neg\alpha) = \gamma(B\perp_C\neg\beta)$.

(b) Let $B\perp_C\neg\alpha \neq \emptyset$. It follows from monotony that $\neg\alpha \notin C(\emptyset)$. *Non-contradiction* yields $\neg\alpha \notin C(B \pm \alpha)$. Due to Observation 5.2.6 there is $X \in B\perp_C\neg\alpha$ such that $B \pm \alpha \subseteq X$. By the definition of γ, $X \in \gamma(B\perp_C\neg\alpha)$.

(c) There are two cases to be considered. First case, if $(B \cup \{\alpha\})\perp_C\neg\alpha = \emptyset$, then it follows directly from Definition 5.2.19 that $B \pm_{C,\gamma} \alpha = B + \alpha$. *Inclusion* yields $B \pm \alpha \subseteq B + \alpha$, and *success* yields $\alpha \in B \pm \alpha$. It remains to show that $B \subseteq B \pm \alpha$. Suppose to the contrary that this is not the case. Then there is some $\beta \in B \setminus B \pm \alpha$, and due to *relevance*, there is some B' such that $B \pm \alpha \subseteq B' \subseteq B + \alpha$, $\neg\alpha \notin C(B')$ and $\neg\alpha \in C(B' \cup \{\beta\})$. However, it follows from $(B \cup \{\alpha\})\perp_C\neg\alpha = \emptyset$, due to Observation 5.2.6, that $\neg\alpha \in C(\emptyset)$, hence by monotony $\neg\alpha \in C(B')$. This contradiction concludes the proof. Second case, if $(B \cup \{\alpha\})\perp_C\neg\alpha \neq \emptyset$, then, by Definition 5.2.19, $B \pm_{C,\gamma} \alpha = \bigcap\gamma((B \cup \{\alpha\})\perp_C\neg\alpha)$ $= \bigcap\{X \in (B \cup \{\alpha\})\perp_C\neg\alpha | (B \cup \{\alpha\}) \pm \alpha \subseteq X\}$ = (by *pre-expansion*) $\bigcap\{X \in (B \cup \{\alpha\})\perp_C\neg\alpha | B \pm \alpha \subseteq X\}$. It follows set-theoretically that $B \pm \alpha \subseteq B \pm_{C,\gamma} \alpha$.

For the other direction, let $\beta \in B \pm_{C,\gamma} \alpha \setminus B \pm \alpha$. Then due to *pre-expansion*, $\beta \in B \cup \{\alpha\} \setminus (B \cup \{\alpha\}) \pm \alpha$. Due to *relevance*, there is some B' such that $(B \cup \{\alpha\}) \pm \alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\neg\alpha \notin C(B')$, and $\neg\alpha \in C(B' \cup \{\beta\})$. Hence, due to Observation 5.2.6, there is some $X \in \gamma((B \cup \{\alpha\})\perp_C\neg\alpha)$ such that $\beta \notin B' \subseteq X$ and so, $\beta \notin \bigcap\gamma((B \cup \{\alpha\})\perp_C\neg\alpha) = B \pm_{C,\gamma} \alpha$. □

Theorem 5.2.22: Let C be an inference operation satisfying monotony, compactness, and $\perp \notin C(\emptyset)$. An operator ? is an operator of kernel semi-revision based on C if and only if for all sets B of sentences and sentences α:

- $\perp \notin C(B?\alpha)$ (consistency)
- $B?\alpha \subseteq B \cup \{\alpha\}$ (inclusion)
- If $\beta \in B \setminus B?\alpha$, then there is some $B' \subseteq B \cup \{\alpha\}$ such that $\perp \notin C(B')$ and $\perp \in C(B' \cup \{\beta\})$ (core-retainment)

- $(B+\alpha)?\alpha = B?\alpha$ (pre-expansion)

- If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)

Proof: (i) construction $\Rightarrow$ postulates: Let $?_{C,\sigma}$ be an operator of kernel semi-revision based on an inference operator C and an incision function σ. It follows directly from the construction that *inclusion*, *pre-expansion*, and *internal exchange* are satisfied. Due to monotony, so is *consistency*. Finally, for *core-retainment*, let $\beta \in B \setminus B?_{C,\sigma}\alpha$. Then by construction $\beta \in \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \bot)$. This means that for some set $X \in (B \cup \{\alpha\}) \perp\!\!\!\perp_C \bot$, $\beta \in X$. Let $B' = X \setminus \{\beta\}$. We have $B' \subseteq B \cup \{\alpha\}$, $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\beta\})$.

(ii) postulates $\Rightarrow$ construction: Let ? be an operator satisfying the postulates above and let σ be such that for every set of sentences B:

$\sigma(B \perp\!\!\!\perp_C \bot) = B \setminus \{\beta | \beta \in B?\alpha \text{ for some } \alpha \in B\}$

We have to show (1) that σ is an incision function for the given domain and (2) that $B?\alpha = B?_{C,\sigma}\alpha$.

(1) To show that σ is an incision function we have to show that conditions 1 and 2 of Definition 5.2.1 are satisfied. For condition 1, let $\delta \in \sigma(B \perp\!\!\!\perp_C \bot)$. Then it holds for all $\alpha \in B$ that $\delta \notin B?\alpha$ hence, $\delta \notin B?\delta$. It follows from *core-retainment* that there is some $B' \subseteq B$ such that $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\delta\})$. It follows that there is some subset B'' of B' such that $B'' \cup \{\delta\} \in B \perp\!\!\!\perp_C \bot$. For condition 2, let $\emptyset \neq X \in B \perp\!\!\!\perp_C \bot$. Suppose that $X \cap \sigma(B \perp\!\!\!\perp_C \bot) = \emptyset$. Then $X \subseteq \{\beta | \beta \in B?\alpha$ for some $\alpha \in B\}$. Let α be an element of B such that $\bot \notin C(\{\alpha\})$. It follows by *internal exchange* that $X \subseteq B?\alpha$. Since $\bot \in C(X)$, it follows from monotony that $\bot \in C(B?\alpha)$, contrary to *consistency*. This contradiction is sufficient to prove that $X \cap \sigma(B \perp\!\!\!\perp_C \bot) \neq \emptyset$.

