Three Studies on Belief Removal

Raghav Ramachandran
M. Sc. Mathematics

June 21, 2011

Department of Computing
Faculty of Science
Macquarie University
North Ryde, NSW 2109
Australia

Thesis Supervisors:
Dr. Abhaya Nayak, Department of Computing
Prof. Mehmet Orgun, Department of Computing

This thesis is presented for the degree of
Doctor of Philosophy.
**Keywords:** Belief Contraction, Belief Erasure, Iterated Belief Contraction, Degrees of Belief, Probabilistic Belief Contraction, MaxEnt, Shannon Entropy, Hartley Entropy, Probabilistic Belief Erasure, Lewis’ Imaging, Conditional Update, Conditional Erasure.
It is not disbelief that is dangerous to our society; it is belief.

G. B. Shaw
# Contents

<table>
<thead>
<tr>
<th>Contents</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>Dedication</td>
<td>ix</td>
</tr>
<tr>
<td>Preface</td>
<td>ix</td>
</tr>
</tbody>
</table>

## 1 Introduction  
1.1 Factors affecting Belief Change  ........................................ 2  
1.2 Rationality Postulates .................................................. 8  
  1.2.1 AGM theory .......................................................... 9  
  1.2.2 Construction of belief change functions ........................... 12  
  1.2.3 System of Spheres .................................................. 14  
  1.2.4 KM theory .......................................................... 18  
1.3 Probabilistic belief change .............................................. 21  
  1.3.1 Probabilistic static change .................................... 21  
  1.3.2 Probabilistic dynamic change ................................... 27  
1.4 Research contribution .................................................. 28  
  1.4.1 Iterated belief contraction ................................... 29  
  1.4.2 Probabilistic belief erasure ................................... 30  
  1.4.3 Probabilistic belief contraction ............................... 31  
1.5 Concluding remarks ................................................... 32

## 2 Iterated Belief Contraction  
2.1 Introduction ............................................................. 35  
2.2 Three Contraction Functions .......................................... 39  
  2.2.1 Moderate contraction function ................................ 40  
  2.2.2 Natural contraction function .................................. 41  
  2.2.3 Lexicographic contraction function ......................... 42  
  2.2.4 Comparison of the three contraction functions ............ 44  
  2.2.5 Principled factored insertion ................................. 46  
2.3 Degrees of Belief ..................................................... 48  
  2.3.1 Conditional degrees of belief ................................. 50  
2.4 Plausible Properties of Iterable Contraction ....................... 52  
  2.4.1 Properties based on Degrees of belief ....................... 54
Abstract

It is important to study how a rational agent should respond to a piece of new information. Two such responses are: (1) incorporation of the new information into its body of beliefs, and (2) removal from it of some old information. This thesis deals with this latter response. It consists of three self-contained studies on belief removal, preceded by an extended introduction. Hence, this thesis might be viewed as a compendium.

The first study examines a particular problem related to belief removal in a non-probabilistic framework. The problem in question is that the recent proposals for iterated belief removal lack complete characterization. The two subsequent essays deal with issues relating to belief removal in a probabilistic context, which would involve a procedure to reduce the prior of some hypothesis from 1 to something less than 1. The problem is: this cannot be done in a purely Bayesian fashion, and is a problem not very well researched in the community.

The first study aims at providing the characterization for three existing proposals for iterated belief removal, namely, Priority Contraction, Natural Contraction and Lexicographic Contraction. The second study deals with the problem of probabilistic belief removal in a dynamic domain where the need to remove (erase) some an old belief indicates that the environment has changed and ones knowledge is dated. This study is based upon a generalization of David Lewiss account of imaging just as Richard Jefferys account of probability kinematics is a generalization of Bayesian updating. Finally, the last study deals with the problem of probabilistic belief removal in a static domain where the need to remove (contract) an existing belief indicates that some of ones current beliefs are incorrect. This study further develops a proposal for probability de-conditioning by Ramer, motivated by the principle of Maximum Entropy.
Preface

This
Chapter 1

Introduction

Initially, I used to believe Hercules was immortal. Sometime later I was informed that he was, in fact, killed with the help of Hydra’s poison by a scheming centaur. This information was inconsistent with my beliefs. One cannot believe both Hercules is immortal and that he was killed. In order to maintain consistency, I had to make a choice between my belief and the information received. Since I felt my belief was more questionable than the new information, I decided to make changes to my system of beliefs. This involved retracting my belief in Hercules being immortal. There have been numerous other instances when I needed to make changes to my belief system. Such instances arose because some of my beliefs were not well founded and also because I had no opinion about certain topics. There were also cases when I had to change my beliefs because the objects described in the beliefs had evolved over time. Such changes to the belief system are quite common with all reasoning agents, be they human or robots.

In this dissertation I will focus on the study of belief removal. In particular, I will investigate the problem of iterated belief contraction and the problem of probabilistic belief removal. These problems have been either left unanswered or have been just partially addressed. In fact, this dissertation is a compilation of three self-contained contributions that focus on iterated contraction and probabilistic belief removal. Along with the required background, this introductory chapter will describe the contribution these works make to the research field.
1.1 Factors affecting Belief Change

The belief system I mentioned earlier is referred to as a *belief state* [6, 16]. The belief state of an agent primarily contains information as to what beliefs it holds. When available, the belief state also includes other related information such as preferences among the beliefs and ranking among non-beliefs, which play a central role in belief management.

As agents modify their belief state, new beliefs come into existence, some existing beliefs are discarded and the preferences among the beliefs are changed. When new beliefs are added, they could conflict with the existing ones leading to inconsistency. For instance, believing in ‘Hercules is dead’ when the agent already believes that ‘Hercules is immortal’ leads to inconsistency. Such a conflict is not desirable. A reasoning agent should strive to maintain consistency in its belief state. It is evident that while striving for consistency, some existing beliefs could be lost and hence the changes need not be monotonic with respect to the set of beliefs.\(^1\) Several theories of *non-monotonic reasoning* have been proposed to model the different changes to a belief state. Detailed surveys of research in this topic are provided in [27, 30, 39, 43, 65] among many others.

The study of belief change is governed by many factors. Some of the important factors are as follows:

I Is information being added or some belief is being removed?

II Is the required change concerned with a dynamic or static environment?

III How is the belief state represented?

IV Is the belief change procedure repeatable?

I Is information being added or some belief is being removed?

As a result of some change to the belief state, new beliefs could be added and/or existing beliefs could be removed.

\(^1\)Monotonicity indicates that learning a new piece of knowledge cannot reduce the set of what is already known. However, when learning a new piece of information that contradicts existing pieces of knowledge, in order to maintain consistency, the set of existing knowledge is reduced.
(i) In the instance presented earlier relating to Hercules being mortal, the sentence ‘Hercules is immortal’ was initially considered certain. That sentence was made uncertain and its negation, that is, ‘Hercules is not immortal’, was made certain in its place. This change is termed *replacement*, where an existing belief is replaced by a new one.

(ii) There are instances where the agent will be uncertain about a sentence and its negation, clearly unable to believe in either of them due to lack of information. When informed one way or another, the agent could become certain about either the sentence or its negation, and add it to its belief state. Addition of a new belief in this case amounts to ending the uncertainty associated with it. No belief is retracted in such an instance.

(iii) In a situation where the agent receives information which casts doubt on a particular belief, the sentence which is considered to be certain is made uncertain. Such cases correspond to retracting belief in the sentence.

II *Is the required change concerned with a dynamic or static environment?*

An agent may need to change its belief state due to various reasons. We describe two possible reasons here.

(i) Sometimes the agent is informed that a particular sentence holds true and the agent includes this new sentence in its belief state. Similarly, when informed otherwise the agent might retract some belief from its belief state.

(ii) However, on occasion the agent could be informed that some action has taken place which changes the state of the world. For instance, a worker from the cleaning service informs the agent that its office has been cleaned. This action might change the truth value of many relevant sentences. The sentence ‘the bin is not empty’ which could have been true prior to the action of cleaning taking place, is likely to be false after cleaning. Hence the agent needs to change its belief state in order to keep it up to date. A single factor differentiates the two types of changes: the *state of the real world*. In the former case the state of the real world is static while in the latter case the state of the world is dynamic.

Different belief inclusion and belief removal operations have been studied in the lit-
Belief expansion, belief revision and belief contraction are three widely studied static belief change operations [2]. Belief expansion is a simple belief addition operation. When the agent learns about or receives information, expansion leads to adding a new belief to the current belief state. Contraction is a belief removal operation in a static environment. Revision is a belief replacement operation in a static environment. If the information received is consistent with the existing beliefs, then revision reduces to expansion. Revision and expansion differ from each other when the information is not consistent with the belief state. Expansion is not capable of removing the logical inconsistency caused by such information. In such a case, expansion adds the new belief to the belief state leading to logical inconsistency. However, revision is capable of handling logical inconsistency. When the information is inconsistent with the existing beliefs, revision results in a new belief state which incorporates the received information while maintaining the consistency of the belief state (assuming that the new information is self-consistent).

Levi sees revision as a two step process wherein the first step involves contraction which is followed by an expansion [58]. When revising by a sentence \( \alpha \) the belief state first undergoes contraction by its negation \( \neg \alpha \) in a bid to avoid possible inconsistencies later. This contracted belief state then undergoes expansion by \( \alpha \). This is generally considered to be as close to a definition of belief revision as possible, and is also referred to as the Levi identity. This definition of belief revision suggests that the existing belief is always sacrificed to accommodate the new information. Some researchers have questioned this approach and have proposed alternatives where the existing belief is given fair opportunity to retain its place in the belief state, for instance, non-prioritized belief revision [42] and screened revision [66]. Since this dissertation focuses on belief removal, we do not discuss these functions here.

There is another well-known identity, due to Harper [45], called Harper Identity, which defines contraction via revision. According to the Harper identity, the result of
contraction by a belief should be the mixture of the initial belief state and the result of its revision by the negation of that particular belief. A sentence considered uncertain in any of those two states is considered uncertain in the contracted state. Harper identity will be re-visited in chapters 2 and 4.

The belief change operations in a dynamic environment are called *belief update* and *belief erasure* [53]. The modification or the change made to the agent’s belief state in order to incorporate a new belief as a result of a changed environment is termed belief update. Belief erasure is the corresponding belief removal operation, that is, erasure captures loss of beliefs that is triggered by possible changes to the environment. The definition of these operations depends on how a belief state is represented. This is a major factor that governs the study of belief change.

**III How is the belief state represented?**

The belief state of an agent has been represented in many different ways. The simplest representation being a set of sentences closed under logical consequence operation [2]. This set is called a *belief set*. There have been other qualitative representations of a belief state such as a *belief base* [40, 41], an *epistemic entrenchment* relation [29], and a *system of spheres* [32]. These representations generally give a set of sentences representing the beliefs of the agent, along with some kind of mechanism that facilitates belief change. In case of the *belief base* representation, the sentences in the base are more highly regarded than the beliefs that are logically deduced from the base. An epistemic entrenchent relation gives an explicit preference ordering among the beliefs. Under the system of spheres representation, a preference ordering among the beliefs is presented indirectly in terms of a plausibility ordering among the possible worlds. Chapter 2 of this dissertation deals with a system of spheres representation of the belief state.

Apart from these representations, numerical representations are representations based on possibilistic logic [19, 21, 22, 95] and ranking functions [46, 92] have also been studied. In the possibilistic framework, a possibility measure determines which sentences are
beliefs. This measure also forms an ordering among the beliefs. In the representations
based on the ranking theory, a measure of disbelief is assigned to the sentences of the
language, based on which a degree of beliefs is derived.

Belief states are also represented in terms of probability functions. Probabilities of
sentences represent the uncertainty attached to them. The agent could either assign strict
probabilities to each of the sentences with the aid of a subjective probability function or
assign vague probabilities with the aid of a set of probability functions [49, 57]. Accord-
ingly, changing the belief state corresponds to changing the initial probability function(s).
A detailed survey of the probabilistic logic can be found in [36, 37]. Chapters 3 and 4
deal with a probabilistic representation of the belief state.

IV Is the belief change procedure repeatable?

An agent does not change its belief state just once. Any reasonable agent can be expected
to change its belief state many number of times. Without stretching our imagination, we
can visualize situations that demand successive inclusion of new beliefs, or successive
removal of existing beliefs or consecutive inclusions and removals.

In case of retracting a belief, alongside the retracted belief some other related beliefs
may need to be withdrawn from the belief state as well. The agent might lose all the
related beliefs in order to retract a particular belief which would result in unnecessary
loss of beliefs. This is against the principle of informational economy [29, 87], according
to which the loss of beliefs as a result of any change should be kept at a minimum. Instead
the agent might choose to retain as much of the existing beliefs as possible and still hope
to be successful in retracting the given belief. The manner by which the agent retains
the belief is based on the agent’s preferences over beliefs. Belief change is dictated by
the principle of preference and indifference [87]. The principle of preference suggests
that when given a choice the belief that is more preferred is retained at the expense of
less preferred one; while the principle of indifference suggests that when the agent is
indifferent between two beliefs, they share the same fortune as a result of the change.
However, as a result of belief change, it is possible that the preferences of the agent undergo certain changes [73, 85].

A belief change function is said to be *iterable* if and only if it is capable of handling successive changes to the belief state. Due to the uncompromising role of these preferences in belief change, their representation as part of the belief state is necessary. Moreover, it is also necessary that an iterable belief change function specifies how the preferences of the agent might change as a result of a belief change, failing which the belief change function cannot be applied on the partially specified belief state, namely, where the contracted set of beliefs is specified but not the changed preference ordering. This is captured as the *principle of categorial matching*\(^2\) (PCM) [30, 84] :

**PCM** *The representation of a belief state after a change has taken place should be of the same format as the representation of the belief state before the change.*

This principle is identified as a necessary characteristic for any *iterable belief change* function [84]. A one-shot belief change function need not satisfy **PCM** since it does not worry about successive change. It only needs to mention the resultant set of beliefs, since the predominant role of the preference ordering is in guiding the change. I discuss more on this principle later in this chapter.

**Notations and Preliminaries**

We assume throughout that the agent’s language \( \mathcal{L} \), in which its beliefs are expressed, is generated by the connectives \( \lor, \land, \neg \) and \( \rightarrow \) from a finite set of alphabet. The sentences of the language are denoted by lowercase Greek letters \( \alpha, \beta, \gamma \) with or without decorations. We denote sets of sentences by the upper case Roman letters \( A \) and \( B \). The set of all possible worlds \( \Omega \) over \( \mathcal{L} \) is a finite set and the individual worlds are denoted by lowercase Greek characters such as \( \omega, \upsilon \) with possible decorations. A classical consequence operator \( \mathsf{Cn} \) governs the background logic. It is defined by \( \mathsf{Cn}(A) = \{ \alpha | A \vdash \alpha \} \), where \( \vdash \) is the

\(^2\)This principle is also known as the *principle of adequacy of representation* in [15].
classical consequence relation for propositional logic. A model of a set of sentences $A$ is an interpretation/possible world in which all the sentences in $A$ are true. For any set of sentences $A$, $[A]$ denotes the set of such models. When representing a singleton set $\{\alpha\}$, we simply denote it by $[\alpha]$.

Probability distributions over $\Omega$ are denoted by $P$ and $Q$ (with and without decorations) and the set of all probability distributions is denoted by $\mathcal{P}$. The probability distribution defined over $\Omega$ is such that, $\sum_{\Omega} P(\omega) = 1$. The probability function on $\mathcal{L}$ is defined as follows: for any sentence $\alpha$,

$$P(\alpha) = \sum_{\omega \in [\alpha]} P(\omega).$$

A total preorder relation is a connected, reflexive and transitive relation. A preorder relation over $\Omega$ is denoted by $\sqsubseteq$. This represents a total preorder relation unless specified otherwise. We use such a relation to denote a plausibility ordering among the set of all possible worlds. The strict and the symmetric parts of the relation $\sqsubseteq$ are given by $\sqsubset$ and $\approx$ respectively.

## 1.2 Rationality Postulates

In this section, I recount two important accounts of belief change: (1) AGM theory of static belief change, proposed by Alchourrón, Gärdenfors and Makinson in [2]; and (2) KM theory of dynamic belief change, proposed by Katsuno and Mendelzon in [53]. The formal study of the belief change operations has been founded upon a list of rationality postulates. These rationality postulates bring out the characteristics that define a particular operation and differentiate it from others. These postulates have become the benchmark for comparing different approaches or proposing equivalent operations in different frameworks. They will play an important role in judging the efficacy of the approaches presented later in this dissertation.
1.2.1 AGM theory

In a seminal work on belief change [2], the authors present a list of rationality postulates, that expansion, contraction and revision functions should adhere to. In the AGM theory, the belief state is represented by a belief set, that is, a set of sentences closed under logical consequence, generally represented by $\mathcal{K}$. The absurd belief set, that is, an inconsistent set of beliefs, is denoted by $\mathcal{K}_\bot$, and the belief set that only contains tautologies is denoted by $\mathcal{K}_\top$.

When a belief set $\mathcal{K}$ is expanded with a sentence $\alpha$, then the resulting set is denoted by $\mathcal{K}_\alpha^+$. The set of rationality postulates for expansion result in a unique definition for expansion. It is obtained by adding the sentence $\alpha$ to $\mathcal{K}$ and closing it under the logical consequence. Therefore the result of expansion of $\mathcal{K}$ by $\alpha$ is given by:

$$\mathcal{K}_\alpha^+ = \text{Cn}(\mathcal{K} \cup \{\alpha\}).$$

Revision of $\mathcal{K}$ by a sentence $\alpha$ ensures that the sentence $\alpha$ is a belief after the change and that the result of revision is a consistent belief set: it is denoted by $\mathcal{K}_\alpha^*$. Revision is governed by the following set of rationality postulates [2, 29]:

**R1** $\mathcal{K}_\alpha^*$ is a belief set.

**R2** $\alpha \in \mathcal{K}_\alpha^*$.

**R3** $\mathcal{K}_\alpha^* \subseteq \mathcal{K}_\alpha^+$.

**R4** If $\neg \alpha \notin \mathcal{K}$, then $\mathcal{K}_\alpha^+ \subseteq \mathcal{K}_\alpha^*$.

**R5** $\mathcal{K}_\alpha^* = \mathcal{K}_\bot$ iff $\vdash \neg \alpha$.

**R6** If $\text{Cn}(\alpha) = \text{Cn}(\beta)$, then $\mathcal{K}_\alpha^* = \mathcal{K}_\beta^*$.

Postulates **R1** and **R2** state that when revising $\mathcal{K}$ by $\alpha$, then the result of revision is a belief set that contains $\alpha$, that is, revision is successful. Posutales **R3** and **R4** explore the relation between expansion and revision. When revising $\mathcal{K}$ by a sentence $\alpha$, which is consistent in itself, the result is a consistent set of beliefs; whereas consistency in the result of expansion is dependent on whether $\alpha$ is consistent with $\mathcal{K}$ or not. If $\alpha$ is consistent with
\( K \), then the result of expansion is also consistent and coincides with the result of revision as depicted by \( R3 \) and \( R4 \). If \( \alpha \) is inconsistent with \( K \), then expansion results in the absurd state \( K_\perp \) and hence subsumes the result of revision as depicted in \( R3 \). Postulate \( R5 \) states that the only scenario under which revision results in inconsistency is when the sentence \( \alpha \) is self-contradictory. According to \( R6 \), revision is not sensitive to the syntax. These six posulates are termed the basic postulates of revision. Two supplementary postulates are also given for revision.

\( R7 \) \( K^*_\alpha \wedge \beta \subseteq (K^*_\alpha)_\beta^+ \).

\( R8 \) If \( \neg \beta \notin K^*_\alpha \), then \( (K^*_\alpha)_\beta^+ \subseteq K^*_\alpha \wedge \beta \).

These postulates discuss the effect of revising by a conjunction of sentences. Postulate \( R7 \) suggests that the beliefs obtained from revision by \( \alpha \wedge \beta \) should also be obtained when first revising by \( \alpha \) and then expanding by \( \beta \). The postulate \( R8 \) gives a conditional converse of \( R7 \). It states that, if \( \beta \) is consistent with the belief set that results from revision by \( \alpha \), then the result of expanding the revised belief set by \( \beta \) should be obtained from revising the initial set by \( \alpha \wedge \beta \).

Unlike expansion, the postulates of revision do not ensure a unique construction of the revision function. In fact it is shown in [2, 29] that an AGM revision function, that is, a revision function that satisfies \( R1 \) to \( R8 \) can be obtained from the Levi identity, provided the contraction function satisfies the postulates of contraction \( C1 \) to \( C8 \) (see below). Levi identity states that the result of revising a belief set \( K \) by a sentence \( \alpha \) is a two-step process, first contracting \( K \) by \( \neg \alpha \) and then expand it by \( \alpha \).

\[
K^*_\alpha = (K_{\neg \alpha})^+_\alpha
\]  

(Levi identity)

Contraction results in some sentence of \( K \) being retracted from the belief set without adding any new belief. In order for the result of contraction to be considered as a belief set, to be closed under logical consequence, it is necessary to give up some other sentences from \( K \) as well. The result of contracting \( K \) with respect to \( \alpha \) is denoted by \( K^-_\alpha \).
Contraction is defined by the following postulates:

**C1 Closure:** \( \mathcal{K}_\alpha^- \) is a theory whenever \( \mathcal{K} \) is a theory.\(^3\)

**C2 Inclusion:** \( \mathcal{K}_\alpha^- \subseteq \mathcal{K} \).

**C3 Vacuity:** If \( \alpha \notin \text{Cn}(\mathcal{K}) \), then \( \mathcal{K}_\alpha^- = \mathcal{K} \).

**C4 Success:** If \( \emptyset \not\vdash \alpha \), then \( \alpha \notin \mathcal{K}_\alpha^- \).

**C5 Preservation:** If \( \text{Cn}(\alpha) = \text{Cn}(\beta) \), then \( \mathcal{K}_\alpha^- = \mathcal{K}_\beta^- \).

**C6 Recovery:** If \( \alpha \in \mathcal{K} \), then \( \mathcal{K} \subseteq (\mathcal{K}_\alpha^-)^+ \).

The *closure* postulate ensures that the result of contraction is also a belief set; while the postulate of *inclusion* ensures no new belief is added to the belief set when contracting \( \mathcal{K} \). When the sentence being removed from \( \mathcal{K} \) does not belong to the set, then the act of contraction becomes vacuous and this is captured by the postulate of *vacuity*. Since a belief set is considered to be closed under logical consequence, a tautology cannot be contracted from \( \mathcal{K} \). The *success* postulate ensures that contracting any non-tautological belief from \( \mathcal{K} \) is successful. Contraction of a belief set is not syntax-sensitive and this is ensured by the postulate of *preservation*. *Recovery* states that contraction of the belief set by \( \alpha \), followed by expansion by \( \alpha \) recovers all the initial beliefs. Just as in the case of revision, these six postulates are called the basic AGM postulates for contraction [2]. Two more postulates for contraction were proposed to study the relation between contraction by \( \alpha \) and contraction by \( \alpha \land \beta \). When contracting a belief \( \alpha \land \beta \), then there is a choice between removing \( \alpha \) or \( \beta \) or both.

**C7 Intersection:** \( \mathcal{K}_\alpha^- \cap \mathcal{K}_\beta^- \subseteq \mathcal{K}_{\alpha \land \beta}^- \), for any theory \( \mathcal{K} \).

**C8 Conjunction:** If \( \beta \notin \mathcal{K}_{\alpha \land \beta}^- \), then \( \mathcal{K}_{\alpha \land \beta}^- \subseteq \mathcal{K}_\beta^- \), for any theory \( \mathcal{K} \).

The Harper identity, extracted from Harper’s analysis of the relation between contraction and revision in [45], amounts to the following: a sentence belongs to the contracted set \( \mathcal{K}_\alpha^- \) if and only if it belongs to both the initial set \( \mathcal{K} \) and the revised set \( \mathcal{K}_{\alpha}^- \). Harper

\(^3\)A theory is a set of sentences closed under the classical consequence operation, that is, it is a belief set. We use ‘theory’ here following the convention in [2, 29].
identity is formulated as follows [29]:

\[ K^-\alpha = K \cap K^\perp\alpha. \]  

(Harper identity)

A series of results are presented in [2, 29] which connect an AGM contraction and an AGM revision function via the Levi identity and the Harper identity. For instance, if the revision function satisfies the basic revision postulates, R1 to R6, then the contraction function defined via Harper identity satisfies the basic postulates of contraction, that is, C1 to C6. As noted earlier, unlike the case of expansion, these postulates of revision and contraction do not result in a unique definition of revision or contraction. I outline below some of the proposed efforts to give a construction of a contraction and revision function given in [2, 29].

1.2.2 Construction of belief change functions

The rationality postulates are backed by a construction of an AGM contraction function.

The principle of informational economy [29] suggests that the initial belief set should be preserved as much as possible. Thus it is understandable to expect the result of contraction of \( K \) by \( \alpha \) to be a maximal subset of \( K \) from which \( \alpha \) cannot be deduced logically. However, in general there could be many maximal subsets of \( K \) which do not imply \( \alpha \).

The set of all maximally consistent subsets of \( K \) which do not entail \( \alpha \) is denoted by \( K^\perp\alpha \).

If the sentence \( \alpha \) is a tautology, that is \( \vdash \alpha \), then the set \( K^\perp\alpha \) is an empty set. Three different contraction functions have been described using this set \( K^\perp\alpha \) as the basis. These functions are based on a ‘selection function’ \( S \).

The contraction \( K^-\alpha \) of \( K \) with respect to \( \alpha \) is based on a selection function \( S \) that picks out an element from \( K^\perp\alpha \), \( S(K^\perp\alpha) \). A family of contraction functions, namely maxichoice contractions, is defined based on the selection function \( S \) as follows [2, 29]:
Definition 1.1: $\mathcal{K}_\alpha^- = \begin{cases} S(\mathcal{K} \perp \alpha) & \text{when } \not\vdash \alpha \\ \mathcal{K} & \text{otherwise.} \end{cases}$

It has been shown that a maxichoice contraction function as given by the Definition 1.1 satisfies the basic AGM postulates [2, 29]. A maxichoice revision function is the revision function generated by Levi identity where the contraction function involved is a maxichoice contraction function. As shown in [2, 3, 29], such a maxichoice revision function results in a maximal belief set $\mathcal{K}^*_\alpha$ when revising $\mathcal{K}$ by $\alpha$ in the sense that given a sentence $\beta$ in the language, the agent either believes $\beta$ or its negation.

Another idea is to retain only those beliefs that are common to all the elements in $\mathcal{K} \perp \alpha$ as a result of contraction of $\mathcal{K}$ by $\alpha$. Such a contraction $\mathcal{K}_\alpha^-$ is termed full meet contraction [2, 29]. It is defined as follows:

Definition 1.2: $\mathcal{K}_\alpha^- = \begin{cases} \bigcap(\mathcal{K} \perp \alpha) & \text{when } \not\vdash \alpha \\ \mathcal{K} & \text{otherwise.} \end{cases}$

In [29] it is shown that as a result of full meet contraction of $\mathcal{K}$ by $\alpha$ only those beliefs that can be logically deduced from $\neg \alpha$ are retained. Thus, a full meet contraction results in a drastic change to the belief set. A full meet revision is obtained via Levi identity where the contraction function is a full meet contraction function. The result of full meet revision of $\mathcal{K}$ by $\alpha$ is $\mathcal{Cn}(\alpha)$. Only those sentences that are deducible from $\alpha$ are believed by the agent. Such a change is very restrictive.

A third construction of contraction is aimed to find a middle ground between the family of maxichoice contraction functions and full meet contraction. This contraction is based on a selection function $S$ which selects a subset of $\mathcal{K} \perp \alpha$ instead of just a single element of $\mathcal{K} \perp \alpha$. A family of contraction functions, namely partial meet contraction [2, 29], is based on the selection function $S$ that selects a subset of $\mathcal{K} \perp \alpha$ and is defined as follows:
Definition 1.3: $K^-_\alpha = \begin{cases} \bigcap S(K_{\perp\alpha}) & \text{when } \alpha \notin \alpha \\ K & \text{otherwise.} \end{cases}$

Partial meet contraction is a generalization of both maxichoice and full meet contractions. The corresponding revision function is termed partial meet revision. When the selection function $S$ returns the set of all elements of $K_{\perp\alpha}$, partial meet contraction reduces to a full meet contraction; while when the selection function returns a single element of $K_{\perp\alpha}$, the contraction reduces to maxichoice contraction. These three contraction operations satisfy the basic AGM postulates.

Moreover, a contraction (revision) function that satisfies all the eight AGM postulates (basic + supplementary postulates) is called a transitivity relational partial meet contraction (revision) [2, 29]. They are so called because, in order to satisfy the supplementary postulates, the selection function $S$ needs to have certain structure. The selection function $S$ is dependent on a relation, $\leq$, among the elements in the set $K_{\perp\alpha}$, that is, $S$ selects only those elements that are minimally related wrt $\leq$. When this relation, $\leq$, satisfies the marking-off identity and transitivity, it is observed that the corresponding contraction satisfies all the AGM postulates.

1.2.3 System of Spheres

Over the years, deficiencies of the AGM theory have been recognised. The major deficiency of the AGM theory is that of iterated belief change. In the AGM framework, the selection function $S$ helps choose which beliefs are retained and which beliefs are lost. In simpler terms, the selection function $S$ represents the preferences of the agent. But the AGM belief change functions do not specify how the selection function changes as a result of belief change, which obstructs the usage of the same belief change function on the new belief state. The AGM belief change functions do not satisfy the principle of categorial matching which is necessary for a belief change function to be iterable.
There are some accounts of belief change that have tried to address this issue [4, 16, 40, 69, 70, 71, 85]. In [40] this problem is overcome by assuming a superselector which assigns a new selection function for each possible belief set. In [4], a method of deriving a changed selection function from the given selection function is proposed, which enables the AGM construction to be iterable. It is argued that the AGM rationality postulates are too permissive to model iterated change and need to be strengthened by additional postulates [16, 51]. In [16, 51] the AGM theory of belief change is extended by adding new postulates which enable giving a plausible account of iterated belief revision.

Iterable belief change functions have been presented in epistemic entrenchment [69, 85] and system of spheres [16, 70, 71, 72] representations. In the epistemic entrenchment representation, the belief state is given by a binary relation over the sentences of the language which is called the entrenchment relation, denoted by \( \preceq \). This relation reflects the preferences among the beliefs of the agent. The relation \( \alpha \preceq \beta \) is read as: \( \alpha \) is at most as entrenched as \( \beta \). Belief change functions in this representation change the initial entrenchment relation into a new entrenchment relation. These functions do not use any external selection function to make the change and hence do not need to specify a new selection function as a result of belief change. Hence they satisfy PCM and therefore are iterable.

In the system of spheres representation, the belief state of the agent is given by a total preorder relation on the set of possible worlds over the language which is generally called a plausibility ordering.\(^4\) In this representation, the belief change functions change the given total preorder relation to a new total preorder relation. Therefore they are iterable.

In chapter 2 of this dissertation, the belief state is represented by a total preorder relation \( \sqsubseteq \). This relation orders the set of possible worlds in accordance to its plausibility as a representation of the real world. The relation \( \omega \sqsubseteq \omega' \) is read as the world \( \omega \) is

\(^4\)A total preorder relation is a totally connected, reflexive and transitive relation; while a partial preorder relation is a partially connected, reflexive and transitive relation. Since every preorder relation is reflexive, the term 'empty preorder' appears to be a wrong usage. But by an 'empty preorder' it is taken to mean a relation where no two different elements are related to each other. If \( \omega, \nu \) are two different elements, then neither \( \omega R \nu \) nor \( \nu R \omega \), where \( R \) is an empty preorder relation. However, both \( \omega R \omega \) and \( \nu R \nu \).
at least as plausible as the world $\omega'$. Those worlds which are considered as the best representations of the real world are the minimally related worlds with respect to $\sqsubseteq$. These are represented by $\min_{\sqsubseteq}(\Omega)$. A total preorder relation can be translated to Grove’s System of spheres as given in [32]. The worlds which are minimally related with respect to $\sqsubseteq$ are the worlds in the inner-most sphere. The sphere immediately enclosing the inner-most sphere includes worlds in the inner-most sphere and those that are next most minimally related with respect to $\sqsubseteq$. A system of spheres is thus built from a total preorder relation, see Figure 1.1. The outer-most sphere includes those worlds which are maximally related with respect to the relation. In a similar fashion, the system of spheres representation can be given as a total preorder relation. The beliefs of the agent are given by the worlds in the inner-most sphere, that is, by the set of minimal worlds under the relation $\sqsubseteq$. In particular, the belief set associated with the belief state $\sqsubseteq$, denoted by $\mathcal{K}_{\sqsubseteq}$, is given by

\[ S_1 = \{K\} \]

The subscript is avoided when there is no ambiguity.
Expansion of the belief state \( \sqsubseteq \) by \( \alpha \) is denoted by \( \sqsubseteq^+ \) and the associated set of beliefs is given by \( K^+ \). The AGM postulates of expansion are satisfied when:

\[
[K^+] = \min_{\sqsubseteq^+}(\Omega) = \min_{\sqsubseteq}(\Omega) \cap [\alpha].
\] (1.1)

This presents a constraint on how the preorder relation should change, yet not completely determine the change to the preorder.

Contraction of \( \sqsubseteq \) by \( \alpha \) is denoted by \( \sqsubseteq^- \). Contraction satisfies all the AGM postulates (including the supplementary postulates) if and only if

\[
[K^-] = \min_{\sqsubseteq^-}(\Omega) = \begin{cases} 
\min_{\sqsubseteq}(\Omega) \cup \min_{\sqsubseteq}[-\alpha] & \text{if } [-\alpha] \neq \emptyset \\
\min_{\sqsubseteq}(\Omega) & \text{otherwise.}
\end{cases}
\] (1.2)

Revision of \( \sqsubseteq \) by \( \alpha \), denoted by \( \sqsubseteq^* \), satisfies all the AGM postulates of revision (including the supplementary postulates) if and only if

\[
[K^*] = \min_{\sqsubseteq^*}(\Omega) = \min_{\sqsubseteq}[\alpha].
\] (1.3)

Again, the constraints on contraction and revision do not determine completely how the preorder relation should be changed.

When it comes to changing the preference orderings, there are numerous ways to do it, each giving rise to a new contraction or revision function. In [85] many belief contraction and belief revision functions based on epistemic entrenchment representation and system of spheres representation are studied. Some of the belief contraction functions given in [85] have been proposed in an earlier work [71] but under different names, such as Moderate or Priority contraction, Natural or Conservative contraction and Lexicographic contraction. The studies into these contraction functions stop with their semantic descriptions, that is, description of how a given preorder relation is changed. The prop-
erties of these contraction functions are only partially known. We provide the axiomatic characterizations of these contraction functions in chapter 2.

1.2.4 KM theory

In the theory of belief change presented by Katsuno and Mendelzon [53], the belief state is represented by belief bases. A belief base is a theory under propositional logic and is denoted by a sentence $\psi$ which is the conjunctions of all the beliefs. They introduce two belief change operations, namely, update and erasure. As described earlier, the operation of update is the inclusion of a belief in a dynamic setting and erasure corresponds to the removal of a belief in a dynamic setting. These belief change operations cannot be accounted for with the aid of the AGM postulates for revision/contraction [53]. The difference between update and revision is explored in detail in [5, 21, 53]. Updating a belief base $\psi$ by a sentence $\mu$, denoted by $\psi \triangleright \mu$, is governed by the following postulates:

U1 $\psi \triangleright \mu$ implies $\mu$.

U2 If $\psi$ implies $\mu$ then $\psi \triangleright \mu$ is equivalent to $\mu$.

U3 If both $\psi$ and $\mu$ are satisfiable then $\psi \triangleright \mu$ is also satisfiable.

U4 If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$ then $(\psi_1 \triangleright \mu_1) \leftrightarrow (\psi_2 \triangleright \mu_2)$.

U5 $(\psi \triangleright \mu) \land \phi$ implies $\psi \triangleright (\mu \land \phi)$.

U6 If $\psi \triangleright \mu_1$ implies $\mu_2$ and $\psi \triangleright \mu_2$ implies $\mu_1$ then $(\psi \triangleright \mu_1) \leftrightarrow (\psi \triangleright \mu_2)$.

U7 If $\psi$ is complete then $(\psi \triangleright \mu_1) \land (\psi \triangleright \mu_2)$ implies $\psi \triangleright (\mu_1 \lor \mu_2)$.

U8 $(\psi_1 \lor \psi_2) \triangleright \mu \leftrightarrow (\psi_1 \triangleright \mu) \lor (\psi_2 \triangleright \mu)$.

The postulates U1 to U5 directly correspond to the postulates of revision given in the AGM theory [2]. Postulate U6 states that if by updating the belief base by $\mu_1$, $\mu_2$ can be deduced and vice-versa, then both updates have the same result. Postulate U7 states that, when $\psi$ is complete, that is, it corresponds to a single world, if some possible world results from updating by $\mu_1$ and also results from updating by $\mu_2$, then it must result from
updating by $\mu_1 \lor \mu_2$. The postulate $\text{U8}$ is termed the disjunction rule. This postulate is regarded as a nonprobabilistic counterpart of the homomorphic condition on the probabilistic revision functions given in [29].