(2) By definition, $\sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \bot) = (B \cup \{\alpha\}) \setminus \{\beta | \beta \in B \cup \{\alpha\}?\varepsilon$ for some $\varepsilon \in B \cup \{\alpha\}\} = (B \cup \{\alpha\}) \setminus (B \cup \{\alpha\})?\alpha$ (*internal exchange*) $= (B \cup \{\alpha\}) \setminus B?\alpha$ (*pre-expansion*). Hence, $B?_{C,\sigma}\alpha = (B \cup \{\alpha\}) \setminus \sigma((B \cup \{\alpha\}) \perp\!\!\!\perp_C \bot)$ (Definition 5.2.21) $= (B \cup \{\alpha\}) \setminus ((B \cup \{\alpha\}) \setminus B?\alpha) = B?\alpha$ (*inclusion*). $\Box$

Theorem 5.2.24: Let C be an inference operation satisfying monotony, compactness, and $\bot \notin C(\emptyset)$. An operator ? is an operator of partial meet semi-revision based on C if and only if for all sets B of sentences and sentences α:

- $\bot \notin C(B?\alpha)$ (consistency)

- $B?\alpha \subseteq B \cup \{\alpha\}$ (inclusion)

- If $\beta \in B \setminus B?\alpha$, then there is some B' such that $B?\alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\beta\})$ (relevance)

- $(B+\alpha)?\alpha = B?\alpha$ (pre-expansion)

- If $\alpha, \beta \in B$, then $B?\alpha = B?\beta$ (internal exchange)

Proof: (i) construction $\Rightarrow$ postulates: Let $?_{C,\gamma}$ be a partial meet revision operator based on an inference operator C and a selection function γ. It follows directly from the construction that *inclusion*, *pre-expansion*, and *internal exchange* are satisfied. Due to monotony, so is *consistency*. It remains to show that $?_{C,\gamma}$ satisfies *relevance*. Let $\beta \in B \setminus B?_{C,\gamma}\alpha$. Then by construction there is an $X \in \gamma((B \cup \{\alpha\})\bot_C\bot)$ such that $\beta \notin X$. Since $\beta \in B \cup \{\alpha\}$, it follows that $B?_{C,\gamma}\alpha \subseteq X \subseteq B \cup \{\alpha\}$, $\bot \notin C(X)$ and $\bot \in C(X \cup \{\beta\})$.

(ii) postulates $\Rightarrow$ construction: Let $?$ be an operator satisfying the postulates above and let γ be defined as follows for every set of sentences B:

$\gamma(B\bot_C\bot) = \{X \in B\bot_C\bot | B?\alpha \subseteq X \text{ for some } \alpha \in A\}$

We have to show that (1) γ is a selection function in the given domain, and (2) $B?\alpha = B?_{C,\gamma}\alpha$.

(1) Due to Definition 5.2.5 it is sufficient to show that $\gamma(B\bot_C\bot)$ is non-empty. It follows from *consistency* and Observation 5.2.6 that there is some $X \in B\bot_C\bot$ such that $B?\alpha \subseteq X$. By the construction of γ, $X \in \gamma(B\bot_C\bot)$.

(2) $B?_{C,\gamma}\alpha = \bigcap\gamma((B \cup \{\alpha\})\bot_C\bot) = \bigcap\{X \in (B \cup \{\alpha\}\bot_C\bot | (B+\alpha)?\beta \subseteq X$ for some $\beta \in A\} =$ *internal exchange* $\bigcap\{X \in (B \cup \{\alpha\}\bot_C\bot | (B+\alpha)?\alpha \subseteq X\} =$ (*pre-expansion*) $\bigcap\{X \in (B \cup \{\alpha\}\bot_C\bot | B?\alpha \subseteq X\}$. Hence, $B?\alpha \subseteq B?_{C,\gamma}\alpha$.

To see that $B?_{C,\gamma}\alpha \subseteq B?\alpha$, let $\beta \notin B?\alpha$. If $\beta \notin B+\alpha$, it is easy to see that $\beta \notin B?_{C,\gamma}\alpha$. Suppose that $\beta \in B+\alpha$. From $\beta \notin B?\alpha$ it follows by *pre-expansion* that $\beta \notin (B+\alpha)?\alpha$. From $\beta \in B+\alpha \setminus (B+\alpha)?\alpha$ it follows by *relevance* and *pre-expansion* that there is some B' such that $B?\alpha \subseteq B' \subseteq B \cup \{\alpha\}$, $\bot \notin C(B')$ and $\bot \in C(B' \cup \{\beta\})$. It follows from Observation 5.2.6 that there is some X such that $B' \subseteq X \in (B+\alpha)\bot_C\bot$. But then, since $B?\alpha \subseteq X$, $X \in \gamma((B+\alpha)\bot_C\bot)$. But since $\beta \notin X$, $\beta \notin (B+\alpha)?_{C,\gamma}\alpha =$ (by *pre-expansion*) $B?_{C,\gamma}\alpha$. □

Theorem 5.2.25: Let $\dot{-}_C$ satisfy the inclusion and core-retainment postulates for contraction, and let $*_C$ be the internal revision operator based on $\dot{-}_C$ via the Levi identity. Let C satisfy $\neg\alpha$-local non-contravention. Then:

$B\dot{-}_C\alpha = B \cap (B * \neg\alpha)$ (the Harper identity)

Proof: Using the Levi identity, what we have to show is that $B\dot{-}\alpha = B \cap ((B\dot{-}\alpha) \cup \{\neg\alpha\})$. The left-to-right inclusion follows directly from the inclusion postulate. For the right-to-left inclusion, let $\beta \in B \cap ((B\dot{-}\alpha) \cup \{\neg\alpha\})$. There are two cases.