The construction of belief update operation is based on a set of partial preorder relations over the set of possible worlds [53]. We denote these partial preorder relations by $\sqsubseteq$ as well, with appropriate subscripts. It is assumed that there exists a partial preorder relation over the set of all possible worlds, $\Omega$, with respect to each possible world $\omega$, which is denoted by $\sqsubseteq_\omega$. As a result of update of $\psi$ by $\mu$, each model $\omega$ of the belief base $\psi$ is replaced by the minimal models of $\mu$ according to $\sqsubseteq_\omega$. The result of updating is defined as follows:

**Definition 1.4:**

$$[\psi \diamond \mu] = \bigcup_{\omega \in [\psi]} \min_{\sqsubseteq_\omega}[\mu].$$

It is shown in [53] that such a definition of an update operator satisfies all postulates $\text{U1}$ to $\text{U8}$. This contrasts with revision, where the result of revision by $\alpha$ is based on a single total preorder relation over $\Omega$ and the models of the revised belief set is given by the minimal models of $\alpha$ as given in Equation 1.3; while the result of update is based on many partial preorder relations over $\Omega$ and the minimal models are chosen with respect to each prior model of the belief base as shown in Definition 1.4.

Update can be understood as a belief change operation where the agent learns that some sentence is true as a result of some action changing the state of the world. However, uncertainty associated with the occurrence of some action or the state of the world changing could also require removal of a particular belief and admit uncertainty regarding that sentence. Such a change is termed belief erasure. Erasure of the belief base $\psi$ by $\mu$ is denoted by $\psi \blacklozenge \mu$. Again a list of rationality postulates are presented to describe belief erasure [53] and they are as follows:

**E1** $\psi$ implies $\psi \blacklozenge \mu$. 

19
E2 If $\psi$ implies $\neg \mu$ then $\psi \mu$ is equivalent to $\psi$.

E3 If $\psi$ is satisfiable and $\mu$ is not a tautology then $\psi \mu$ does not imply $\mu$.

E4 If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$ then $\psi_1 \mu_1 \leftrightarrow \psi_2 \mu_2$.

E5 $(\psi \mu) \land \mu$ implies $\mu$.

E6 $(\psi_1 \lor \psi_2) \mu$ is equivalent to $(\psi_1 \mu) \lor (\psi_2 \mu)$.

The postulate E2 is significantly different from that of C2; while contraction does not change the belief state when the sentence is not a belief, erasure might change the belief state even when the sentence is not a belief. The postulate E6 is the counterpart of postulate U8. This postulate forms another point of divergence from the notion of contraction. A more detailed study of these postulates can be found in [53].

Erasure of a belief base is defined similar to the Harper identity. Erasure of $\psi$ by sentence $\mu$ is defined as follows:

**Definition 1.5:**

$$\psi \mu \leftrightarrow \psi \lor (\psi \diamond \neg \mu).$$

According to Definition 1.5, erasure of $\psi$ by $\mu$ is defined as a combination of the initial belief base and the result of updating $\psi$ by $\neg \mu$. A belief is removed from the belief state because a sentence which is initially considered certain by the agent is realized to be uncertain. Definition 1.5 suggests that the result of erasure is a mixture of two belief states, namely, a belief state where $\mu$ is certain and the state $\psi \diamond \neg \mu$ where $\neg \mu$ is certain, thus resulting in a state $\psi \mu$ where $\mu$ is uncertain. While belief update has received wide attention, erasure has not. Chapter 3 of this dissertation presents an account of belief erasure in a probabilistic framework. The following section introduces the notion of belief change in a probabilistic framework.
1.3 Probabilistic belief change

Subjective probability assignments are often used to represent the belief state [37, 49, 76]. The probability given to a sentence reflects the agent’s attitude towards the sentence. Bayesian tradition considers only logical truths to have probability 1 and all other sentences to have varying probability between 0 and 1, with falsities having 0 probability. A notion of *full belief* where every belief is certain, that is, has probability 1, even when they are not logical truths is discussed in [26, 60]. A sentence is said to be *accepted* if and only if it has the maximum probability, that is, a probability of 1. A probability assignment is said to be *absurd* or *irregular* when it assigns probability of 1 to every sentence in the language. An agent’s belief state is represented by a given probability function. When the agent makes any observation or receives some information, the probability function which represents its belief state needs to be changed accordingly. The nature of the information and/or observation dictates the type of change the probability function needs to undergo.

1.3.1 Probabilistic static change

To illustrate the different forms of information and/or observations that demand different changes to the probability function we use a simple example.

**Example 1.1:** Adam takes part in a coin toss experiment. He considers that both *heads* and *tails* are equi-probable. The probability function, $\text{prob}$, representing Adam’s belief state assigns equal probability to both *heads* and *tails*.

$$\text{prob}(\text{heads}) = \frac{1}{2}$$

$$\text{prob}(\text{tails}) = \frac{1}{2}.$$  

Before the coin toss, Adam overhears that the coin has been manipulated to favour a *heads* result. How does Adam react to this information?

Upon gaining this information, Adam changes the probability function $\text{prob}$ associ-
ated with his belief state such that the probability assigned to heads is increased and the probability assigned to tails is reduced. Before we discuss this instance we turn our attention to a simpler instance.

**Example 1.2:** Adam overhears that the coin is manipulated to the extent that heads is the only possible result. Hence prob is changed to prob′ such that the probability assigned to heads is made 1, and as a result, the probability assigned to tails is zero.

It is clear that such a transformation of the probability function prob, as expected from Example 1.2, can be modelled by conditionalizing prob with heads. Thus prob′(·) is given by prob(·|heads). Moreover, since a sentence is said to be accepted if it has probability 1, Adam is said to accept that the result of a coin toss will be heads. Thus above described change to prob, that is, belief expansion, corresponds to Bayesian conditionalization.

In the first instance of overhearing (Example 1.1) Adam learns that the coin is manipulated to favour a heads result. The information does not demand that the probability of heads be 1, but suggests that it is greater than $\frac{1}{2}$. This change neither corresponds to addition of a new belief nor to removal of an existing belief. An account of this change to Adam’s belief state, prob, is given by Jeffrey conditionalization [50]. Jeffrey conditionalization gives an account of how the probability function should change when the information received is uncertain. Suppose the agent receives information that the probability of heads should be $\frac{3}{4}$ instead of $\frac{1}{2}$. Changing the probability function with respect

---

6Conditionalization is defined as follows: $\text{prob}(h|e) = \frac{\text{prob}(h \cap e)}{\text{prob}(e)}$ when $\text{prob}(e) > 0$ and is not defined otherwise. The reason why the probabilities associated with the event heads taking place should increase proportionally is as follows. The coin toss could be a part of a bigger experiment the result of which is dependent on the result of the coin toss. For instance, it could be that if the coin toss results in heads, then a fair die is cast and if it is a tails no die is cast. Hence the probability of getting the number two on a die is dependent on the result of coin toss being heads. The probability of getting two upon casting the die increases as a result of probability of heads becoming 1. But the information received does not indicate that two is more plausible than say three and vice-versa. This applies for the other possibilities as well. The result of casting the die still remains equi-probable. Hence in the new probability function, two is still equi-plausible as any other possibility in casting the die. Thus the increase in the probabilities of different events should maintain the original proportionality between them. When heads is the only possible result, the probability of getting two on the die increases proportional to the increase in the probability of heads.
to the information received, \( \text{prob}(\text{heads}) = \frac{3}{4} \) and \( \text{prob}(\text{tails}) = \frac{1}{4} \), does not dictate any changes to the proportion among the different possibilities associated with getting heads and similarly does not dictate any changes to the proportion among the different possibilities associated with getting tails. According to Jeffrey conditionalization, \( \text{prob}' \) is given by the following equation:

\[
\text{prob}'(.) = \text{prob}(.,|\text{heads}) \cdot \text{prob}'(\text{heads}) + \text{prob}(.,|\text{tails}) \cdot \text{prob}'(\text{tails}) \quad (J)
\]

Bayesian conditionalization is a special case of Jeffrey conditionalization. Jeffrey conditionalization assumes that the proportion among the \( \alpha \)-worlds (and \( \neg \alpha \)-worlds) is maintained as before when conditionalizing with respect to \( \alpha \) (uncertain evidence) [17, 18].

Lewis [63] offers an alternative to Bayesian conditionalization, namely, Imaging. It does not assume that the proportion among \( \alpha \)-worlds (and \( \neg \alpha \)-worlds) is maintained, while still ensuring that the evidence is assigned probability 1 as a result of the change. Imaging is discussed in greater detail in Section 1.3.2 and in chapter 3.

The above discussed instances in Example 1.1 and 1.2 are not the only instances where Adam finds it necessary to change his belief state. Sometimes the agent realizes that it is mistaken with some of its beliefs and that those beliefs must be removed from its belief state. Now consider the following development regarding Adam’s coin toss.

**Example 1.3:** After the second instance of overhearing (that is, Example 1.2), Adam has changed the probability function representing his belief state to \( \text{prob}' \) which is the result of conditionalizing \( \text{prob} \) with \( \text{heads} \). Afterwards, Adam’s knowledgeable friend, Ben, informs him that the conversation Adam overheard was aimed to mislead him.

This piece of information can be taken to imply either that the coin was not doctored at all or that the coin was doctored but not to the extent that \( \text{heads} \) is the only possible result. Whichever of these two choices is the intended one, it is clear that the probability
assigned to heads needs to be reduced from 1 to strictly less than 1 and the probability assigned to tails should be increased from zero to a positive non-zero value. It should be noted that such a change, that is, reducing the probability associated with heads from 1 to strictly less than 1, has not received the required attention in the literature except for in [28, 29, 82]. The above described instance corresponds to Adam discarding his belief that the result of a coin toss would be heads. The change from the probability function $prob'$ to a new probability function $prob''$ models removal of an incorrect or a questionable belief, that is, models probabilistic belief contraction.

The change required suggests that the probability associated with heads be reduced from 1. This requires that the probabilities associated with the events which predict a heads result are proportionally reduced. Some probability is assigned to the events which predict a tails result as well. Thus to model probabilistic contraction, a reversal of the effect of Bayesian conditionalization is needed. The certainty associated with a belief is changed to uncertainty, an account of which is modelled in terms of Harper identity. Harper identity in terms of probabilistic representation is given by the following [29]: let $\alpha$ be a belief, that is $P(\alpha) = 1$ and $0 < a < 1$,

$$P_{\alpha}^{-} = PaP_{\neg\alpha}^{*}. \quad (8)$$

Here $P_{\alpha}^{*}$ represents revision of $P$ with respect to the sentence $\neg\alpha$. When $\alpha$ is a belief under $P$, then $\alpha$ has probability 1 under $P$ and hence $P(\neg\alpha) = 0$. Harper identity defines probabilistic contraction in terms of revision by a sentence with initial probability of zero.

Revising by a sentence with probability zero has received a lot of attention from philosophers who lament at the definition of conditional probability in terms of the ratio rule, detailed discussions of which are presented in [38, 67]. Ratio rule is not defined

7Because it does not change the fact that the die is still fair and each of the possibilities of die casting are equi-probable.

8$PaP_{\neg\alpha}^{*} = a \cdot P + (1 - a) \cdot P_{\neg\alpha}^{*}$.
There is a school of thought that conditional probability is more than just the ratio rule and that it is in fact a more primitive notion than the absolute probability [24, 47, 78]. This is reflected in the alternative descriptions of conditional probability presented by them. Popper functions assume conditional probabilities as the primitive notions and that absolute probabilities are derived from it [78]. An advantage with Popper functions is that conditional probabilities of the form $P(\alpha, \beta)$ are defined for every sentence $\beta$. Therefore, when we define a revision operation in terms of Popper functions, revising $P$ by a sentence $\alpha$ is defined as follows:

$$P^*_\alpha(.) = P(., \alpha).$$

A probability function is said to be irregular if and only if it assigns a probability 1 to every sentence in the language. A probability function is said to be regular if it is not irregular. The Popper system does not guarantee that the conditional probability $P(\alpha, \beta)$ is a regular probability function. It is possible that when conditionalizing with some logically consistent sentence $\beta$, $P(., \beta)$ is an irregular function. An alternative system known as Hosiasson-Lindenbaum system (HL-system) given in [47] guarantees that the result of conditionalizing with any logically consistent sentence is a regular probability function. Having defined revision by a sentence which has zero initial probability, the result of contraction by a sentence is straightforward given the Harper identity. However, to exactly prescribe the contracted probability function $P^{-}_\alpha$, it is necessary to know exactly what $P^{-}_\alpha$ is. Since the HL-system only guarantees the existence of a regular probability function but does not specify how to identify the function, we do not know the exact result of revision by $\neg \alpha$ and in turn the Harper identity does not exactly give the contracted probability function. In other words, given a probability function and a belief $\alpha$, the HL-system guarantees the existence of a regular probability function as a result of the

---

9This issue is termed as the ‘zero denominator problem’ in [38] and the logically consistent sentences with zero probability are said to belong to a ‘critical zone’ in [67].
revision, and hence, guarantees the existence of a regular probability function as a result of contraction, in turn. To identify the result of contraction, however, we need to look for some other methods.

The study of probabilistic belief contraction in [28, 29] assumes the existence of an ordering among the set of all possible probability functions on the space. Given such an ordering, $P^*_\alpha$ is chosen as the best probability function $P'$ with $P'(\alpha) = 1$. Again, the result of contraction $P^-_{\lnot \alpha}$ is obtained employing the Harper identity. However, it is difficult to motivate the existence of such an ordering among the probability functions. On the other hand, [82] presents a simple account of belief contraction based on the notion of entropy. Entropy measures have been used in the study of belief change often [25, 33, 35, 56]. Entropy measures give an indication of the uncertainty associated with a probability function. There have been different entropy measures proposed in the literature such as Renyi entropy [83], Hartley entropy [79] and Shannon entropy [89]. Shannon entropy is widely used in the information theory and is defined as follows [89]:

$$H(P) = - \sum_{\omega \in \Omega} P(\omega) \cdot \log P(\omega)$$

The entropy is maximum for a uniform probability distribution and is least for a probability distribution which assigns a probability 1 to a particular world $\omega$.

Contraction results in loss of information. Olsson [74] reads this loss of information as increase in the uncertainty associated with the probability distribution. Hence he postulates that the contracted probability function should have a higher entropy value than the initial probability function. Ramer [82] suggests that the contracted probability function, that is, the result of contraction, should be the one which satisfies the postulates of contraction and also has the maximum possible entropy value. When contracting $P$ by a belief $\alpha$, this results in a probability distribution where every model of $\lnot \alpha$ is assigned equal non-zero probability and the probability associated with the $\alpha$-worlds are reduced proportionally. But this account of contraction corresponds to a full meet contraction
which does not satisfy the principle of informational economy. In chapter 4, we present alternative constructions of contracted probability functions some of which are partial meet contractions and maxichoice contractions.

1.3.2 Probabilistic dynamic change

Changes to the environment or the world described by Adam’s beliefs are capable of forcing Adam to add new beliefs to his belief state as well. Such belief change functions are called dynamic belief change functions. These were introduced in [53] and have been studied extensively, especially the dynamic belief inclusion functions. Dynamic belief inclusion function is termed belief update. It is necessary at this point to differentiate between the term update as used in probabilistic reasoning and update as used in belief change. The term update is used to refer to Bayesian conditionalization in the probabilistic reasoning context. Updating a probability function by a sentence results in conditionalizing the probability function with respect to the sentence. On the other hand, update in the belief change context refers to the belief change function called belief update. Belief update function includes new beliefs to the belief state as a result of possible changes to the world described by the agent’s belief state [53].

As mentioned earlier, belief update is a nonprobabilistic version of probabilistic imaging. Imaging was proposed to give an account of Stalnaker’s conditionals [93].

Lewis [63] assumes that there is a partial preorder relation over \( \Omega \) centred on each possible world \( \omega \). This partial preorder relation \( \sqsubseteq \omega \) represents the similarity relation between two worlds with respect to \( \omega \). The relation \( \upsilon \sqsubseteq \omega \upsilon' \) is taken to mean that the world \( \upsilon \) is at least as similar to the world \( \omega \) as the world \( \upsilon' \). Given a probability distribution \( P \) over the set of all possible worlds, that is, \( \Omega \), imaging \( P \) by a sentence \( \alpha \), denoted by \( P^\#_\alpha \), changes \( P \)

\(^{10}\)Slatnaker suggests that in order to evaluate the truth value of a conditional sentence \( \alpha \supset \beta \), change the given set of beliefs minimally in order to accommodate the antecedent \( \alpha \), while ensuring the consistency, and check whether the consequent \( \beta \) is then true [93].

\(^{11}\)Stalnaker’s analysis of conditional sentences thus amounts to the following: given a world \( \omega \), the truth value of a conditional sentence \( \alpha \supset \beta \) is determined by the truth value of the consequent \( \beta \) in an \( \alpha \)-world that is most similar to \( \omega \).
such that the probability associated with $\alpha$ is increased to 1. The probability associated with each world $\omega$ is moved to their corresponding most similar $\alpha$-world as a result of imaging by $\alpha$. Due to such a re-distribution of probability, imaging does not preserve the initial proportion among the $\alpha$-worlds, except for in the trivial case as shown in Lewis’ triviality result [63]. When the belief state, represented by the probability function $P$, undergoes belief update by a sentence $\alpha$, the resultant state is given by $P_{\alpha}^\#$.

To illustrate an instance of belief erasure, let us go back to the second instance of the overhearing (Example 1.2), after which the probability function $prob'$ assigns a probability of 1 to $heads$.

**Example 1.4:** Ben, Adam’s knowledgeable friend, informs Adam that the coin-tosser was seen dropping the coin which was to be used for the coin toss experiment amidst a pile of coins of same denomination. As a result the coin to be used for coin toss could be a different one from the doctored coin and hence the result need not be $heads$. This information demands that Adam change his belief that the coin toss would result in $heads$.

The change here is demanded by the possible change to the environment described by Adam’s beliefs. To the best of my knowledge, no known method of probabilistic change can be applied to model the change in Example 1.4. There is no account of probabilistic erasure. It is clear that the study of probabilistic belief change is incomplete without an account of probabilistic contraction and probabilistic erasure. Chapter 3 and 4 address this problem and provide a solution for probabilistic erasure and contraction, respectively.

### 1.4 Research contribution

The discussion put forward in the preceding sections highlights three open problems in the study of belief removal. They are:
1 Complete axiomatic characterizations of Moderate or Priority contraction, Conservative or Natural contraction and Lexicographic contraction functions have not been provided,

2 An account of probabilistic belief erasure has not been given, and

3 An account of probabilistic belief contraction that satisfies the principle of informational economy is lacking.

This dissertation is a compilation of three works that address these issues. I briefly discuss them in the following subsections.

1.4.1 Iterated belief contraction

A contraction function on a belief state is said to be AGM-rational if it satisfies the contraction postulates given in [2]. When the belief state is given in terms of a system of spheres, or equivalently, in terms of a total preorder relation over the set of all possible worlds $\Omega$, Equation 1.2 gives the necessary condition for a contraction function to be AGM-rational. But Equation 1.2 is not strong enough to give a unique contracted belief state. There are many ways of changing the initial preorder relation, each giving rise to a new contraction. The priority or moderate contraction, the natural or conservative contraction and the lexicographic contraction functions [71, 85] change the initial preorder relation in uniquely. The properties of these contraction functions have not been studied comprehensively. There was an attempt to completely axiomatize one of these contraction functions, namely lexicographic contraction, in [70]. However, the attempt was only partially successful. This attempt showed that a lexicographic contraction function satisfies a condition termed principled factored insertion apart from the regular AGM postulates of contraction.

In [80] we show that indeed all the three contraction functions satisfy the property of principled factored insertion. We build on the notion of principled factored insertion and manage to completely axiomatize priority or moderate contraction function and natural
or conservative contraction function. But for characterizing lexicographic contraction, we find the need for more tools. Consequently, we define a numerical notion of degree of belief among the sentences in the language based on this preference relation. This notion, that we introduce, is observed to be a counter-part of the entrenchment ranking function proposed by Hans Rott [86]. With the help of a more restricted notion of conditional degree of belief we manage to characterize the lexicographic contraction function. This work is presented in chapter 2. The proofs of results in chapter 2 are presented in Appendix A.

1.4.2 Probabilistic belief erasure

While probabilistic belief update is accounted for by Lewis’ imaging [63], an account of probabilistic belief erasure remains an open problem. Towards solving this problem, we present a notion of partial imaging in chapter 3. Partial imaging is a generalization of Lewis’ imaging; it also mirrors the application of Harper identity over imaging. We find that partial imaging satisfies all the equivalent formalizations of postulates of belief erasure in a probabilistic framework. Furthermore, partial imaging is shown to generalize imaging just as Jeffrey conditionalization generalizes Bayesian conditionalization.

Our study of generalizations of imaging leads to the introduction of another generalization of imaging, namely selective imaging. Selective imaging changes the given probability function in a unique way which does not correspond to any known belief change operation. We show that, in fact, selective imaging gives a probabilistic account of what we term as conditional update. Often we handle instructions, that is, sentences that suggest an action. A belief state undergoes conditional update upon receiving a conditional instruction, namely, an instruction to perform an action subject to some condition. We show that the change required for a conditional update is modelled by selective imaging.

A combination of the notions of partial imaging and selective imaging gives rise to a more generalized version of imaging, namely selective partial imaging. We find that se-
lective partial imaging gives an account of conditional erasure, the belief removal counterpart of conditional update. This work is presented in chapter 3 and the proofs of the results from this chapter are presented in Appendix B.

1.4.3 Probabilistic belief contraction

Ramer [82] presents an elegant account of probabilistic contraction based on principle of maximum entropy [90]. According to this proposal, the result of contraction is given by the probability function that has maximum possible entropy value among all those probability functions that satisfy the postulates of contraction [29]. In chapter 4, we show that such an account of probabilistic contraction corresponds to a full meet contraction [2, 29], resulting in excessive loss of beliefs. By forcing various constraints on the Shannon entropy value of the contracted probability function, we are able to get different solutions to the problem of probabilistic contraction. We obtain, as a result of minimizing the entropy value of the contracted probability function, a contraction that is similar to that of a maxichoice contraction. However, such a contraction does not satisfy Olsson’s postulate [74]. We introduce a notion of submaximal entropy contraction, wherein we search for a probability function that satisfies the postulates of contraction but has entropy value less than the maximum possible value.

We show that we need extra-logical constraints in order to give an account of partial meet contraction. This is because partial meet contraction is based on making a choice between beliefs that cannot co-exist in the belief state if the belief change were to be successful. Such a choice is made depending on the agent’s preference. But a probabilistic representation is not rich enough to incorporate a preference relation among the full beliefs, that is, sentences with probability 1. Therefore, we make use of Grove’s system of spheres [32] to represent the agent’s preferences among the beliefs. With the help of this preference relation, we give a cogent account of partial meet contraction.

All the above mentioned contractions are based on Shannon entropy measure. We also
present a discussion on the counterparts of these contractions based on Hartley entropy measure [1, 79]. The proofs of the results from this chapter are presented in Appendix C.

1.5 Concluding remarks

In this dissertation, we have highlighted various aspects associated with belief removal that need to be studied in greater detail. Belief removal, simply put, is the process of discarding questionable beliefs. Everyone needs to perform belief removal at some time or other. The import of belief removal in artificial intelligence cannot be unduly stressed.

In the process of our investigations into belief removal, we have worked on different representations of the belief state and have used different theories and tools to propose an answer to the problem at hand. We have answered some open problems, but have in turn opened up a few more. We aim to take up these problems in the future research.

In chapter 2 we work with a system of spheres representation of the belief state. Our aim here is to completely characterize the three contraction functions, namely, priority contraction, natural contraction and lexicographic contraction. We are successful in this regard. We characterize these contraction functions based on a notion of ‘degrees of belief’. We also find that the priority contraction and natural contraction can be characterized even without this notion. Curiously, neither are we able to give a characterization of lexicographic contraction without the aid of this notion nor are we able find an explanation for why this is so.

Chapter 3 is concerned with giving an account of probabilistic belief removal. Here we work with a probabilistic representation where the belief state is given in terms of a probability distribution over the set of all possible worlds. In this work, we present three different generalizations of Lewis’ imaging, namely, partial imaging, selective imaging and selective partial imaging. We show that partial imaging can be used to model an account of belief erasure. Moreover, we introduce two new notions, that of conditional update and conditional erasure. We argue that these two belief change operations can be
modelled with the help of selective imaging and selective partial imaging, respectively. Conditional update and conditional erasure, clearly, play an important role in the field of artificial intelligence, and their various facets need to be explored.

The chapter 4 of this dissertation also assumes a probabilistic representation of the belief state. In this chapter, we begin by re-introducing indifferent contraction, originally due to Ramer. Our study of indifferent contraction shows that it corresponds to a probabilistic version of full meet contraction. We then propose a few alternative accounts of contraction, namely, submaximal entropy contraction, minimal contraction and preferential contraction. In our discussion of submaximal entropy contraction, we were unable to propose a numerical constraint on the maximum number of worlds over which the given probability distribution is spread out. Such a constraint could make an account of probabilistic partial meet contraction simple.

Instead, we propose preferential contraction as an account of probabilistic partial meet contraction by using an “enriched” representation of the belief state. In this “enriched” belief state, the beliefs of the agent are obtained from the probability distribution, and we make use of a total preorder relation over the set of possible worlds to indirectly represent a preference relation among these beliefs. However, we do not investigate the properties of this “enriched” belief state in detail. It would be interesting to study how a contraction function on this belief state would change the total preorder relation, if it does.
Chapter 2

Iterated Belief Contraction

Lo! those who believe, then disbelieve and then (again) believe, then disbelieve, and then increase in disbelief, Allah will never pardon them, nor will He guide them unto a way.

(Sura 4:137).

2.1 Introduction

It is now taken for granted that any rational theory of belief change, such as the classic AGM theory [2] and those that extended it in various directions, must provide accounts of how new beliefs are added (belief expansion), old beliefs are removed (belief contraction), as well as how current beliefs are modified in light of new information that is deemed accepted (belief revision). One of the problems that the classical AGM account has been extended to deal with is the problem of iterated belief change – the problem of how an agent is supposed to continually modify its beliefs in light of a sequence of observations made (or pieces of information received). However, for some reason or other, the emphasis in this context has historically been on iterated revision; the problem of iterated contraction having received a rather step-motherly treatment. In more recent times, however, the research community appears to be taking measures to redress
this inequity [46, 70, 72, 85]. Three approaches to such an account examined by Nayak and colleagues [70, 72] are called the Natural Contraction, the Priority Contraction and the Lexicographic Contraction.\(^2\) The accounts provided of these operations in the literature are by and large semantic; and the attempt made to characterize the Lexicographic Contraction via rationality postulates in [70] is rather incomplete. In this paper we aim to characterize these three operations. Unlike the Natural and Priority contractions, we found Lexicographic contraction difficult to characterize in the expected fashion using rationality postulates of the sort well known in the literature starting from [2]. We instead use a more “information-rich” measure that we call degree of belief for this purpose.

Throughout this paper we will assume a propositional object language \(\mathcal{L}\) generated from a finite alphabet, whose sentences, denoted by lower case Greek letters such as \(\alpha\) and \(\beta\), with or without decorations, will be used to represent individual beliefs. Sets of such sentences will be represented by uppercase Latin letters, such as \(A, B, \ldots\); in particular, \(K\) with or without decorations, will be used to denote sets of beliefs. For the sake of simplicity we take the background logic governing this language to be the classical propositional logic, identified with the classical deducibility relation \(\vdash\).\(^3\) Given a set \(A\) of sentences, we will represent by \([A]\) the set of models of \(A\); and for readability, given a sentence \(\alpha\), the set of models \([\{\alpha\}]\) will be represented simply by \(][\\alpha]\). Other such measures will be adopted for readability where no confusion is imminent; for instance, for some function \(f\) and sentence \(\alpha\), the expression \(f([\alpha])\) will be simplified to \(f[\alpha]\), whose official representation is in fact \(f([\{\alpha\}])\). We will be using the words “interpretation” and “world” interchangeably, with the understanding that models of some sentence \(\alpha\) are the worlds that satisfy \(\alpha\); individual worlds would be denoted by \(\omega\) with possible decorations, the set of all worlds will be denoted by \(\Omega\), and subsets of \(\Omega\) by upper case Greek letters such as \(\Gamma\) and \(\Delta\).

\(^2\)Please note that the Natural Contraction and the Priority Contraction are respectively termed the Conservative Contraction and the Moderate Contraction in [85].

\(^3\)Alternatively, as standard in the literature, we can take the background logic to be a propositional, supra-classical logic satisfying the deduction theorem and compactness.
As is standard, we will assume that a belief set, that is the body of beliefs of an epistemic agent, is represented by a set of sentences, typically \( K \) that is closed under \( \vdash \), that is, \( K = \{ \alpha \in L \mid K \vdash \alpha \} \). A belief set such as \( K \) is distinguished from a belief state such as \( K \): the latter is a richer representation of the relevant information. In particular, we assume that the belief set \( K_K \) associated with a belief state \( K \) can be extracted from the latter using some appropriate operation, say, \( \text{bel} \); thus, \( \text{bel}(K) = K_K \). Furthermore, we assume that the belief state \( K \) incorporates a relevant belief (set) contraction operation, say \( -K \) that determines the outcome \( K' = -K(K, \alpha) \) of removing some information \( \alpha \) from the associated belief set \( K_K \). In general, in response to the removal of \( \alpha \), the belief state \( K \) will undergo modification to, say, \( K' = -(K, \alpha) \) such that the new belief state \( K' \) and the new belief set \( K' \) are appropriately aligned, that is, \( \text{bel}(K') = K' \) such that further belief removal can be carried out as and when necessary.

In the literature, belief states have been represented in many different ways, including a relational epistemic entrenchment measure [29], a numerical possibility measure [19] and an ordinal ranking function [92]. In this paper we will represent a belief state as a total preorder (that is, a reflexive, connected and transitive relation) \( \sqsubseteq \) over \( \Omega \), with the understanding that \( \omega \sqsubseteq \omega' \) means \( \omega \) is at least as plausible as \( \omega' \). It is well known in the literature that a total preorder such as \( \sqsubseteq \) can be directly translated to Grove’s system of spheres as propounded in [32] and vice versa. Furthermore, the plausibility measure \( \sqsubseteq \) over \( \Omega \) and the entrenchment measure \( \preceq \) over sentences of \( L \) are inter-translatable.\(^4\) We will denote by \( \sqsubset \) the strict part of \( \sqsubseteq \), and by \( \approx \) its symmetric part. It is worth noting that on occasion we will need to refer to modified plausibility preorders such as \( \sqsubseteq_{\alpha} \); and in

\(^4\)Epistemic entrenchment is a relational measure, roughly indicating how hard it is to remove a given belief. It is typically defined as a binary relation \( \preceq \) over sentences of a language satisfying certain standard conditions such as those provided in [29]. How an entrenchment relation can be obtained from a given plausibility preorder is well-known in the literature: for all sentences \( \alpha, \beta \in L \), \( \alpha \preceq \beta \) iff \( \omega \sqsubseteq \omega' \) for all \( \omega \in \text{min}_{\sqsubseteq} [\neg \alpha] \) and \( \omega' \in \text{min}_{\sqsubseteq} [\neg \beta] \). The other side, how the plausibility preorder can be obtained from a given entrenchment relation can be found on page 252 of [84] in a slightly different framework. In our notation, it would be: for all worlds \( \omega, \omega' \in \Omega \), \( \omega \sqsubseteq \omega' \) iff for every sentence \( \alpha \in L \) such that \( \omega \models \neg \alpha \), there exists a sentence \( \alpha' \in L \) such that both \( \omega' \models \neg \alpha' \) and \( \alpha \preceq \alpha' \). The proof is easily verified. For \((\Rightarrow)\), assume \( \omega \sqsubseteq \omega' \), \( \omega \in [\neg \alpha] \) and set \([\neg \alpha'] = \{ \omega' \} \). For \((\Leftarrow)\), let \([\neg \alpha] = \{ \omega \}, \alpha \preceq \alpha' \) and \( \omega' \in [\neg \alpha'] \). The proof will use the well known definition of \( \preceq \) via \( \sqsubseteq \) just mentioned.
such cases we will refer to their strict part and symmetric part by $\sqsubseteq^\sim$ and $\approx^\sim$ respectively.

Given any set of worlds $\Delta \subseteq \Omega$ we will denote by $\text{min}_{\sqsubseteq}(\Delta) = \{\omega \in \Delta \mid \omega \sqsubseteq \omega', \text{ for all } \omega' \in \Delta\}$ the set of $\sqsubseteq$-minimal worlds of $\Delta$ that the epistemic agent considers most plausible among those in $\Delta$. In particular, $\text{min}_{\sqsubseteq}(\Omega)$ will represent the set of most plausible worlds among all possible worlds as viewed by the agent and $\text{min}_{\sqsubseteq^\sim}(\Omega)$ will represent the most plausible worlds after it has removed information $\alpha$ from its belief state (set). The beliefs of the agent are those that are true in all these most plausible worlds. This is captured in the following equation that we name after Grove. Note that from here onwards, the set of beliefs $\text{bel}(\sqsubseteq)$ extracted from the belief state $\sqsubseteq$ will be denoted by $K_{\sqsubseteq}$.

$$\text{bel}(\sqsubseteq) = K_{\sqsubseteq} = \{\alpha \in \mathcal{L} \mid \text{min}_{\sqsubseteq}(\Omega) \subseteq [\alpha]\} \quad (G1)$$

At this point we seek to clarify certain notational device that is potentially confusing. The contraction operation — that we are concerned with is a state contraction operation: given a belief state $\sqsubseteq$ and a sentence $\alpha$ to be removed from it, this operation returns a belief state $\sqsubseteq^\sim$ in which, in general, $\alpha$ is not believed. The corresponding propositional content of these two belief states, that is the associated belief sets, are $K_{\sqsubseteq}$ and $K_{\sqsubseteq^\sim}$. When the prior state $\sqsubseteq$ can be contextually determined, and the intended reading is clear, for the sake of simplicity we will refer to these belief sets as $K$ and $K_{\sqsubseteq^\sim}$ instead. In a similar fashion, $(K_{\sqsubseteq^\sim})_{\bar{\beta}}$ will refer to the belief set associated with the belief state $(\sqsubseteq^\sim)_{\bar{\beta}}$. In other words, the same symbol — is used for both state contraction operation as in $K_{\sqsubseteq^\sim}$, as well as the corresponding set contraction operation $-_{\sqsubseteq}$ (as in $K_{\sqsubseteq^\sim}$, which should officially be written as: $K_{\sqsubseteq^\sim}$), but since the belief set and belief state in question are assumed to be appropriately co-related, it is not problematic.

An inconsistent belief state is represented by an empty relation $\sqsubseteq_{\bot}$: for any two distinct worlds $\omega, \omega' \in \Omega$, $\omega \not\sqsubseteq_{\bot} \omega'$. On the other hand $\sqsubseteq_{\top}$ denotes the full pre-order relation: for every $\omega, \omega' \in \Omega$, $\omega \sqsubseteq_{\top} \omega'$. The relation $\sqsubseteq_{\top}$ denotes the belief state where

5Here by empty preorder we mean a reflexive and transitive relation which is "completely disconnected", that is, no two distinct worlds are related to each other.
the agent believes only in logical tautologies and the relation \( \sqsubseteq \) denotes the state where the agent believes in every sentence.

A notion that we would be using throughout this paper is that of a chain of worlds. Given a belief state \( \sqsubseteq \) and a set of worlds \( \Delta \subseteq \Omega \), a chain of worlds in \( \Delta \) is a sequence of worlds in \( \Delta \) ordered by the strict part \( \sqsubset \) of \( \sqsubseteq \), e.g., \( \omega_0 \sqsubset \omega_1 \sqsubset \ldots \sqsubset \omega_n \). Based on this notion, we define a complete chain of worlds in \( \Delta \) as follows. A chain of worlds \( \omega_0 \sqsubset \omega_1 \sqsubset \ldots \sqsubset \omega_n \) in \( \Delta \) is said to be complete iff \( \omega_0 \) is a \( \sqsubseteq \)-minimal world of \( \Delta \) and for any \( \omega_{i-1} \sqsubset \omega_i \) \( (1 \leq i \leq n) \) there does not exist any world \( \omega' \) in \( \Delta \) such that \( \omega_{i-1} \sqsubset \omega' \sqsubset \omega_i \). The length of a complete chain of worlds \( C = \omega_0 \sqsubset \omega_1 \sqsubset \ldots \sqsubset \omega_n \) is defined to be \( n \), and is denoted by \( |C| \). Two special cases of this notion are of interest: when \( \Delta \) is the universal set \( \Omega \), and when \( \Delta \) is exactly the set of models \( [\alpha] \) of some sentence \( \alpha \).