Case i: $\beta \in B \cap (B\dot{-}\alpha)$. In this case, $\beta \in B\dot{-}\alpha$ follows directly.

Case ii: $\beta \in B \cap \{\neg\alpha\}$. Then $\beta = \neg\alpha$ and $\neg\alpha \in B$. Suppose that $\neg\alpha \notin B\dot{-}\alpha$. It follows from core-retainment for contraction that there is some $B' \subseteq B$ such that $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{\neg\alpha\})$. This contradicts $\neg\alpha$-local non-contravention, and we can conclude that $\neg\alpha \in B\dot{-}\alpha$, i.e. $\beta \in B\dot{-}\alpha$. □

Proposition 5.2.26: Let B and A be sets of formulas and, let C_A be a local inference operator. If $\dot{-}_{C_A}$ is a partial meet base contraction on B based on C_A, then for every α it holds that $B \setminus c(A, B) \subseteq B\dot{-}_{C_A}\alpha$.

Proof: If $\dot{-}_{C_A}$ is a partial meet base contraction on B based on C_A, then there exists a selection function γ such that $B\dot{-}_{C_A}\alpha = \bigcap\gamma(B\perp_{C_A}\alpha)$. It suffices to show that for every $X \in (B\perp_{C_A}\alpha)$, $B \setminus c(A, B) \subseteq X$. Let $X \in (B\perp_{C_A}\alpha)$. Then X is an inclusion-maximal subset of B such that $\alpha \notin C_A(X) = C(c(A, X))$. Suppose by contradiction that $B \setminus c(A, B) \not\subseteq X$. Let $\beta \in (B \setminus c(A, B)) \setminus X$. By part 2 of Observation 5.1.5, $c(A, X \cup \{\beta\}) \subseteq c(A, B)$ and hence, since $\beta \notin c(A, B)$, we have $\beta \notin c(A, X \cup \{\beta\})$. From this it follows that $c(A, X \cup \{\beta\}) = c(A, X)$ and $\alpha \notin C_A(X \cup \{\beta\})$, contradicting the maximality of X. □

Proposition 5.2.27: Let B and A be sets of formulas and, let C_A be a local inference operator. If $\dot{-}_{C_A}$ is a kernel contraction on B based on C_A, then for every α it holds that $B \setminus c(A, B) \subseteq B\dot{-}_{C_A}\alpha$.

Proof: If $\dot{-}_{C_A}$ is a kernel contraction on B based on C_A, then there exists an incision function σ such that $B\dot{-}_{C_A}\alpha = B \setminus \sigma(B \perp\!\!\!\perp_{C_A}\alpha)$. To show that $\sigma(B \perp\!\!\!\perp_{C_A}\alpha) \subseteq c(A, B)$, it suffices to show that for every $Y \in (B \perp\!\!\!\perp_{C_A}\alpha)$, $Y \subseteq c(A, B)$. Let $Y \in (B \perp\!\!\!\perp_{C_A}\alpha)$. Then Y is an inclusion-minimal subset of B such that $\alpha \in C_A(Y) = C(c(A, Y))$. Since by Part 1 of Observation 5.1.5 $c(A, Y) = c(A, c(A, Y))$ and $c(A, Y) \subseteq Y$, we have $\alpha \in C_A(c(A, Y)) = C(c(A, c(A, Y))) = C(c(A, Y)) = C_A(Y)$. From the minimality of Y it follows that $c(A, Y) = Y$. Since $Y \subseteq B$, by Part 2 of Observation 5.1.5 it follows that $c(A, Y) = Y \subseteq c(A, B)$. □

Lemma 5.4.4: If A and B are sets of formulas, α is a formula and there is no maximum size for any set involved, then $f(B -_A \alpha) = f(B) -_A \alpha$.

Proof: We can see what happens to each argument of a belief state when it goes through the operation defined in 5.4.3. The second argument (Cn) does not change. The first argument is not affected by the retrieval ($+_r$) operation. After the doubting ($+_d$) operation, we have $B \setminus \delta_\gamma(B, \alpha, A)$. The rejection operation does not affect the first argument and the acceptance ($+_a$) operation only adds to the first argument formulas that were already part of it. The third argument is empty before the operation. Retrieval adds $c(A, B)$ to it, doubting does not affect it, rejection deletes $\delta_\gamma(B, \alpha, A) = c(A, B) \setminus \bigcap\gamma(B \perp_{C_A} \alpha)$ and acceptance deletes $\rho(B, \alpha, A) = \bigcap\gamma(B \perp_{C_A} \alpha)$. After the operation, the third argument is empty again.