2.2 Three Contraction Functions

The modified preorder \( \sqsubseteq^- \) represents the result of contracting the prior belief state \( \sqsubseteq \) by a sentence \( \alpha \). From Equation G1 it is evident that, if \( \alpha \) is to be successfully removed from the corresponding belief set in this process, then there must exist at least one model of \( \neg \alpha \) among the minimal worlds in \( \sqsubseteq^- \). This would mean that \( \text{min}_{\sqsubseteq^-}(\Omega) \) is not contained in \( [\alpha] \) and hence \( \alpha \) is not retained. We will say that a belief state contraction operation \( \neg \) is AGM rational just in case the corresponding belief set contraction operation satisfies the standard AGM postulates [2]. Combining the semantic account of belief (set) contraction provided in [16, 29, 32] with the account of belief (state) contraction we are espousing, we get that a belief state contraction operation \( \neg \) is AGM rational when

\[
\text{min}_{\sqsubseteq^-}(\Omega) = \text{min}_\sqsubseteq(\Omega) \cup \text{min}_\sqsubseteq[\neg \alpha].
\] (G2)
Henceforth in this paper, any reference to an AGM-rational contraction function will refer either to a state contraction function which satisfies Equation G2, or a belief set contraction operation appropriately obtained from such a state contraction function. The context should sufficiently disambiguate the usage.

It is to be noted that the condition G2 on $\sqsubseteq_\alpha$ does not guarantee a unique way of changing $\sqsubseteq$ to $\sqsubseteq_\alpha$. Different ways of changing $\sqsubseteq$ have been proposed, each giving a different contraction function. Three such contraction functions, namely the moderate contraction, the natural contraction and the lexicographic contraction have been proposed in [71, 72]. In this section we briefly outline these functions.

### 2.2.1 Moderate contraction function

This contraction function is referred to as *moderate contraction* function in [85] and as *priority contraction* function in [71, 72]. The basic idea behind moderate contraction is simple. Suppose we want to remove a non-trivial belief $\alpha$ (that is, $\not\vdash \alpha$). When we remove $\alpha$ from a belief set, under usual circumstances, $\alpha$ becomes a non-belief. Every sentence of the form $\beta \rightarrow \alpha$ can be viewed as contributing to the agent’s epistemic attitude towards $\alpha$. This demands that out of every pair $\beta \rightarrow \alpha$ and $\neg \beta \rightarrow \alpha$, at least one be removed.

The question is, what happens to those sentences $\beta \rightarrow \alpha$ which are not removed from the belief set in the process. Moderate contraction is based on the intuition that even if some sentence $\beta \rightarrow \alpha$ is not removed via removal of $\alpha$, its entrenchment should be reduced, and its status be demoted. The degree of entrenchment of $\beta \rightarrow \alpha$ for some particular $\beta$ is reflected by the most plausible worlds in $[\beta] \cap [\neg \alpha]$ in the belief state $\sqsubseteq$. By promoting all the worlds in $[\neg \alpha]$ the entrenchment of every $\beta \rightarrow \alpha$ is reduced. This is captured by the condition MC3 in the following formalization.

Given a total pre-order relation $\sqsubseteq$ on $\Omega$, and a sentence $\alpha$ to be removed from the belief set, the degree of belief on the sentences of the language. We can show that the degree of belief of the sentences of the form $\beta \rightarrow \alpha$ decreases when $\alpha$ is removed using moderate contraction.
associated belief set, a moderate contraction function — changes the relation $\sqsubseteq$ to a new relation $\sqsubseteq^-$. If $\alpha$ is not believed in the state represented by the prior $\sqsubseteq$, or it is a tautology, then no change is effected in $\sqsubseteq$, that is, $\sqsubseteq^- = \sqsubseteq$. In the principal, non-trivial case, where $\alpha \in K$ and $\not\forall \alpha$, the new total pre-order $\sqsubseteq^-$ satisfies the following conditions: for any $\omega_1, \omega_2 \in \Omega$,

**MC1** When $\omega_1 \models \alpha$ and $\omega_2 \models \alpha$ then $\omega_1 \sqsubseteq^- \omega_2$ if and only if $\omega_1 \sqsubseteq \omega_2$.

**MC2** When $\omega_1 \models \neg \alpha$ and $\omega_2 \models \neg \alpha$ then $\omega_1 \sqsubseteq^- \omega_2$ if and only if $\omega_1 \sqsubseteq \omega_2$.

**MC3** When $\omega_1 \models \alpha$, $\omega_1 \not\in \min_{\sqsubseteq}[\alpha]$ and $\omega_2 \models \neg \alpha$ then $\omega_2 \sqsubseteq^- \omega_1$.

**MC4** When $\omega_1 \in \min_{\sqsubseteq}[\Omega]$ or $\omega_1 \in \min_{\sqsubseteq}[\neg \alpha]$, then $\omega_1 \sqsubseteq^- \omega_2$, for any $\omega_2 \in \Omega$.

It is clear from conditions **MC1** and **MC2** that the relative plausibility of worlds in $[\alpha]$ (and respectively in $[\neg \alpha]$) are not affected. We come across such invariance a number of times in this paper, and need a name for it. Accordingly we introduce the notion of *Order Preservation* in Section 2.2.4 which is closely related to the postulates of iterated revision (CR1 and CR2) as presented in [16]. Condition **MC3** captures the uniform promotion of worlds in $[\neg \alpha]$ and condition **MC4** reflects G2. The change to the preorder relation as prescribed here is presented pictorially in Figure 2.1(a). In Figure 2.1 preorder relations are depicted as systems of spheres since they are equivalent, and representation of preorders as systems of spheres is well understood.

### 2.2.2 Natural contraction function

Boutilier [7] introduced *natural revision* in an attempt to account for iterated belief revision while remaining true to the dictum of “minimal change” to the belief state. Upon receiving information $\gamma$, natural revision changes the belief state $\sqsubseteq$ such that only the $\min_{\sqsubseteq}[\gamma]$ worlds become the minimal worlds in the revised belief state; no other change is made to the relation $\sqsubseteq$. Formally the natural revision function is defined as follows: considering the non-trivial cases, when $[\gamma] \neq \emptyset$ and $\not\sqsubseteq \not\sqsubseteq_\perp$,

**NR1** If $\omega_1 \in \min_{\sqsubseteq}[\gamma]$, and $\omega_2 \not\in \min_{\sqsubseteq}[\gamma]$, then $\omega_1 \sqsubseteq^- \gamma \omega_2$. 

41
NR2 If $\omega_1 \not\in \min_{[\gamma]} \Omega$ and $\omega_2 \not\in \min_{[\gamma]} \Omega$, then $\omega_1 \sqsubseteq \omega_2$ iff $\omega_1 \sqsubseteq \omega_2$.

NR3 If $\omega_1, \omega_2 \in \min_{[\gamma]} \Omega$, then $\omega_1 \approx_{[\gamma]} \omega_2$.

Natural contraction follows natural revision in changing the relation on $\Omega$ only with respect to the most plausible worlds in $[\neg \alpha]$ when contracting by $\alpha$. Given a total pre-order relation $\sqsubseteq$ on $\Omega$, and a sentence $\alpha$, the changed total pre-order relation $\sqsubseteq_{\alpha}$ satisfies the following conditions: for any $\omega_1, \omega_2 \in \Omega$,

NC1 If $\omega_1 \in \min_{[\alpha]} \Omega$ or $\omega_1 \in \min_{[\neg \alpha]} \Omega$, then $\omega_1 \sqsubseteq_{\alpha} \omega_2$, for any $\omega_2 \in \Omega$.

NC2 If $\omega_1, \omega_2 \notin \min_{[\alpha]} \Omega$ and $\omega_1, \omega_2 \notin \min_{[\neg \alpha]} \Omega$ then $\omega_1 \sqsubseteq_{\alpha} \omega_2$ iff $\omega_1 \sqsubseteq \omega_2$.

Natural contraction is pictorially depicted in Figure 2.1(b). It is clear that since a natural contraction function satisfies G2, the belief set corresponding to the contracted belief state will satisfy the standard AGM contraction postulates given in [2, 29]. The relative orderings of worlds in $[\alpha]$ (and respectively in $[\neg \alpha]$) are not changed; thus it satisfies Order Preservation (which we discuss in Section 2.2.4) just like the moderate contraction. It should be noted that Rott [85] has provided an equivalent definition of natural contraction when the belief state of the agent is given by an epistemic entrenchment relation.

2.2.3 Lexicographic contraction function

The Harper identity gives a relation between contraction and revision [45]. It states that $\mathcal{K}_{\alpha} = \mathcal{K} \cap \mathcal{K}_{\neg \alpha}$ [29]. Semantically the Harper identity may be taken to say that the $\sqsubseteq$-minimal worlds of $\Omega$ and the $\sqsubseteq_{\alpha}$-minimal worlds of $[\neg \alpha]$ are to be given equivalent status in the state resulting from the contraction of $\sqsubseteq$ by $\alpha$. We present below a mildly modified version of the generalized Harper identity that was presented in [72]:

**Generalized Harper Identity.** In the trivial case, where $\alpha \not\in \mathcal{K}_{\sqsubseteq}$, the preorder $\sqsubseteq$ does not change under contraction by $\alpha$. As to the principal case, let $B_i, 0 \leq i \leq n - 1$ be the $n$ bands ($\sqsubseteq$-equivalence classes) of worlds.

7The modification in question deals with certain emptysets that were not properly dealt with in the version presented in [72].
generated by the pre-contraction state \( \sqsubseteq \), where \( B_0 \) consists of the \( \sqsubseteq \)-minimal worlds in \( \Omega \) and \( \omega \sqsubseteq \omega' \) for all \( \omega \in B_i, \omega' \in B_j \) and \( i < j \). Let \( C_i, \; 0 \leq i < k \leq n \) be the \( k \) \( \sqsubseteq \)-equivalent classes of worlds in \( \neg \alpha \), that is, \( \bigcup_{i=0}^{k-1} C_i = [\neg \alpha] \), and \( \omega \sqsubseteq \omega' \) for all \( \omega \in C_i, \omega' \in C_j \) and \( i < j \). Similarly, let \( C'_i, \; 0 \leq i < k' \leq n \) be the \( k' \) \( \sqsubseteq \)-equivalent classes of worlds in \( [\alpha] \), that is, \( \bigcup_{i=0}^{k'-1} C'_i = [\alpha] \), and \( \omega \sqsubseteq \omega' \) for all \( \omega \in C'_i, \omega' \in C'_j \) and \( i < j \). Define \( C_{i+1} = \emptyset \) for \( k-1 \leq i < k'-1 \), if \( k < k' \); otherwise, if \( k > k' \), then define \( C'_{i+1} = \emptyset \) for \( k'-1 \leq i < k-1 \). The bands in \( \sqsubseteq_{\alpha} \) are given by \( D_i = C_i \cup C'_i \) for \( 0 \leq i < \max(k, k') - 1 \).

It can be easily noted that neither moderate nor natural contraction function satisfies the generalized Harper identity. On the other hand, lexicographic contraction is defined to satisfy the generalized Harper identity. Given a total pre-order relation \( \sqsubseteq \) on \( \Omega \) and a belief \( \alpha \in K \), when contracting \( \alpha \) from \( K \), the lexicographic contraction function \( \sqsubseteq_{\alpha} \) changes the total pre-order to \( \sqsubseteq_{\alpha} \). The changed relation \( \sqsubseteq_{\alpha} \) is given as follows:

\[ \text{LC1} \quad \text{If } \omega \models \alpha \text{ and } \omega' \models \alpha, \text{ then } \omega \sqsubseteq_{\alpha} \omega' \text{ iff } \omega \sqsubseteq \omega' \]

\[ \text{LC2} \quad \text{If } \omega \models \neg \alpha \text{ and } \omega' \models \neg \alpha, \text{ then } \omega \sqsubseteq_{\alpha} \omega' \text{ iff } \omega \sqsubseteq \omega' \]

\[ \text{LC3} \quad \text{Let } \chi \text{ be one member of } \{ \alpha, \neg \alpha \} \text{ and } \overline{\chi} \text{ the other. If } \omega \models \chi \text{ and } \omega' \models \overline{\chi}, \text{ then } \omega \sqsubseteq_{\alpha} \omega' \text{ iff the length of a complete chain of worlds in } [\chi] \text{ which ends in } \omega \text{ is less than or equal to the length of a complete chain of worlds in } \overline{\chi} \text{ which ends in } \omega'. \]

For the special case where \( \alpha \notin \mathcal{K} \) or \( \vdash \alpha \), the lexicographic contraction results in an unchanged pre-order relation, \( \sqsubseteq_{\alpha} = \sqsubseteq \). Conditions LC1 and LC2 state that the prior ordering between any two worlds \( \omega, \omega' \) that both satisfy \( \alpha \) (or \( \neg \alpha \)) is not changed upon contraction by \( \alpha \). LC3 states that the worlds of \( [\neg \alpha] \) are simultaneously shifted so that the worlds in \( [\neg \alpha] \) and worlds in \( [\alpha] \) which are ranked equally within the respective sets (after removing the empty sets) are equally placed in the preorder resulting from contraction. A pictorial representation of lexicographic contraction is given in Figure 2.1(c).

\[ ^8 \text{Please note that empty "layers" are ignored in the process.} \]

\[ ^9 \text{This is equivalent to the original condition presented in [70].} \]
Figure 2.1: (a) Priority/Moderate contraction, (b) Natural/Conservative contraction, and (c) Lexicographic contraction. The numbers indicate how the “system of spheres” should be constructed after contraction by \( \alpha \). The cells numbered 1 would jointly constitute the centre; the ones numbered 2 will be the next layer, and so on.

### 2.2.4 Comparison of the three contraction functions

Moderate, natural and lexicographic contraction functions offer three different ways of changing the total preorder on \( \Omega \). They share some similarities. The three contraction functions behave in the same way when the sentence \( \alpha \) (which is being contracted) is either a logical tautology or is not an existing belief of the agent. In the non-trivial case, where \( \alpha \) is a contingent belief, that is, both \( \alpha \in \mathcal{K} \) and \( \vdash \alpha \), all three of them preserve the ordering of the worlds in \([\alpha]\) after contraction. Furthermore, in all the three cases of contraction, ordering of worlds in \([\neg \alpha]\) is also preserved. We call these properties as \textit{Order Preservation} in \([\alpha]\) and \textit{Order Preservation} in \([\neg \alpha]\).

**Definition 2.1 (Order Preservation):** Let \( \sqsubseteq \) be a total preorder on \( \Omega \) representing the belief state, and \( \alpha \) be a sentence. A contraction function \( - \) is said to obey \textit{Order Preservation in} \( \Delta \), for any \( \Delta \subseteq \Omega \), if and only if, for every \( \omega, \omega' \in \Delta \), \( \omega \sqsubseteq \omega' \) iff \( \omega \sqsubseteq \omega' \).

When \( - \) obeys \textit{Order Preservation} in \( \Delta \) upon contraction by \( \alpha \), we write this in shorthand by \( \text{OP}_{\alpha}(\Delta) \). It is worth noting that \textit{Order Preservation} has been studied in the literature in different contexts. Order Preservation property was proposed in the context of iterated belief revision in [16]. Chopra and colleagues [10] provide order preservation properties based on iterated belief change in the context of presenting variations of
the recovery axiom [2].\(^10\) We study order preservation purely in the context of iterated contraction.

When contracting a belief \(\alpha\), the contraction function \(-\) might preserve the ordering of worlds in two different sets \([\alpha]\) and \([-\alpha]\). The following two lemmas show that if the contraction function is an AGM-rational state contraction function that satisfies both \(OP_\alpha[\alpha]\) and \(OP_\alpha[-\alpha]\), then the result of consecutive contractions using this function follows a predictable pattern under special circumstances. We provide the proofs of these and other claims in Appendix A.

**Lemma 2.1:** Let \(\sqsubseteq\) be a consistent belief state and \(\mathcal{K}\) its associated belief set. An AGM-rational state contraction function \(-\) satisfies \(OP_\alpha[\alpha]\) for any sentence \(\alpha\) iff for every sentence \(\beta\) such that \(\vdash \alpha \lor \beta\), \((\mathcal{K}_\alpha)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \lor \beta}^-\).

**Lemma 2.2:** Let \(\sqsubseteq\) be a consistent belief state and \(\mathcal{K}\) its associated belief set. An AGM-rational state contraction function \(-\) satisfies \(OP_\alpha[-\alpha]\) for any sentence \(\alpha\) iff for every sentence \(\beta\) such that \(\vdash \alpha \rightarrow \beta\), \((\mathcal{K}_\alpha)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-\).

Lemmas 2.1 and 2.2 highlight the similarity between the three contraction functions. To illustrate the differences between the three contraction functions, we draw attention to an example provided in [71] which is a variation of the well known example given in [16].

**Example 2.1:** Let the agent believe that Mr. Craig is rich and Mr. Craig is smart. The result of removing the belief smart followed by the removal of the belief rich in terms of the three state contraction functions is given in Figure 2.2.

In this example it is shown that given an initial belief state represented as a total pre-order on \(\Omega\), shown as the four-layered box on the leftmost column, all the three contraction

\(^{10}\)The recovery axiom: \(\mathcal{K} \subseteq (\mathcal{K}_\alpha)^+\).
functions retain the same set of beliefs after first contraction, represented in Figure 2.2 as the lowest layers in the three boxes in the central column (consisting of worlds 01 and 11). But in the process of contracting the belief, the three contraction functions change the belief state in different ways, thus resulting in different belief states. Upon second contraction, according to moderate contraction function, the agent will end up believing *if smart then rich*. When the agent uses conservative contraction function, the agent will end up believing *either smart or rich*. On the other hand, if the contraction function used is lexicographic contraction, the agent will end up believing only in tautologies. The difference in how the non-minimal [\(\alpha\)] and [\(\neg\alpha\)] worlds are shifted relative to each other by each contraction function gives rise to different results upon iterated contraction. In Section 2.4, we study the effect of different changes to the preorder relation on iterated contraction.

### 2.2.5 Principled factored insertion

In their attempt to characterize lexicographic contraction functions, the authors in [70] studied *Principled Factored Insertion* which is derived from *Qualified Intersection* [84]
Factoring [2].

**Principled Factored Insertion (PFI).** Given $\beta \in K_{\alpha}^-$

1. If $\alpha \rightarrow \beta \in (K_{\alpha}^-)_\beta$, then $(K_{\alpha}^-)_\beta = K_{\alpha}^- \cap K_{\alpha \lor \beta}^-$

2. If $\beta \lor \alpha \in (K_{\alpha}^-)_\beta$, then $(K_{\alpha}^-)_\beta = K_{\alpha}^- \cap K_{\alpha \rightarrow \beta}^-$

3. If neither $\alpha \rightarrow \beta \in (K_{\alpha}^-)_\beta$ nor $\alpha \lor \beta \in (K_{\alpha}^-)_\beta$, then $(K_{\alpha}^-)_\beta = K_{\alpha}^- \cap K_{\alpha \lor \beta}^- \cap K_{\alpha \rightarrow \beta}^-$. 

Any AGM contraction function that satisfies PFI is said to be a *principled iterated contraction operation*. In [70] it was shown that every lexicographic contraction function is a principled iterated contraction operation. However, it was also observed that even moderate contraction function is a principled iterated contraction operation. Here we present sufficiency conditions for a contraction function to satisfy PFI.

**Theorem 2.1:** Every contraction function — satisfying G2, $\mathcal{OP}_{\alpha}[\alpha]$ and $\mathcal{OP}_{\alpha}[-\alpha]$ satisfies the principled factored insertion (PFI).

We have already seen that moderate, natural and lexicographic contraction functions satisfy $\mathcal{OP}_{\alpha}[\alpha]$ and $\mathcal{OP}_{\alpha}[-\alpha]$. They also satisfy G2. Hence all the three contraction functions satisfy PFI, and therefore, they are all principled iterated contraction operations. These contraction functions differ from each other in how the ordering between the worlds of $[\alpha]$ and $[-\alpha]$ are changed relative to each other. To characterize these contraction functions we need to identify properties that capture these changes. We investigate such properties in Section 2.4. But before that, we present a measure called *degree of belief* based on the belief state of the agent. This measure would be helpful in investigating the characteristics of different ways of changing the preorder on $\Omega$. 

47
2.3 Degrees of Belief

All the beliefs of an agent need not be considered to be equally important. Prioritization of beliefs has been studied in detail in many works, [20, 22, 52, 86]. The quality or strength of beliefs is captured in the literature in different ways, as ‘incorrigibility’ of a belief [59, 61], degrees of potential surprise or disbelief [88], epistemic entrenchment relations [29], possibility measures [20] and ordinal ranking functions [92]. These offer a measure of gradation among the beliefs of an agent, be it qualitative or quantitative. Here we present a quantitative measure that we call degree of belief, and denote it by the function \( d(\cdot) \).

Before we define the function \( d \), we recall that a chain of worlds \( \omega_0 \sqsubseteq \omega_1 \sqsubseteq \ldots \sqsubseteq \omega_n \) is said to be complete iff \( \omega_0 \) is a \( \sqsubseteq \)-minimal world of \( \Omega \) and for any \( \omega_{i-1} \sqsubseteq \omega_i \) (\( 1 \leq i \leq n \)) there does not exist any world \( \omega' \) such that \( \omega_{i-1} \sqsubseteq \omega' \sqsubseteq \omega_i \). Degree of belief, \( d \) is a function that takes the sentences in the language to a non-negative integer. To every sentence \( \alpha \), the function \( d \) assigns a value based on the belief state \( \sqsubseteq \).

**Definition 2.2:** Given a consistent belief state \( \sqsubseteq \) (that is, \( \sqsubseteq \neq \bot \)), the degree of belief \( d \) for any sentence \( \alpha \in L \) is defined as

1. \( d(\alpha) = \infty \) when \( \vdash \alpha \),
2. \( d(\alpha) = 0 \) when \( \alpha \vdash \bot \),
3. \( d(\alpha) = \min \{ \| C \| : C \) is a complete chain of worlds in \( \Omega \) ending in some model of \( \neg \alpha \} \), for any \( \alpha \) such that \( \nvdash \alpha \) and \( \alpha \nvDash \bot \).

One way to get a better grasp of this notion of degree of belief is the following. Imagine that the belief state \( \sqsubseteq \) that typically is viewed as a system of spheres in fact represents a graduated cup [68], shown in Figure 2.3.\(^{11}\) The central sphere, that is, \( \min_{\sqsubseteq}(\Omega) \), constitutes the base of this cup and other subsequent \( \sqsubseteq \)-equivalence classes correspond to

\(^{11}\)In [68], the notion of graduated cup was used to motivate the semantics of epistemic entrenchment (footnote 12 in Section 1.4.3). Here we use it to motivate the degree of belief which is similar.
different levels as marked in the cup. The set of worlds $[\neg \alpha]$ then will correspond to some “hole” in this cup and $[\alpha]$ to a graduated cup with a hole in it. Given this picture, the measure $d(\alpha)$ in fact specifies the lowest point of the hole(s) represented by $[\neg \alpha]$, and hence, equivalently, tells how much liquid the damaged graduated cup $[\alpha]$ can hold. Alternatively, we may say $d(\alpha)$ specifies where the belief $\alpha$ leaks, and to that effect specifies the degree of belief in $\alpha$. Definition 2.2 above refers only to the case where the belief state is consistent. An inconsistent belief state is represented by $\sqsubseteq \bot$ which is totally disconnected. Then the set of minimal worlds $\min_{\sqsubseteq \bot}(\Omega)$ is an empty set. This is a special case, and needs to be dealt with as such. Hence we postulate that the degree of belief of every sentence in an inconsistent belief state is $\infty$. Alternatively, when the belief state is denoted by a full preorder relation $\sqsubseteq \top$, we get $\min_{\sqsubseteq \top}(\Omega) = \Omega$. In this case, for every contingent sentence there is a model of its negation which is present in $\min_{\sqsubseteq}(\Omega)$. By our definition, every contingent sentence is, therefore, assigned zero as degree of belief. This is as expected, since when a belief state is represented by a full preorder every contingent sentence is a non-belief and the only beliefs are tautologies.

The following properties of $d$ are easily established. Given two sentences $\alpha$ and $\beta$ in the language:

1. $d(\alpha \land \beta) = \min\{d(\alpha), d(\beta)\}$.

2. $d(\beta) \geq d(\alpha)$, when $\alpha$ and $\beta$ are such that $\alpha \vdash \beta$. 

Figure 2.3: The cup and the hole (denoted by the shaded region) together form the set of all possible worlds. The cup in (a) holds more liquid than the cup in (b); hence the degree of belief in the proposition $\alpha$ is higher than that of $\beta$. 

Figure 2.3: The cup and the hole (denoted by the shaded region) together form the set of all possible worlds. The cup in (a) holds more liquid than the cup in (b); hence the degree of belief in the proposition $\alpha$ is higher than that of $\beta$. 

The following properties of $d$ are easily established. Given two sentences $\alpha$ and $\beta$ in the language:

1. $d(\alpha \land \beta) = \min\{d(\alpha), d(\beta)\}$.

2. $d(\beta) \geq d(\alpha)$, when $\alpha$ and $\beta$ are such that $\alpha \vdash \beta$. 

49
Further, we have defined the degree of belief of $\bot$ to be zero. By observing these properties, we realize that our definition of degree of belief bears a striking resemblance to the Entrenchment Ranking Function [86]. In fact, our definition is a translation of the notion of entrenchment ranking function in to a system of spheres framework. We now extend our definition of degree of belief to some restricted belief states.

### 2.3.1 Conditional degrees of belief

The degree of belief of a sentence is the dual of a measure of doubt or uncertainty on that sentence. The higher the degree of belief of a sentence the lower is its measure of doubt. When a sentence has maximum degree of belief, there is no trace of doubt attached to it. The agent assigns a degree of belief zero to any sentence it maximally doubts. In general, tautologies are the only sentences to have maximum degree of belief. However, it is quite conceivable that an agent unreservedly believes a sentence which is not a tautology. For instance a highly opinionated agent might refuse to see the fallibility of some of its beliefs. Alternatively, a reasonable agent might take a certain physical law such as the Second Law of Thermodynamics\(^{12}\) as a given, and beyond the reach of any doubt. In that case the agent considers the sentence in question to have maximum degree of belief. We propose to capture such phenomena by a restriction of the belief state such that a given contingent sentence has maximum degree of belief.

Let $\sqsubseteq$ be the initial belief state of the agent (where only tautologies have maximum degree of belief). Suppose that the agent decides that a set of contingent sentences are true beyond any doubt. We denote this set of sentences by $\mathcal{R}$. The agent then assigns a maximum degree of belief to the sentences in $\mathcal{R}$. To this end, we restrict the initial belief state $\sqsubseteq$ of the agent to $[\mathcal{R}]$, denoting the resultant state by $\sqsubseteq_{R}$. In other words, for any two possible worlds $\omega_1$ and $\omega_2$, $\omega_1 \sqsubseteq_R \omega_2$ if and only if $\omega_1, \omega_2 \in [\mathcal{R}]$ and $\omega_1 \sqsubseteq \omega_2$.

\(^{12}\)This law says the the overall entropy of a closed system will never decrease.
call this restriction of the belief state by $\mathcal{R}$ as *conditionalization* of the belief state.\textsuperscript{13} We denote the degree of belief in the conditionalized belief state as $d_{\mathcal{R}}$ or $d(\cdot | \mathcal{R})$.

To gain a better understanding of the issue at hand, let us consider a very simple case. Suppose that the set $\mathcal{R}$ is a singleton set, $\mathcal{R} = \{\alpha\}$. There are three cases to consider while conditionalizing by $\alpha$.

Case 1. When $\alpha$ is an existing belief of the agent, every $\sqsubseteq$-minimal world of $\Omega$ is a model of $\alpha$. Upon conditionalization, all these worlds remain to be minimal worlds of $[\mathcal{R}]$ (now $[\mathcal{R}] = [\alpha]$) based on $\sqsubseteq_{\mathcal{R}}$. Therefore the set of beliefs of the agent remains the same after conditionalization. However, the degree of belief assigned to sentences have changed. For instance, degree of belief of $\alpha$ is changed from some non-zero value to $d_{\mathcal{R}}(\alpha) = \infty$.

Case 2. Suppose neither $\alpha$ nor $\neg\alpha$ is believed in the initial belief state. The set of $\sqsubseteq_{\mathcal{R}}$-minimal worlds is given by the set of all $\sqsubseteq$-minimal worlds of $\Omega$ which are also models of $\alpha$. Therefore, from G1 we see that no beliefs are lost but the belief set is expanded to include $\alpha$. The degree of belief of sentence $\alpha$ is changed from zero to $\infty$ upon conditionalization while the degree of belief of $\neg\alpha$ remains zero.

Case 3. Suppose the agent initially believes in $\neg\alpha$. Conditionalisation by $\alpha$ in this case changes the set of beliefs. Some beliefs are lost and some beliefs are gained. The minimal worlds of $[\alpha]$ based on the relation $\sqsubseteq$ become the minimal worlds based on the relation $\sqsubseteq_{\mathcal{R}}$. The degree of belief of $\alpha$ changes from zero to $\infty$ and the degree of belief of $\neg\alpha$ changes from non-zero positive value to zero.

When the set $\mathcal{R}$ is empty, this reduces to the case where only logical tautologies are the sentences with maximum degree of belief. It can be seen that $d(\alpha) = d(\alpha | \top)$, where $d$ is the degree of belief based on $\sqsubseteq$. On the other hand suppose $\mathcal{R}$ contains an inconsistent sentence. Conditionalization by $\mathcal{R}$ in this case results in an inconsistent belief state. We have already defined the degree of belief of any sentence in an inconsistent state to be $\infty$.

\textsuperscript{13}This is very similar to conditional probability. In probabilistic conditionalization, the posterior state ignores the simple events that are inconsistent with the evidence; similarly here worlds outside $[\mathcal{R}]$ are ignored.
Hence the degree of belief of every sentence upon conditionalization by an inconsistent sentence becomes $\infty$. Some of the properties of conditional degrees of belief are as follows:

1. \( d(\beta|\alpha) = \infty \) for every \( \beta \) such that \( \alpha \vdash \beta \). In particular:
   
   (a) \( d(\alpha|\alpha) = \infty \).
   
   (b) \( d(\top|\alpha) = \infty \).
   
   (c) \( d(\beta|\bot) = \infty \) for every \( \beta \).

2. \( d(\beta|\alpha) = 0 \) for every \( \beta \) such that \( \alpha \vdash \neg \beta \). As special cases, we have:
   
   (a) \( d(\bot|\alpha) = 0 \).
   
   (b) \( d(\neg \alpha|\alpha) = 0 \).

Although the degree of belief is similar to epistemic entrenchment, it is richer than a pure relational measure. In a certain sense, it allows us to “compare” the effects of different epistemic changes on one or more beliefs. For instance, intuitively, I would more firmly believe that kangaroos are man eaters if I observe a killer kangaroo than if I observe a grass munching kangaroo. While this intuition cannot be captured via relational measures such as epistemic entrenchment, it can be captured using the degree of belief. We will make use of such conditional degrees of belief in characterising lexicographic contraction in Section 2.5. But before that we need to discuss various cases that arise in sequential contraction of \( \alpha \) followed by \( \beta \); and that is the topic we take up in the next section.

### 2.4 Plausible Properties of Iterable Contraction

As we saw in Example 2.1, the difference between the three contraction functions is evident in the resultant sets of beliefs only after repeated contractions. Therefore we aim to completely characterize these state contraction functions with the help of the properties based on iterated contractions. Towards this goal, in this section we study various cases
that arise when contracting two beliefs \( \alpha \) and \( \beta \) one after the other. From Equation G2 it is evident that when the belief being withdrawn is not present in the belief set, the result of contraction by any of these three contraction functions does not change the resulting set of beliefs. Hence in the following discussion, we will assume that \( \alpha, \beta \in \mathcal{K} \) and also \( \beta \in \mathcal{K}_\alpha^- \). Sentences \( \alpha \lor \beta \) and \( \alpha \rightarrow \beta \) are two important factors to be considered when removing \( \beta \) following removal of \( \alpha \). Let us suppose \( \alpha \lor \beta \) survives the individual removal of \( \alpha \) and \( \beta \), that is, \( \alpha \lor \beta \in \mathcal{K}_\alpha^- \) and \( \alpha \lor \beta \in \mathcal{K}_\beta^- \). In such a case one might want \( \alpha \lor \beta \in (\mathcal{K}_\alpha^-)_\beta^- \). Given the assumption that \( \beta \in \mathcal{K}_\alpha^- \), we get:\(^{14}\)

\[
\alpha \lor \beta \in \mathcal{K}_\beta^- \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-.
\] (2.1)

As a special case we have \( \alpha \) and \( \beta \) such that \( \vdash \alpha \lor \beta \). Thus,

\[
\vdash \alpha \lor \beta \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-.
\] (2.2)

When the belief \( \alpha \rightarrow \beta \) is preferred to \( \alpha \lor \beta \), the agent might retain \( \alpha \rightarrow \beta \) in \( \mathcal{K}_\beta^- \), at the cost of \( \alpha \lor \beta \). Since both \( \alpha \rightarrow \beta \in \mathcal{K}_\alpha^- \) and \( \alpha \lor \beta \in \mathcal{K}_\alpha^- \) (courtesy the assumption that \( \beta \) in \( \mathcal{K}_\alpha^- \)), the agent might make the same choice when contracting \( \beta \) from \( \mathcal{K}_\alpha^- \); that is, \( \alpha \rightarrow \beta \in (\mathcal{K}_\alpha^-)_\beta^- \). Therefore \( \alpha \rightarrow \beta \in \mathcal{K}_\beta^- \) suggests that \( \alpha \rightarrow \beta \in (\mathcal{K}_\alpha^-)_\beta^- \) and hence from PFI

\[
\alpha \rightarrow \beta \in \mathcal{K}_\beta^- \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \lor \beta}^-.
\] (2.3)

As a special case of Equation 2.3, we get:

\[
\vdash \alpha \rightarrow \beta \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \lor \beta}^-.
\] (2.4)

\(^{14}\)Note that the symbol \( \Rightarrow \) in the properties listed henceforth is not a logical connective. For readability, we use this symbol rather than natural language “if . . . , then . . .”.

53
We recall that Lemmas 2.1 and 2.2 concern Equations 2.2 and 2.4 respectively. Furthermore, as a dual of Equation 2.2 we get:

\[ \forall \alpha \vee \beta \Rightarrow (K_{\alpha})_{\beta} = K_{\alpha} \cap K_{\alpha \vee \beta}. \] (2.5)

**Lemma 2.3:** An AGM-rational contraction function satisfies Equation 2.5 iff for every sentence \( \alpha \), both: (a) \( -satisfies OP_{\alpha}[\neg \alpha] \), and (b) \( \omega \sqsubseteq_{\alpha} \omega' \) for every \( \omega, \omega' \in \Omega \) such that \( \omega \in [\neg \alpha] \) and \( \omega' \in [\alpha] \setminus \min_{\subseteq}[\alpha] \).

Finally, Equations 2.1 and 2.3 jointly suggest:

\[ \alpha \vee \beta \notin K_{\beta}, \alpha \rightarrow \beta \notin K_{\beta} \Rightarrow (K_{\alpha})_{\beta} = K_{\alpha} \cap K_{\alpha \vee \beta} \cap K_{\alpha \rightarrow \beta}. \] (2.6)

**Lemma 2.4:** For any arbitrary beliefs \( \alpha \) and \( \beta \) such that \( \beta \in K_{\alpha} \), an AGM-rational contraction function satisfies Equations 2.1, 2.3 and 2.6 if and only if \( \forall \omega, \omega' \notin (\min_{\subseteq}(\Omega) \cup \min_{\subseteq}[\neg \alpha]) \), \( \omega \sqsubseteq_{\alpha} \omega' \) iff \( \omega \sqsubseteq \omega' \).

The Equations 2.1, 2.3 and 2.6 have been presented as *Naive factored insertion* in [70]. It is worth noting that there is a certain amount of conflict among the Equations 2.1-2.6. For instance, one can imagine cases where the preconditions of Equations 2.1 and 2.5 are jointly satisfied, but their consequences are not. In fact, it may be argued that these equations naturally fall into two exclusive sets: \{2.2 and 2.5\} is one and \{2.1, 2.3 and 2.6\} is the other, since their respective preconditions partition the hypothesis space in different manner.

### 2.4.1 Properties based on Degrees of belief

We now present a translation of Equations 2.1 to 2.6 in terms of degrees of belief. As we have already seen, \( \alpha \vee \beta \) and \( \alpha \rightarrow \beta \) are two important factors to be considered when
removing $\beta$ following the removal of $\alpha$. The choice between contracting by $\alpha \lor \beta$ and $\alpha \rightarrow \beta$ can be resolved with the help of degrees of belief. When the degree of belief in $\alpha \lor \beta$ is greater than the degree of belief in $\alpha \rightarrow \beta$, then the agent contracts $\alpha \rightarrow \beta$, and similarly it contracts $\alpha \lor \beta$ when the degree of belief in $\alpha \lor \beta$ is less than that of $\alpha \rightarrow \beta$. When both have equal degrees of belief then both are contracted. This can be formalized in terms of the following:

$$d(\alpha \lor \beta) > d(\alpha \rightarrow \beta) \Rightarrow (\mathcal{C}_\alpha^-)_\beta = \mathcal{C}_\alpha^- \cap \mathcal{C}_{\alpha \rightarrow \beta}^-.$$  \hspace{1cm} (2.7)

$$d(\alpha \lor \beta) < d(\alpha \rightarrow \beta) \Rightarrow (\mathcal{C}_\alpha^-)_\beta = \mathcal{C}_\alpha^- \cap \mathcal{C}_{\alpha \lor \beta}^-.$$  \hspace{1cm} (2.8)

$$d(\alpha \lor \beta) = d(\alpha \rightarrow \beta) \Rightarrow (\mathcal{C}_\alpha^-)_\beta = \mathcal{C}_\alpha^- \cap \mathcal{C}_{\alpha \rightarrow \beta}^- \cap \mathcal{C}_{\alpha \lor \beta}^-.$$  \hspace{1cm} (2.9)

With the help of following observation we note that Equations 2.7, 2.8 and 2.9 are equivalent to Equations 2.1, 2.3 and 2.6 respectively when the contraction function involved is AGM-rational.