So, if we apply a local partial meet contraction to $f(B) = \langle B, Cn, \emptyset\rangle$, we obtain $\langle B \setminus \delta_\gamma(B, \alpha, A), Cn, \emptyset\rangle = \langle\bigcup\gamma(B\perp_{C_A}\alpha), Cn, \emptyset\rangle = f(B -_A \alpha)$. □

Bibliography

[AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[AM82] Carlos Alchourrón and David Makinson. On the logic of theory change: contraction functions and their associated revision functions. *Theoria*, 48:14–37, 1982.

[AM85] Carlos Alchourrón and David Makinson. On the logic of theory change: Safe contraction. *Studia Logica*, 44:405–422, 1985.

[And80] John R. Anderson. *Cognitive Psychology and its Implications.* W. H. Freeman, 1980.

[AB99] Carlos Areces and Veronica Becher. Iterable AGM functions. In Hans Rott and Mary-Anne Williams, editors, *Frontiers in Belief Revision.* Kluwer, 1999. to appear.

[BDP97] Salem Benferhat, Didier Dubois, and Henri Prade. Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study. Part 1: The flat case. *Studia Logica*, 58:17–45, 1997.

[Bil98] Dorrit Billman. Representations. In William Bechtel and George Graham, editors, *A Companion to Cognitive Science*, pages 649–659. Blackwell Publishers, 1998.

[Bou96] Craig Boutilier. Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25(3):262–305, 1996.

[dCH96] Luis Fariñas del Cerro and Andreas Herzig. Belief change and dependence. In Yohav Shoham, editor, *Proceedings of TARK 96*, 1996.

[Che86] Christopher Cherniak. *Minimal Rationality.* MIT Press, 1986.

[CGP99] Samir Chopra, Konstantinos Georgatos, and Rohit Parikh. Nonmonotonic inference on belief sequences (preliminary report). In Phokion Kolaitis and George Koletsos, editors, *Proceedings of the Second Pan-Hellenic Symposium on Logic*, 1999.

[CP99] Samir Chopra and Rohit Parikh. An inconsistency tolerant model for belief representation and belief revision. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI'99)*, 1999.

[CLR90] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. MIT Press, 1990.

[DP97] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89(1-2):1–29, 1997.

[Dix94] Simon E. Dixon. *Belief Revision: A Computational Approach*. PhD thesis, Basser Department of Computer Science, University of Sydney, 1994.

[vD97] Stijn van Dongen. Graph clustering and information structuring. In E. Kraak and R. Wassermann, editors, *Proceedings of the Accolade 97*. Dutch Graduate School in Logic, 1997.

[DP86] Jennifer Drapkin and Don Perlis. Step-logics: An alternative approach to limited reasoning. In Wolfgang Wahlster, editor, *Proceedings of the European Conference on Artificial Intelligence (ECAI)*. Wiley, 1986.

[Dun95] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-persons games. *Artificial Intelligence*, 77:321–357, 1995.

[Dun86] Michael Dunn. Relevance logic and entailment. In F. Guenthner and D. Gabbay, editors, *Handbook of Philosophical Logic*, volume 3, pages 117–224. D. Reidel, 1986.

[Eps90] Richard L. Epstein. *The semantic foundations of logic, volume 1: Propositional Logic*. Nijhoff International Philosophy Series. Kluwer Academic Publishers, 1990.

[FH88] Ronald Fagin and Joseph Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

[Fer99] Eduardo Fermé. Five faces of recovery. In Hans Rott and Mary-Anne Williams, editors, *Frontiers in Belief Revision*. Kluwer, 1999. to appear.

[Fuh91] André Fuhrmann. Theory contraction through base contraction. *Journal of Philosophical Logic*, 20:175–203, 1991.

[Gab85] Dov Gabbay. Theoretical foundations for nonmonotonic reasoning in expert systems. In K. Apt, editor, *Expert Systems, Logics and Models of Concurrent Systems*, pages 439–459, 1985.

[Gab96] Dov Gabbay. *Labelled Deductive Systems (Vol.1: Introduction).* Oxford University Press, 1996.

[Gal92] Julia Rose Galliers. Autonomous belief revision and communication. In Peter Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 220–246. Cambridge University Press, 1992.

[Gar93] Kevin Gary. RABIT – A spreading activation approach to real-time commonsense reasoning. Master's thesis, Arizona State University, Tempe, Arizona, 1993.

[GED96] Kevin Gary and Jennifer Elgot-Drapkin. RABIT– Bridging formal and implementational approaches to commonsense reasoning. Technical report, Arizona State University, 1996.

[Gär88] Peter Gärdenfors. *Knowledge in Flux - Modeling the Dynamics of Epistemic States.* MIT Press, 1988.

[GM88] Peter Gärdenfors and David Makinson. Revision of knowledge systems using epistemic entrenchment. In Moshe Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge (TARK)*, pages 83–95. Morgan Kaufmann, 1988.

[GM94] Peter Gärdenfors and David Makinson. Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65:197–245, 1994.

[GR95] Peter Gärdenfors and Hans Rott. Belief revision. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume IV, chapter 4.2. Oxford University Press, 1995.

[GSW89] Russell Greiner, Barbara A. Smith, and Ralph W. Wilkerson. A correction to the algorithm in Reiter's theory of diagnosis. *Artificial Intelligence*, 41(1):79–88, 1989.

[HCdK92] Walter Hamscher, Luca Console, and Johan de Kleer, editors. *Readings in Model-Based Diagnosis.* Morgan Kaufmann, 1992.

[Han91] Sven Ove Hansson. *Belief Base Dynamics.* PhD thesis, Uppsala University, 1991.

[Han92a] Sven Ove Hansson. In defense of base contraction. *Synthese*, 91:239–245, 1992.

[Han92b] Sven Ove Hansson. Reversing the Levi identity. *Journal of Philosophical Logic*, 22:637–639, 1992.

[Han94] Sven Ove Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59:845–859, 1994.

[Han97a] Sven Ove Hansson. Semi-revision. *Journal of Applied Non-Classical Logic*, 7(1-2):151–175, 1997.

[Han97b] Sven Ove Hansson, editor. *Theoria – special issue on non-prioritized belief revision*, volume 63, 1997.

[Han99a] Sven Ove Hansson. A survey of non-prioritized belief revision. *Erkenntnis*, 50(2/3):413–427, 1999.

[Han99b] Sven Ove Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Press, 1999.

[HW98] Sven Ove Hansson and Renata Wassermann. Local change: A preliminary report. In *Fourth Symposium on Logical Formalizations of Commonsense Reasoning, London*, 1998.

[Har86] Gilbert Harman. *Change in View – Principles of Reasoning*. MIT Press, 1986.

[Her97] Andreas Herzig. How to change factual beliefs using laws and independence information. In Dov M. Gabbay, Rudolf Kruse, Andreas Nonnengart, and Hans Juergen Ohlbach, editors, *Qualitative and Quantitative Practical Reasoning - First Int. Joint Conf. on Qual. and Quant. Practical Reasoning; ECSQARU-FAPR'97*, number 1244 in LNCS, pages 311–321. Springer, jun 1997.

[Jaś69] Stanislaw Jaśkowski. Propositional calculus for contradictory deductive systems. *Studia Logica*, 24:143–157, 1969. Translation from the Polish original that appeared in Studia Societatis Scientiarum Torunensis, Vol I, No. 5, 1948.

[KM92] Hirofumi Katsuno and Alberto O. 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, 1992.

[dK86] Johanes de Kleer. An assumption-based truth maintenance system. *Artificial Intelligence*, 28:127–162, 1986.

[LM98] Jérôme Lang and Pierre Marquis. Complexity results for independence and definability in propositional logic. In Anthony G. Cohn, Lenhart Schubert, and Stuart Shapiro, editors, *Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, pages 356–367, 1998.

[Lev84] Hector Levesque. A logic of implicit and explicit belief. In *Proceedings of AAAI-84*, 1984.

[Lev91] Isaac Levi. *The fixation of belief and its undoing.* Cambridge University Press, 1991.

[Lew82] David Lewis. Logic for equivocators. *Noûs*, 16:431–441, 1982.

[Lin91] Sten Lindström. A semantic approach to nonmonotonic reasoning: Inference operations and choice. (Unpublished Manuscript), 1991.

[Lou98] Ronald P. Loui. Process and police: Resource-bounded non-demonstrative reasoning. *Computational Intelligence*, 1998.

[Lug96] Valérie Lugardon. *Sur les Fondements de la Notion de Dépendance et de son Application a la Theorie de l'Action.* PhD thesis, Université Paul Sabatier de Toulouse, 1996.

[Mak87] David Makinson. On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic*, 16:383–394, 1987.

[Mak89] David Makinson. General theory of cumulative inference. In M. Reinfrank, J. de Kleer, M. L. Ginsberg, and E. Sandewall, editors, *Proceedings of the Second International Workshop on Non-Monotonic Reasoning*, LNAI. Springer-Verlag, 1989.

[Mak97] David Makinson. Screened revision. *Theoria*, 63:14–23, 1997.

[McC80] John McCarthy. Circumscription - a form of nonmonotonic resoning. *Artificial Intelligence*, 13(1+2):27–39, 1980.

[Neb90] Bernhard Nebel. *Reasoning and Revision in Hybrid Representation Systems*, volume 422 of *Lecture Notes in Artificial Intelligence.* Springer-Verlag, 1990.

[Neb98] Bernhard Nebel. How hard is it to revise a belief base? In D. Dubois and H. Prade, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 3: Belief Change*, pages 77–145. Kluwer Academic, 1998.

[Ols97] Erik J. Olsson. Coherence and the modularity of mind. *Australasian Journal of Philosophy*, 1997.

[Ori99] Francesco Orilia. Belief revision and the alethic paradoxes. In A. Chapuis and A. Gupta, editors, *Circularity, Definition, and Truth.* Indian Council of Philosophical Research, 1999. to appear.

[Par96] Rohit Parikh. Beliefs, belief revision and splitting languages. In *Proceedings of Itallc-96*, 1996.

[Pei77] Charles S. Peirce. The fixation of belief. *Popular Science Monthly*, 12:1–15, 1877. Reprinted in "Collected Papers of Charles S. Peirce", C. Hartshorne and P. Weiss (eds.), Harvard University Press, 1931-1935.

[Poo89] David Poole. Normality and faults in logic-based diagnosis. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 1304–1310, Detroit, 1989.

[PRN89] Graham Priest, Richard Routley, and Jean Norman, editors. *Paraconsistent logic: essays on the inconsistent.* Philosophia, 1989.

[Rei87] Raymond Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987. Reprinted in [HCdK92].

[RS95] Greg Restall and John Slaney. Realistic belief revision. Technical Report TR-ARP-2-95, Automatic Reasoning Project, Australian National University, 1995.

[Rod97] Odinaldo T. Rodrigues. *A Methodology for Iterated Information Change.* PhD thesis, Imperial College, University of London, 1997.

[Rot91] Hans Rott. A nonmonotonic conditional logic for belief revision. In André Fuhrmann and Michael Morreau, editors, *The logic of theory change.* Springer Verlag, 1991. LNAI 465.

[Rot96] Hans Rott. Making up one's mind: Foundations, coherence, nonmonotonicity. Habilitationsschrift, University of Konstanz., October 1996.