**Observation 2.1:** Let $\rightarrow$ be an AGM-rational contraction function, and the degree of belief function $d$ is appropriately related with the presumed belief state $\sqsubseteq$. Then:

(a) Equation 2.1 is satisfied iff Equation 2.7 is,

(b) Equation 2.3 is satisfied iff Equation 2.8 is, and

(c) Equation 2.6 is satisfied iff Equation 2.9 is.

Equation 2.5, which states that when $\alpha \lor \beta$ is not a tautology the contraction by $\beta$ after contraction by $\alpha$ is given by the meet of contraction by $\alpha$ and $\alpha \lor \beta$, can be captured in terms of degrees of belief by

$$d(\alpha \lor \beta) < \infty \Rightarrow (\mathcal{C}_\alpha^-)_\beta = \mathcal{C}_\alpha^- \cap \mathcal{C}_{\alpha \lor \beta}^-.$$  \hspace{1cm} (2.10)
The Equation 2.2, on the other hand, can be translated as follows:

\[ d(\alpha \lor \beta) = \infty \Rightarrow (K^\bot_{\alpha})_{\beta} = K^\bot_{\alpha} \cap K^\bot_{\alpha \rightarrow \beta}. \quad (2.11) \]

By using conditional degrees of belief, we can list some more properties for iterated contraction. The degree of belief of \( \alpha \lor \beta \) in the belief state conditionalised with respect to \( \neg \alpha \) can be taken as indicating the degree to which the belief \( \alpha \lor \beta \) is independent of \( \alpha \).

Similarly the conditional degree of belief \( d(\alpha \rightarrow \beta | \alpha) \) can be considered as indicating the degree of independence of belief \( \alpha \rightarrow \beta \) from \( \neg \alpha \). Suppose we have \( d(\alpha \rightarrow \beta | \alpha) > d(\alpha \lor \beta | \neg \alpha) \), we can interpret this as \( \alpha \rightarrow \beta \) is more “self-sufficient” than \( \alpha \lor \beta \) and hence there is more reason to retain the belief in \( \alpha \rightarrow \beta \) after iterated contraction of \( \alpha \) followed by \( \beta \).

Going back to PFI in case the agent decides to retain \( \alpha \rightarrow \beta \) in \( (K^\bot_{\alpha})_{\beta} \), we say that the result of iterated contraction is given by the meet of contraction by \( \alpha \) and \( \alpha \lor \beta \). Similarly when \( d(\alpha \lor \beta | \neg \alpha) > d(\alpha \rightarrow \beta | \alpha) \), the result of iterated contraction can be derived from PFI to be the meet of contraction by \( \alpha \) and \( \alpha \rightarrow \beta \). If \( d(\alpha \lor \beta | \neg \alpha) = d(\alpha \rightarrow \beta | \alpha) \), then it is indicated that both \( \alpha \lor \beta \) and \( \alpha \rightarrow \beta \) are equally independent of \( \alpha \) and \( \neg \alpha \), respectively.

With equal preference, the agent could decide not to retain either of them and hence the result of the iterated contraction could be the combined meet of contraction by \( \alpha \), \( \alpha \lor \beta \) and \( \alpha \rightarrow \beta \). We list these in the form of following properties.

\[ d(\neg \alpha \lor \beta | \alpha) > d(\alpha \lor \beta | \neg \alpha) \Rightarrow (K^\bot_{\alpha \lor \beta})_{\beta} = K^\bot_{\alpha} \cap K^\bot_{\alpha \lor \beta} \quad (2.12) \]

\[ d(\alpha \lor \beta | \neg \alpha) > d(\neg \alpha \lor \beta | \alpha) \Rightarrow (K^\bot_{\alpha \lor \beta})_{\beta} = K^\bot_{\alpha} \cap K^\bot_{\alpha \rightarrow \beta} \quad (2.13) \]

\[ d(\alpha \lor \beta | \neg \alpha) = d(\neg \alpha \lor \beta | \alpha) \Rightarrow (K^\bot_{\alpha \lor \beta})_{\beta} = K^\bot_{\alpha} \cap K^\bot_{\alpha \rightarrow \beta} \cap K^\bot_{\alpha \lor \beta}. \quad (2.14) \]

We give two lemmas that connect the properties based on degrees of belief and properties 2.1-2.6 described earlier in Section 2.4.
Lemma 2.5: Any AGM-rational contraction function — that satisfies Equation 2.13 also satisfies Equation 2.2.

Lemma 2.6: Any AGM-rational contraction function — that satisfies Equation 2.12 also satisfies Equation 2.4.

2.5 Representation Results

In this section we will provide the representation results which axiomatically characterize moderate, natural and lexicographic contraction functions based on these properties.

Theorem 2.2: An AGM-rational contraction function is a moderate contraction function iff it satisfies Equation 2.2 and 2.5.

The above result gives a very simple characterisation of the moderate contraction function in that, it identifies the AGM-rational contraction functions that also satisfy the equations:

\[ \vdash \alpha \lor \beta \Rightarrow (K^-_{\alpha})^- = K^-_{\alpha} \cap K^-_{\alpha \lor \beta} \]  
\[ \forall \alpha \lor \beta \Rightarrow (K^-_{\alpha})^- = K^-_{\alpha} \cap K^-_{\alpha \lor \beta} \]  

(2.2) \hspace{2cm} (2.5)

Theorem 2.3 identifies the necessary and sufficiency conditions for a contraction function to be exactly the moderate contraction functions. In terms of degrees of belief, any AGM-rational contraction function is a moderate contraction function iff it satisfies Equations 2.10 and 2.11.

Theorem 2.3: An AGM-rational contraction function is a natural contraction function iff it satisfies Equations 2.1, 2.3 and 2.6.

Theorem 2.3 identifies the necessary and sufficiency conditions for a contraction func-
tion to be qualified as a natural contraction function. Keeping in mind how AGM-rational contraction functions are constructed, we claim that a belief (set) contraction function is generated from a natural belief state contraction function iff it satisfies the AGM contraction postulates along with

\[ \alpha \vee \beta \in \mathcal{K}_\beta^- \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^- \quad (2.1) \]

\[ \alpha \rightarrow \beta \in \mathcal{K}_\beta^- \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \vee \beta}^- \quad (2.3) \]

\[ \alpha \vee \beta, \alpha \rightarrow \beta \notin \mathcal{K}_\beta^- \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \vee \beta}^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^- \quad (2.6) \]

Since Equations 2.1, 2.3 and 2.6 form the Naive factored insertion [70], we have that an AGM-rational contraction function is a natural contraction function iff it satisfies naive factored insertion. The natural contraction function can also be characterized in terms of degrees of belief based on a given belief state. Every AGM-rational contraction function is a natural (or conservative) contraction function if and only if it satisfies the Equations 2.7, 2.8 and 2.9.

**Theorem 2.4:** An AGM-rational contraction function is a lexicographic contraction function iff it satisfies Equations 2.12, 2.13 and 2.14

The above theorem gives the necessary and sufficient conditions for a contraction function to be a lexicographic contraction function. Thus we have that an AGM-rational contraction function is lexicographic iff it satisfies the equations:

\[ d(\lnot \alpha \vee \beta | \alpha) > d(\alpha \vee \beta | \lnot \alpha) \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \vee \beta}^- \quad (2.12) \]

\[ d(\alpha \vee \beta | \lnot \alpha) > d(\lnot \alpha \vee \beta | \alpha) \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^- \quad (2.13) \]

\[ d(\alpha \vee \beta | \lnot \alpha) = d(\lnot \alpha \vee \beta | \alpha) \Rightarrow (\mathcal{K}_\alpha^-)_\beta^- = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^- \cap \mathcal{K}_{\alpha \vee \beta}^- \quad (2.14) \]
2.6 Concluding Remarks

In this paper we examined three different iterable contraction functions, namely, the moderate (or priority) contraction, the natural (or conservative) contraction and the lexicographic contraction. The proposals for these contraction functions were originally couched in semantic terms, and their properties have not been studied before. We showed that all these three functions satisfy the *Principled factored insertion* which was studied in [70] in relation to lexicographic contraction. We presented and examined a list of plausible properties (similar to rationality postulates) that one may expect an iterable contraction function to satisfy. Using these properties, we provided representation results for the natural and moderate contraction functions.

In order that we can characterise lexicographic contraction, we introduced a quantitative measure that we call the *degrees of belief*, and its derivative notion called the *conditional degree of belief*. This notion allows simple representation of plausible properties of iterated contraction. Using different sets of such properties we characterised lexicographic, moderate as well as natural contraction functions. Nonetheless, we find it somewhat unsatisfying that a standard characterisation of lexicographic contraction still eludes us. It will be nice if such a characterisation can be given, or it can be established that such standard characterisation of lexicographic contraction is not possible. We leave this task to a future occasion.
Chapter 3

Probabilistic Belief Erasure

*Education is the path from cocky ignorance to miserable uncertainty.*

Mark Twain.

3.1 Introduction

We often need to deal with uncertainty, if for no other reason, because the information that we receive is uncertain. Hence it is important to have a satisfactory account of dealing with uncertain information. The typical way of doing it is founded upon the Bayesian tradition – possibly enhanced by Jeffery conditionalisation [50]. The belief state of an agent is represented by a probability function, say \( P \). Bayesian updating uses a normalisation process according to which the probability mass of all simple events (or possible worlds) in a select subset of the sample space are proportionally blown up so that the sum the probability mass in this subset becomes 1. This normalisation proves to be a stumbling block on the path of developing an account which can deal with new information, uncertain or otherwise, if the evidence-sentence has zero initial probability. On the other hand, if \( P(\alpha) = 1 \) for a given sentence \( \alpha \), then \( P(\neg \alpha) = 0 \) leading to the complementary problem, namely, once a fact is learnt it cannot be unlearnt, that is \( \alpha \) always has probability 1 under any form of conditionalization. A sentence is said to be a full belief if it has
probability 1 [60]. Consequently, removal of full beliefs is not possible, by any account based on Bayesian conditionalization.

David Lewis [63], in order to offer a cogent account of Stalnaker conditionals [93], provides an alternative to this Bayesian rationality (read uniformity) requirement. He proposes imaging as an alternative rational method of redistributing probability mass. This process assumes that there is a mechanism available to determine, given certain world $\omega$, the closest or most similar world $\omega'$ to $\omega$ satisfying certain constraints. The probability associated with the world $\omega$ is moved to the closest world $\omega'$ thus determined. An elaborate discussion of the distinctions between conditionalization and imaging can be found in [13, 29, 61]. Imaging has proved useful in solving many well-known problems such as The Monty Hall Problem [13] and Sleeping Beauty Scenario [11, 23]. Imaging has also been discussed in the context of information retrieval [12].

Since Lewis’ account dispenses with the proportionality requirement, in principle, it should be possible to account for probabilistic belief revision/removal. Nonetheless, problem of probabilistic belief removal have been pretty much ignored in the literature except for the seminal works by Peter Gärdenfors [29]. However, the account that Gärdenfors develops assumes a similarity measure not among worlds, but among probability functions. On the other hand, the well-known account of belief update – inclusion in a dynamic setting – developed by Katsuno and Mendelzon [53], and studied further in [21], deploys imaging in a non-probabilistic framework. Yet the question remains whether an acceptable accounts of probabilistic belief removal can be developed based on Lewis’s original idea of imaging. We address this question in this paper.

Three variations of imaging, namely partial imaging, selective imaging and selective partial imaging were introduced in [81]. Here we re-introduce partial imaging (Section 3.4), selective imaging (Section 3.5) and selective partial imaging (Section 3.6) with more general definitions and study them in greater detail. We argue that these variations are indeed generalizations of imaging and that they model specific changes to the epis-
emic state of an agent. We show that partial imaging, under specific conditions, models removal of a belief (Section 3.4.1). It is worth noting that while the account of probabilistic belief removal developed by Gärdenfors [28, 29] is an account of belief contraction, the account we develop is best viewed as the probabilistic analogue of belief erasure introduced in [53] intended to model belief removal in a dynamic environment. Furthermore, in Section 3.4.2 we show that partial imaging is the generalization of imaging just as Jeffrey’s probability kinematics is a generalization of the orthodox Bayesian conditionalization. We introduce the notions of conditional update and conditional erasure that can handle conditional instructions of the sort: when there is a gas-leak, open the window. These conditional instructions are intended to convey information about possible changes to the state of the world when certain triggers are activated. Conditional update and erasure change the belief state depending on the agent’s stance towards the antecedent of the instruction in question. When the agent considers the antecedent to be false, that is, attaches zero probability to the antecedent, then the conditional instruction in question does not alter the belief state of the agent. The belief state is expected to change only when the agent attaches non-zero probability for the antecedent to be true. We demonstrate that selective imaging models conditional update (Section 3.5.1) and selective partial imaging models conditional erasure (Section 3.6.1).

Before going to any details, we first present the notations to be followed throughout this paper.

3.2 Preliminaries

In this paper we let $L$ represent the propositional language generated from a finite set of propositional atoms with the standard connectives $\land, \lor, \neg$ and $\rightarrow$. The sentences of this language are denoted by lower case Greek letters (with or without any super- or sub-scripts) such as $\alpha, \beta, \gamma$ and $\xi$. The set of all possible worlds is denoted by $\Omega$ whose
individual elements are denoted by $\omega, \mu, \text{ and } \nu$ (with possible decorations). We define $\omega(\alpha) = 1$ when $\omega \models \alpha$ and $\omega(\alpha) = 0$ otherwise. The set of models of a sentence $\alpha$ is the set of interpretations/possible worlds in which $\alpha$ is true and is represented by $[\alpha]$. In our examples we denote a possible world by the set of all literals which are true in that world, for example if $\alpha$ and $\beta$ are the only atoms of $\mathcal{L}$, $\{\alpha, \beta\}$ represents the possible world in which both $\alpha$ and $\beta$ are true. We will use $\{\alpha, \overline{\beta}\}$ as a shortened representation of $\{\alpha, \neg \beta\}$. In the pictorial representations such as Figure 3.1 we drop the braces for readability.

By $\alpha \rightarrow \beta$ we denote a conditional instruction which states that when $\alpha$ holds true, the state of the world needs to be changed such that $\beta$ holds true. The set of all conditional instructions is represented by $\mathcal{C}$.

A classical consequence operator $\mathcal{C}_n$ governs the background logic. It is defined by $\mathcal{C}_n(A) = \{\alpha | A \vdash \alpha\}$, where $\vdash$ is the classical consequence relation for propositional logic and $A$ does not contain any conditional instruction. A probability distribution over $\Omega$ is denoted by $P$ (with and without decorations), such that, $\sum_{\Omega} P(\omega) = 1$. The set of all probability functions is denoted by $\mathcal{P}$. We denote the set of worlds which have non-zero probability under the function $P$ by $\|P\|$. The probability function on $\mathcal{L}$ is defined as follows: for any sentence $\alpha$,

$$P(\alpha) = \sum_{\Omega} \{P(\omega) : \omega(\alpha) = 1\}.$$ 

Since some worlds in $\Omega$ could have a probability of zero, it is not necessary that only tautologies have probability 1. A sentence $\alpha$ is said to be consistent with the probability function $P$ iff $P(\alpha) > 0$, and is inconsistent with $P$ otherwise. A belief set is a body of beliefs of an epistemic agent closed under $\vdash$, denoted by $\mathcal{K}$. The belief set $\mathcal{K}$, corresponding to the given probability function $P$, is given by the set of all sentences in the language which have probability 1, that is, $\mathcal{K}_P = \{\alpha \in \mathcal{L} : P(\alpha) = 1\}$. When the context is clear, we will avoid using subscripts for the belief set.
3.3 Imaging

Consider an agent who employs a domestic robot in its room. We represent the window in the room being open by $\alpha$ and the door to the room being open by $\beta$. The window and the door being closed are represented by $\neg\alpha$ and $\neg\beta$ respectively. The window and the door are the only factors that are considered in this example. The set of all possible worlds is given by

$$\Omega = \{\{\alpha, \beta\}, \{\alpha, \neg\beta\}, \{\neg\alpha, \beta\}, \{\neg\alpha, \neg\beta\}\}.$$ 

**Example 3.1:** Let us suppose that the agent believes it did not close the door when going to work and the door is still open, that is, $\beta$. The agent instructs the robot to close the door. The robot, following the instruction, closes the door and leaves the state of the window unaltered.

![Figure 3.1: A representation of the change to the state of the world as a result of the robot’s action of closing the door.](image)

This action of the robot changes the state of the real world. When the window and the door are both open the robot closes the door leaving the window open. On the other hand, if the window is closed and the door is open, then the robot closes the door, leaving the window closed. In case the door is already closed, the robot does not alter the state of the world. Figure 3.1 shows how each possible world will be changed by the robot’s action. Let the initial belief state be represented by the probability distribution $P_0$ given as follows:
Since the agent believes the robot will carry out the instruction, the probability is transferred from the \( \beta \)-worlds to the corresponding \( \neg \beta \)-worlds as shown in Figure 3.1. The probability distribution \( P_0 \) changes in such a way that the probability associated with each world is transferred to the corresponding new world as shown in Figure 3.1. The new belief state should be given by the probability distribution \( P_1 \) where

\[
\begin{align*}
P_0(\{\alpha, \beta\}) &= 0.75 \\
P_0(\{\bar{\alpha}, \beta\}) &= 0.25 \\
P_0(\{\alpha, \bar{\beta}\}) &= 0 \\
P_0(\{\bar{\alpha}, \bar{\beta}\}) &= 0.
\end{align*}
\]

This process will work equally well if the \( \neg \beta \)-worlds had non-zero probability. Suppose any \( \neg \beta \)-world had positive probability under \( P_0 \) then that probability mass would remain with it under \( P_1 \) along with the probability transferred to it from some \( \beta \)-world according to Figure 3.1. The action of ‘closing the door’ performed by the robot changes the state of the world. The agent includes a new belief \( \neg \beta \) since it believes the robot to have carried out its instructions. Bringing the belief state of the agent up to date, when the world described by it changes, is termed belief update [53]. Update, as observed in [53], is imaging in a non-probabilistic setting.