[Rot99] Hans Rott. Two dogmas of belief revision. In *Spinning Ideas (Eletronic Essays Dedicated to Peter Gärdenfors).* 1999. Available at http://www.lucs.lu.se/spinning/.

[RW91] Stuart Russell and Eric Wefald. *Do the Right Thing.* The MIT Press, 1991.

[Rya92] Mark Ryan. *Ordered Presentations of Theories: Default Reasoning and Belief Revision.* PhD thesis, Department of Computing, Imperial College, 1992.

[SC97] Marialuisa Sanseverino and Fulvio Cascio. Model-based diagnosis for automotive repair. *IEEE Expert*, 12(6):33–37, 1997.

[SC95] Marco Schaerf and Marco Cadoli. Tractable reasoning via approximation. *Artificial Intelligence*, 74(2):249–310, 1995.

[SL92] Guillermo R. Simari and Ronald P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence*, 53(2-3):125–157, 1992.

[Sim82] Herbert A. Simon. *Models of Bounded Rationality*, volume Volume 2: Behavioral Economics and Business Organization. MIT Press, 1982.

[Smi59] T. J. Smiley. Entailment and deducibility. *Proceedings of the Aristotelian Society*, pages 233–254, 1959.

[tTvH96] Annette ten Teije and Frank van Harmelen. Computing approximate diagnoses by using approximate entailment. In *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*, pages 256–265, Boston, Massachusetts, November 1996.

[tTvH97] Annette ten Teije and Frank van Harmelen. Exploiting domain knowledge for approximate diagnosis. In M. Pollack, editor, *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI'97)*, pages 454–459, Nagoya, Japan, August 1997.

[Ten84] Neil Tennant. Perfect validity, entailment and paraconsistency. *Studia Logica*, 43:181–200, 1984.

[Was97] Renata Wassermann. Towards a theory of resource-bounded belief revision. In Joeri Engelfriet and Tigran Spaan, editors, *Accolade'96 Proceedings*, 1997. A shorter version appeared in the Third Dutch/German Workshop on Nonmonotonic Reasoning.

[Was98] Renata Wassermann. On structured belief bases: A preliminary report. In *Seventh International Workshop on Nonmonotonic Reasoning.* Trento, 1998.

[Was99a] Renata Wassermann. On structured belief bases. In Hans Rott and Mary-Anne Williams, editors, *Frontiers in Belief Revision.* Kluwer, 1999. to appear.

[Was99b] Renata Wassermann. Resource-bounded belief revision. *Erkenntnis*, 50(2-3):429–446, 1999.

[Wil94] Mary-Anne Williams. On the logic of theory base change. In *Proceedings of the European Workshop on Logics in Artificial Intelligence (JELIA'95)*, 1994.

[Win90] Marianne Winslett. *Updating Logic Databases*, volume 9 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1990.

[Win95] Marianne Winslett. Epistemic aspects of databases. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume IV, chapter 4.2, pages 133–174. Oxford University Press, 1995.

Samenvatting

De afgelopen twintig jaar wordt Belief Revision uitgebreid bestudeerd. Het probleem waar het bij Belief Revision om gaat komt in het kort hier op neer: Hoe moet een gegeven agent met een verzameling (toegeschreven) geloven zijn geloven veranderen als hij nieuwe informatie binnenkrijgt? Met “agent” bedoelen wij een mens, een computer of een willekeurig systeem waaraan geloven kunnen worden toegeschreven en waarvan rationele reacties kunnen worden verwacht.

Dit is een multidisciplinair probleem met toepassingen in vele gebieden. Een paar voorbeelden laten zien hoe belief revision voorkomt in:

- Dagelijks leven: Ik dacht dat het in Amsterdam altijd regende. Op een ochtend word ik wakker in Amsterdam en het is mooi weer. Ik geloof dat het mooi weer is, en dat is in strijd met mijn vorige geloven. Ik moet dus mijn mening herzien en geloof nu niet meer dat het in Amsterdam altijd regent.

- Gegevensbestand: In een gegevensbestand met data over klanten van een boekhandel bestaat een entry voor Jan Smit, met geboortedatum 20/2/67. Ik krijg een nieuwe bestelling binnen waarbij bij de geboortedatum van Jan Smit 20/2/76 staat. Ik kan geen nieuwe geboortedatum toevoegen aan Jan's entry en Jan's geboortedatum kan ook niet veranderd zijn in de loop van de tijd. Wat moet ik nu doen? De oude datum bewaren? De oude datum door de nieuwe vervangen? Of is dit gewoon een andere Jan Smit die ik aan het gegevensbestand moet toevoegen?

- Robotica: Een mobiele robot heeft een plattegrond van het gebied waar hij moet bewegen. Volgens de plattegrond is er niets dat hem in de weg zal staan wanneer hij rechtdoor gaat. Maar zijn sensoren geven de aanwezigheid van een groot voorwerp aan dat tegenover hem staat. Moet de robot gaan twijfelen aan zijn sensoren en trachten rechtdoor te gaan? Of moet hij de sensoren geloven en aan de plattegrond gaan twijfelen?

- Diagnose: Ik geloof dat, als ik een artikel op de juiste positie op een kopieermachine plaats ik een kopie krijg van het artikel. Stel dat ik bij het kopiëren mijn artikel inderdaad op de juiste positie plaats maar slechts witte bladeren terugkrijg. Moet ik mijn geloof opgeven dat ik de juiste positie koos? Of moet ik gaan betwijfelen dat de machine het goed doet?

Tot nu toe werd Belief Revision meestal bestudeerd voor zeer geidealiseerde agenten. De agenten die daarbij in aanmerking komen zijn oneindige wezens zonder beperkingen op hun geheugen, beschikbare tijd en logisch vermogen. Het is geen triviale kwestie of (en hoe) men deze theorieën kan aanpassen zodat ze bruikbaar worden voor minder geidealiseerde agenten. Om bovengenoemde problemen op te kunnen lossen, hebben wij een theorie nodig die rekening houdt met de eindigheid van een agent en zijn redeneringen.

Uitgaande van de standaard logische modellen voor Belief Revision is het doel van dit proefschrift een theorie te ontwikkelen die op realistischer agenten kan worden toegepast. Wij zijn dus niet op zoek naar een computationele implementatie van bestaande theorieën, maar naar een theorie voor minder geidealizeerde agenten.

De belangrijkste resultaten van ons werk zijn:

1. Het formaliseren van een ruimer begrip van belief state, gebaseerd op het informele werk van Harman en Cherniak (Hoofdstuk 4).

2. Generalisatie van standaardresultaten uit de vakliteratuur, die het mogelijk maakt dat andere logica's kunnen worden gebruikt (Hoofdstuk 5). Dit is gezamenlijk werk met Sven Ove Hansson.

3. Het ontwerpen van een methode, die psychologisch gemotiveerd en computationeel efficient is, om op het relevante deel van een belief state te concentreren (Hoofdstuk 6).

4. Toepassing van het ontwikkelde model op diagnose problemen en het gebruik van computationele tools uit de diagnose literatuur om operatoren voor Belief Revision te implementeren (Hoofdstuk 7).

We beginnen met het ontwikkelen van een formele theorie die een model is voor belief states en voor eenvoudige operaties voor het veranderen van belief states. Ons model van belief states onderscheidt verschillende soorten geloven: expliciete versus impliciete geloven, geloven die op een bepaald moment actief dan wel niet actief zijn, en geloven die tijdelijk dan wel volledig geaccepteerd worden. In de literatuur zijn veel voorstellen gedaan tot het onderscheiden van geloven. Onze theorie verschilt in enkele opzichten van deze voorstellen:

- Anders dan bij Harman's en Cherniak's informele voorstellen hebben wij een formeel framework ontwikkeld, gebaseerd op verzamelingen, waarbij de verhouding tussen verschillende soorten geloven duidelijk wordt.

- In tegenstelling tot de formele benaderingen die onderscheid maken tussen expliciete en impliciete geloven, zoals die van Fagin en Halpern of van Levesque, eisen wij niet dat de verzameling van expliciete geloven consistent is of dat de verzameling impliciete geloven de klassieke logische afsluiting is van de expliciete geloven.

- In tegenstelling tot wat gebruikelijk is bij Belief Revision bouwen wij de operatoren stap voor stap op uit een reeks heel eenvoudige (basis) operatoren op belief states.

Wij hebben laten zien dat traditionele Belief Revision operatoren voor geidealiseerde agenten kunnen worden gemodelleerd in onze theorie, mits wij oneindige verzamelingen en een oneindige reeks basis operatoren toestaan.

De basis operatoren worden gebruikt als bouwstenen om meer ingewikkelde operatoren te modelleren. Wanneer een agent nieuwe informatie krijgt, wordt deze niet meteen volledig geaccepteerd, maar eerst als een tijdelijk geloof bewaard. De agent moet vervolgens, gegeven zijn vorige geloven, beslissen of hij deze tijdelijke geloven volledig zal accepteren. Onze theorie laat enigszins open hoe deze beslissing wordt genomen. In Appendix A laten wij een manier zien waarop deze beslissingsprocedure kan worden geimplementeerd. Wij stellen voor om argumentatie theorie bij een dergelijke beslissing te gebruiken. Wij passen Loui's model voor resource-bounded argumentatie toe, waarbij de agent rekening houdt met zoveel mogelijk, maar in het algemeen niet álle, argumenten.

Een ander belangrijk ingrediënt in onze model is het feit dat de verzameling van actieve geloven, d.w.z. de geloven die beschikbaar zijn voor redeneren, heel klein is in vergelijking met de verzameling expliciete geloven. Dit komt overeen met de intuïtie dat agenten niet over alles tegelijk kunnen denken. Een agent is meestal maar over één bepaald onderwerp aan het denken. Een van onze basis operatoren maakt expliciete geloven actief. Maar hoe beslist de agent welke geloven relevant zijn voor een bepaalde operatie? In Hoofdstukken 5 en 6 geven we twee verschillende oplossingen voor dit probleem. De eerste oplossing, die gepresenteerd wordt in Hoofstuk 5, gebruikt logica — en niets dan logica — om het relevante deel van de geloven van een agent te isoleren. Een geloof wordt als relevant beschouwd voor een bepaalde formule als het helpt bij het bewijzen van die formule of zijn negatie. Wij definiëren locale operaties voor Belief Revision die alleen maar invloed hebben op het relevante deel van de belief state. Deze methode heeft veel tekortkomingen: het vinden van de relevante geloven is computationeel net zo moeilijk als de traditionele Belief Revision operatoren. Bovendien is er geen controle op de grootte van het relevante deel, d.w.z. het kan gebeuren dat de verzameling van relevante geloven gelijk is aan de verzameling expliciete geloven. Maar de methode geeft ons wel enkele interessante formele resultaten. Al de locale operaties die wij hebben gedefinieërd worden geaxiomatiseerd en de representatie stellingen geven precies aan wat nodig is om de elegante eigenschappen van de operaties te krijgen. Omdat deze axiomatiseringen onafhankelijk zijn van het concept relevantie, presenteren wij in Hoofdstuk 6 een computationeel efficiente methode om de relevante geloven te vinden. Voor deze methode hebben wij niet-logische informatie nodig die ons zegt hoe de geloven met elkaar gerelateerd zijn. Dit leidt tot een notie van graden van relevantie. Wij laten zien dat de relatie tussen geloven meestal kan worden afgeleid van de gegeven belief state of van een bepaalde toepassing, zoals die bijvoorbeeld gegeven wordt in Hoofdstuk 7. Behalve de computationele voordelen is de methode die beschreven wordt in Hoofdstuk 6 ook heel intuítief en sluit goed aan bij onderzoek over cognitieve modellen van het geheugen.

In Hoofdstuk 7 beschrijven wij een toepassing van de ontwikkelde theorie op het gebied van diagnose. Diagnose van elektronische circuits wordt gebruikt als concreet voorbeeld en maakt duidelijk wat de theoretische concepten eigenlijk betekenen. Relevantie kan hier een causale relatie tussen input en output van een component van het circuit betekenen. De methode uit Hoofdstuk 6 wordt gebruikt om de diagnose procedure te beperken tot het relevante deel van een circuit. Door het expliciet maken van de relatie tussen Belief Revision en Diagnose laten wij zien dat sommige computationele tools ontwikkeld voor diagnose kunnen worden gebruikt om Belief Revision te implementeren. In het bijzonder laten we zien hoe Reiter's algoritme voor diagnose gebaseerd op consistentie kan worden gebruikt om kernel semi-revision te implementeren.

Titles in the ILLC Dissertation Series:

ILLC DS-1995-01: **Jacob Brunekreef**
On Modular Algebraic Protocol Specification

ILLC DS-1995-02: **Andreja Prijatelj**
Investigating Bounded Contraction

ILLC DS-1995-03: **Maarten Marx**
Algebraic Relativization and Arrow Logic

ILLC DS-1995-04: **Dejuan Wang**
Study on the Formal Semantics of Pictures

ILLC DS-1995-05: **Frank Tip**
Generation of Program Analysis Tools

ILLC DS-1995-06: **Jos van Wamel**
Verification Techniques for Elementary Data Types and Retransmission Protocols

ILLC DS-1995-07: **Sandro Etalle**
Transformation and Analysis of (Constraint) Logic Programs

ILLC DS-1995-08: **Natasha Kurtonina**
Frames and Labels. A Modal Analysis of Categorial Inference

ILLC DS-1995-09: **G.J. Veltink**
Tools for PSF

ILLC DS-1995-10: **Giovanna Cepparello**
Studies in Dynamic Logic

ILLC DS-1995-11: **W.P.M. Meyer Viol**
Instantial Logic. An Investigation into Reasoning with Instances

ILLC DS-1995-12: **Szabolcs Mikulás**
Taming Logics

ILLC DS-1995-13: **Marianne Kalsbeek**
Meta-Logics for Logic Programming

ILLC DS-1995-14: **Rens Bod**
Enriching Linguistics with Statistics: Performance Models of Natural Language

ILLC DS-1995-15: **Marten Trautwein**
Computational Pitfalls in Tractable Grammatical Formalisms

ILLC DS-1995-16: **Sophie Fischer**
The Solution Sets of Local Search Problems

ILLC DS-1995-17: **Michiel Leezenberg**
Contexts of Metaphor

ILLC DS-1995-18: **Willem Groeneveld**
Logical Investigations into Dynamic Semantics

ILLC DS-1995-19: **Erik Aarts**
Investigations in Logic, Language and Computation

ILLC DS-1995-20: **Natasha Alechina**
Modal Quantifiers

ILLC DS-1996-01: **Lex Hendriks**
Computations in Propositional Logic

ILLC DS-1996-02: **Angelo Montanari**
Metric and Layered Temporal Logic for Time Granularity

ILLC DS-1996-03: **Martin H. van den Berg**
Some Aspects of the Internal Structure of Discourse: the Dynamics of Nominal Anaphora

ILLC DS-1996-04: **Jeroen Bruggeman**
Formalizing Organizational Ecology

ILLC DS-1997-01: **Ronald Cramer**
Modular Design of Secure yet Practical Cryptographic Protocols

ILLC DS-1997-02: **Nataša Rakić**
Common Sense Time and Special Relativity

ILLC DS-1997-03: **Arthur Nieuwendijk**
On Logic. Inquiries into the Justification of Deduction

ILLC DS-1997-04: **Atocha Aliseda-LLera**
Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence

ILLC DS-1997-05: **Harry Stein**
The Fiber and the Fabric: An Inquiry into Wittgenstein's Views on Rule-Following and Linguistic Normativity

ILLC DS-1997-06: **Leonie Bosveld - de Smet**
On Mass and Plural Quantification. The Case of French 'des'/'du'-NP's.

ILLC DS-1998-01: **Sebastiaan A. Terwijn**
Computability and Measure

ILLC DS-1998-02: **Sjoerd D. Zwart**
Approach to the Truth: Verisimilitude and Truthlikeness

ILLC DS-1998-03: **Peter Grunwald**
The Minimum Description Length Principle and Reasoning under Uncertainty

ILLC DS-1998-04: **Giovanna d'Agostino**
Modal Logic and Non-Well-Founded Set Theory: Translation, Bisimulation, Interpolation

ILLC DS-1998-05: **Mehdi Dastani**
Languages of Perception

ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke

ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation

ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz

ILLC DS-1999-04: **Barbara Terhal**
Quantum Algotithms and Quantum Entanglement