Imaging changes the probability distribution in order to assign maximum probability (that is, probability of 1) to a given proposition that has been learnt. For this purpose use is made of a comparative similarity system where for every possible world \( \omega \) there exists a total preorder relation \( \sqsubseteq_\omega \). In [62, 64] Lewis presents a detailed discussion on the comparative similarity relation among the worlds. He uses the relation

\[
\mu \sqsubseteq_\omega \nu
\]
to mean that the world $\mu$ is at least as similar to the world $\omega$ as the world $\nu$. The minimal worlds based on the relation $\subseteq_\omega$ which model a sentence $\xi$ are the most similar $\xi$-worlds to $\omega$. Lewis assumes the relations $\subseteq_\omega$ to be a linear order whereby there exists a unique most similar $\xi$-world to $\omega$. We however relax this assumption and take $\subseteq_\omega$ to be a total preorder relation, therefore the minimal $\xi$-worlds in the relation $\subseteq_\omega$ is a set of worlds. We denote the set of $\xi$-worlds that are most similar to $\omega$ by $\text{min}_{\subseteq_\omega}[\xi]$ or simply $\text{min}_\omega[\xi]$. We assume a ‘choice function’ $\#$ to be a function which, for every pair of a world $\omega$ and a sentence $\xi$ returns a world $\omega^\#_\xi$ where $\omega^\#_\xi \in \text{min}_\omega[\xi]$. We refer to $\omega^\#_\xi$ as the $\#$-selected minimal or simply $\#$-minimal $\xi$-world to $\omega$. It is evident that Lewis’s system of comparative similarity relations can be derived with the help of the choice function $\#$ from the more general system of total preorders centred on every possible world. In case $\omega$ itself is a model of $\xi$, we have $\text{min}_\omega[\xi] = \{\omega\}$ since $\subseteq_\omega$ is a total preorder relation centred on $\omega$, and hence $\omega^\#_\xi = \omega$.

When imaging a probability function $P$ by a sentence $\xi$, the probability mass associated with each model of $\neg\xi$ is shifted to some model of $\xi$. This model of $\xi$ is identified using the preorder relations associated with the models of $\neg\xi$. The probability associated with $\omega \in [\neg\xi]$ is shifted to $\omega^\#_\xi$, the $\#$-minimal $\xi$-world to $\omega$. The new probability function is denoted as $P^\#_\xi$ and is defined as follows:

**Definition 3.1:** Imaging a probability distribution $P$ with respect to a sentence $\xi$ results in a probability distribution $P^\#_\xi$ such that

$$
\text{for every } \nu \in \Omega, \quad P^\#_\xi(\nu) = \sum_{\omega} P(\omega) \cdot \omega^\#_\xi(\nu)
$$

where $\omega^\#_\xi(\nu) = 1$ iff $\omega^\#_\xi = \nu$ and 0 otherwise.
Based on the above definition, the new probability of any sentence $\gamma$ is given by

$$P_{\xi}^\#(\gamma) = \sum_{\Omega} \{P(\omega): \omega_{\xi}^\#(\gamma) = 1\} \quad (3.1)$$

According to Figure 3.1, the arrows point to the most similar $\neg\beta$-world with respect to each world chosen by $\#$, that is, $\omega_{\xi}^\#$ for each $\omega$. The arrow from the world $\{\alpha, \beta\}$ to $\{\alpha, \neg\beta\}$ shows that the $\#$-minimal $\neg\beta$-world of $\{\alpha, \beta\}$ is $\{\alpha, \neg\beta\}$. The arrow from $\{\neg\alpha, \beta\}$ to $\{\neg\alpha, \neg\beta\}$ shows that the $\#$-minimal $\neg\beta$-world of $\{\neg\alpha, \beta\}$ is $\{\neg\alpha, \neg\beta\}$. The $\#$-minimal $\neg\beta$-world to any $\neg\beta$-world is itself by definition of the ‘choice function’. When imaging with respect to $\neg\beta$ the probability associated with $\{\alpha, \beta\}$ is shifted to $\{\alpha, \neg\beta\}$. Therefore imaging $P_0$ with respect to $\neg\beta$ gives the required probability distribution $P_1$. In the following section we propose a generalization of imaging, namely partial imaging, and show how it models belief removal.

### 3.4 Partial Imaging

In the previous section we saw that imaging is capable of changing a given probability function in order to accord maximum probability (that is, 1) to any evidence proposition $\xi$. Suppose that the observation made by the agent does not demand that $\xi$ have final probability of 1, but just requires that the probability of $\xi$ be enhanced. In order to increase the probability assigned to the sentence $\xi$, the probability assigned to its negation has to decrease. Let $P$ be the given initial probability function and $\xi$ be the sentence whose probability needs to be enhanced as per the observation made. To facilitate enhancing the probability of $\xi$, there needs to be a shift of some probability mass from every possible world in $\Omega$ to some specific models of $\xi$. However, in contrast to imaging, the probability mass that is shifted need not be the entire probability associated with the possible worlds.

Let $a$ be a real number between 0 and 1. Let there be a shift of $a$-share of the probability associated with each possible world to some specific models of $\xi$. In other words,
each possible world $\omega$ loses $a \cdot P(\omega)$ retaining only $(1 - a) \cdot P(\omega)$. The probability mass lost by a possible world $\omega$ is moved to its corresponding $\#$-minimal $\xi$-world, that is, $\omega^\#_\xi$. When $\omega$ is a model of $\xi$ it follows that $\omega = \omega^\#_\xi$. The probability mass shifted from any $\omega$ which is a model of $\xi$ is returned to $\omega$. As this process involves a partial movement of the probability mass in comparison with imaging, we call this partial imaging. We denote the result of partial imaging of $P$ with respect to a sentence $\xi$ and the parameter $a$ by $P^{p}_{\xi,a}$.

**Definition 3.2:** Partial imaging $P^{p}_{\xi,a}$ of $P$ with respect to a sentence $\xi$ in $\mathcal{L}$ and a real number $a \in [0, 1]$ is a probability distribution defined over $\Omega$ as follows:

$$
\text{for every } v \in \Omega, P^{p}_{\xi,a}(v) = a \cdot \sum_{\omega} P(\omega) \cdot \omega^\#_\xi(v) + (1 - a) \cdot P(v).
$$

The resultant probability of any $\omega$ which is a model of $\neg \xi$ is $(1 - a) \cdot P(\omega)$ compared to zero which would have been the result upon imaging; and the resultant probability of $\omega^\#_\xi$ for any such $\omega$ is $a \cdot P(\omega)$ in comparison to $P(\omega)$ which would be the case under imaging.\footnote{Here we have assumed a very simplistic case, where $\omega^\#_\xi \neq \omega'^\#_\xi$, for any $\omega'$, to highlight how partial imaging differs from imaging.} $P^{p}_{\xi,a}$ reduces to the following:

$$
P^{p}_{\xi,a} = P^{\#}_{\xi,a} aP
$$

(3.2)

where $P^{\#}_{\xi}$ is the image of $P$ with respect to $\xi$.\footnote{\textit{P} = P_{1}aP_{2} is read as $P = a \cdot P_{1} + (1 - a) \cdot P_{2}$.} For the sake of readability, we will use the notation $P^{p}_{\mathcal{I}}$ instead of $P^{p}_{\xi,a}$, with the understanding that $\mathcal{I} = \langle \xi, a \rangle$ denotes the pair of sentence $\xi$ and the real number $a$. The new probability of a sentence $\alpha$ is given by

$$
P^{p}_{\mathcal{I}}(\alpha) = a \cdot P^{\#}_{\xi}(\alpha) + (1 - a) \cdot P(\alpha).
$$

(3.3)
The resultant probability of $\neg \xi$ is

$$P^p_I(\neg \xi) = (1 - a) \cdot P(\neg \xi) \tag{3.3a}$$

and the resultant probability of the sentence $\xi$ is

$$P^p_I(\xi) = a + (1 - a) \cdot P(\xi) \tag{3.3b}$$

This clearly shows that the probability associated with the sentence $\xi$ may be enhanced (it depends on the value of $a$). Two interesting special cases that arise are with respect to the extreme values $a$ can assume:

1. Case $a = 0$. In this case, the probability mass that is shifted from every possible world to their corresponding $\#$-minimal $\xi$-world is zero. There is no change to the initial probability function. It follows from Equation 3.2 that, $P^p_I = P$, where $\mathcal{I} = \langle \xi, 0 \rangle$.

2. Case $a = 1$. The entire probability mass associated with every possible world is shifted to the respective $\#$-minimal model of $\xi$. Therefore the change to the initial probability function is same as the result of imaging with respect to $\xi$. With the help of Equation 3.2 we can observe that $P^p_I = P^\#_\xi$, where $\mathcal{I} = \langle \xi, 1 \rangle$. Thus, imaging is a special case of partial imaging.

More special cases arise when considering the initial probability of sentence $\xi$:

3. Case $P(\xi) = 1$. We know that when partial imaging with respect to $\xi$, no model of $\xi$ loses the probability mass associated with it. Since every model of $\neg \xi$ has probability zero and for every model of $\xi$, $\omega^\#_\xi = \omega$, the probability function does not change as a result of partial imaging. Therefore we have $P^p_I = P$.

4. Case $P(\xi) = 0$. This is a very special case. When $P(\xi) = 0$, we have $P(\neg \xi) = 1$. Any sentence with probability 1 is a belief, therefore $\neg \xi$ is a belief of the agent. When $a > 0$, the probability of $\neg \xi$ is to be reduced from 1 to some value strictly less than 1 as a result of partial imaging. In other words, partial imaging with respect to $\xi$ and
$a > 0$ results in removal of belief in $\neg \xi$. The resultant probability of $\xi$ becomes $a$ and the resultant probability of $\neg \xi$ becomes $(1 - a)$ from Equations 3.3a, 3.3b. We study this particular application of partial imaging in Section 3.4.1.

In the above discussion of partial imaging we have assumed that $\omega^\#_\xi$ for any possible world $\omega$ and sentence $\xi$ is a single possible world. This definition of ‘choice function’ is consistent with Lewis’s assumption. Gärdenfors argues that such an assumption is too restrictive on imaging [29]. According to Gärdenfors, when imaging with respect to $\xi$, the probability associated with the world $\omega$ should be evenly split among all the worlds in $\min_\omega[\xi]$. Here Gärdenfors assumes that $\omega^\#_\xi$ returns the set $\min_\omega[\xi]$ instead of an element of $\min_\omega[\xi]$. This variation of imaging is termed general imaging [29]. The result of general imaging is also denoted by $P^\#_\xi$, where $\#$ represents the choice function that returns the whole set, $\min_\omega[\xi]$. This generalization can be carried over to the notion of partial imaging as well. Regardless of that, we continue with the assumption that the choice function $\#$ returns a single $\xi$-world from $\min_\omega[\xi]$ represented, henceforth, by $\omega^\#_\xi$.

### 3.4.1 Partial Imaging as a Belief Erasure function

Consider the following scenario:

**Example 3.2:** Suppose the robot has developed a fault which restricts its reliability. The robot no longer performs all the instructions given to it, ignoring one-fifth of its instruction set and the agent is aware of this. When the agent instructs the robot to open the door the agent does not completely expect the robot to perform the task.

This possible action of the robot will change the state of the world as given in Figure 3.2. If the door is open in the real world, regardless of whether the instruction is followed or not, the robot’s action does not change the state of the world and this is shown by the loops over the $\beta$-worlds in Figure 3.2. Suppose the door is closed, there
is an 80% chance that the robot follows the instruction and opens the door and a 20% chance that the robot will leave the door as is. To begin with, let us assume that the door is closed and that the belief state is given by $P_1$ (as given in Section 3.3). In case the robot was completely reliable the new probability distribution would have been given by $P_0$. However the robot is only 80% reliable and hence $P_1$ should be transformed to $P_2$ where

\begin{align*}
P_2(\{\alpha, \beta\}) &= 0.6 \\
P_2(\{\overline{\alpha}, \beta\}) &= 0.2 \\
P_2(\{\alpha, \overline{\beta}\}) &= 0.15 \\
P_2(\{\overline{\alpha}, \overline{\beta}\}) &= 0.05.
\end{align*}

In Example 3.2 the observation made demands the enhancement of the initial probability assigned to $\beta$. Partial imaging $P_1$ with respect to $\beta$, and $\alpha = 0.8$, results in a transfer of 80% of probability associated with $\{\overline{\alpha}, \overline{\beta}\}$ to $\{\overline{\alpha}, \beta\}$. Similarly 80% of probability associated with $\{\alpha, \overline{\beta}\}$ is transferred to $\{\alpha, \beta\}$. Therefore, $P^p_\mathcal{I}$, where $\mathcal{I} = \langle \beta, 0.8 \rangle$, is the required $P_2$.\(^3\) In the above example, $\neg \beta$ was initially a belief of the agent but the belief state is altered and represented by $P^p_\mathcal{I}$, and $\neg \beta$ is deposed from the set of beliefs. Thus partial imaging by a sentence $\xi$ models a belief removal function when $\neg \xi$ is initially believed by the agent and $0 < a \leq 1$.

Since imaging models a dynamic belief inclusion function, it is natural to assume that partial imaging would model a dynamic belief removal function. When the agent removes a belief due to possible changes to the state of the world, the belief change is called belief

\(^3\)To be consistent with the notations, we should actually write $P^p_{1z}$, since it is $P_1$ that is being modified, but for the sake of readability, we use $P^p_\mathcal{I}$.
erasure [53]. Therefore we study the relation between belief erasure and partial imaging.

Let \( \gamma \) be a sentence in the language \( \mathcal{L} \). We denote the erasure operator by \( \ominus \) and \( P_\gamma \ominus \) denotes the result of erasure of \( \gamma \) from \( P \). Katsuno and Mendelzon [53] present a set of rationality postulates which guide an erasure function. The probabilistic analogues of these erasure postulates are as follows:

- **E1** If \( P_\gamma \ominus (\delta) = 1 \) then \( P(\delta) = 1 \).
- **E2** If \( P(\gamma) = 0 \) then \( P_\gamma \ominus = P \).
- **E3** If \( \not\vdash \gamma \), then \( P_\gamma \ominus (\gamma) < 1 \).
- **E4** If \( \vdash \gamma \leftrightarrow \delta \) then \( P_\gamma \ominus = P_\delta \ominus \).
- **E5** \( (P_\gamma \oplus)_+ = P \), where \( + \) denotes expansion of the belief state.
- **E6** \( (PaP')_\gamma \ominus = P_\gamma \ominus aP_\gamma \ominus \) where \( 0 \leq a \leq 1 \).

Any function which changes the given probability function such that the properties **E1** to **E6** are satisfied is said to be a probabilistic erasure function. **E1** states that erasure does not add anything new to the set of beliefs. Postulate **E2** presents the trivial case. \( P(\gamma) = 0 \) represents the case where the agent assigns zero probability to every \( \gamma \)-world, that is, the agent considers none of the \( \gamma \)-worlds to be a possible representation of state of the real world. Therefore, erasure by \( \gamma \) makes no change to the initial probability assignment. According to postulate **E3**, as long as the sentence being erased is not a theorem, erasure would be successful. Postulate **E4** states that erasure is not syntax sensitive, that is, when the content of two sentences being erased is the same the result of erasure is also the same. **E5** is the recovery postulate. Belief expansion is modelled in terms of Bayesian conditionalization. If an agent erases a belief from its belief state, and then expands by the same sentence, it gets back the initial belief state. It must be noted that probabilistic expansion is modelled by Bayesian conditionalization. A mix of probability functions \( PaP' \), \( 0 \leq a \leq 1 \), represents a “disjunction” of belief states represented by \( P \) and \( P' \) respectively.

Postulate **E6** states that erasure of a belief from a disjunction of belief states results in
disjunction of erasure of the belief from the individual disjuncts. This postulate ensures that each world that is considered as a possible state of the real world is changed individually. The postulate $E6$ is a counter-part of the postulate $U8$ given in [53] which is regarded as a non-probabilistic version of the homomorphic condition about probabilistic revision functions in [29]. A function $*$ which changes probability functions is said to be homomorphic if and only if, for all probability functions $P$ and sentence $\gamma$, for every $P_1$, $P_2$ and $a \in [0, 1]$ such that $P = P_1 a P_2$ it holds that $P^{*}_\gamma = (P_1)^*_\gamma a (P_2)^*_\gamma$. Gärdenfors shows that imaging is homomorphic [29]. Pearl, in [77], identifies homomorphism as the property which enables modelling actions using imaging. We find that partial imaging is homomorphic as well.

**Observation 3.1:** Partial imaging is homomorphic.

The proofs for all results are presented in Appendix B. Observation 3.1 plays an important role in our discussion on the relation between partial imaging and belief erasure as demonstrated by the following result.

**Theorem 3.1:** Partial imaging is a belief erasure function when $0 < a < 1$.

Therefore, we can define erasure in a probabilistic setting in terms of partial imaging as follows. Let $P$ represent the belief state of the agent and $\xi$ be the belief being erased from the belief state. Then for some $a \in (0, 1)$,

$$P^{\xi} = P^{p}_{\xi, a}$$  \hspace{1cm} (E)
3.4.2 Jeffrey-style Generalization of Imaging

Let us recall Equation 3.2. It states that \( P_{\xi,a} = P_{\xi}^# a P \). When \( P(\neg \xi) = 1 \), clearly \( P_{\neg \xi} = P \). Therefore Equation 3.2 becomes

\[
P_{\xi,a} = P_{\xi}^# a P_{\neg \xi}^#
\]  

(2R)

The probability of \( \xi \) under such \( P_{\xi,a} \), given by Equation 2R, is \( a \) and the probability of \( \neg \xi \) is \( 1 - a \). Imagine a situation where the agent makes an observation which demands the probability of \( \xi \) to be \( a \) (and hence probability of \( \neg \xi \) to be \( 1 - a \)). Equation 2R enables us to change the given probability function \( P \) to a new probability function which satisfies the demands of the observation irrespective of the initial probabilities of \( \xi \) and \( \neg \xi \). This implication of Equation 2R is similar to the generalization of Bayesian conditionalization proposed by Jeffrey [50]. In this section we investigate this connection between partial imaging and Jeffrey conditionalization. We present a brief discussion of conditionalization and Jeffrey’s generalization before showing how partial imaging generalizes imaging in a similar vein. Consider the following example.

**Example 3.3:** Let us assume that the agent initially believes the window is open, that is, \( \alpha \). The neighbour calls the agent to inform that the door is open, that is, \( \beta \). However the neighbour fails to inform the agent regarding the state of the window.

Let the initial belief state be given by \( P_3 \):

\[
P_3(\{\alpha, \beta\}) = 0.6
\]

\[
P_3(\{\alpha, \beta\}) = 0
\]

\[
P_3(\{\alpha, \overline{\beta}\}) = 0.4
\]

\[
P_3(\{\overline{\alpha}, \overline{\beta}\}) = 0.
\]

The information received, \( \beta \), is consistent with the beliefs held by the agent and the required belief change is a simple belief expansion [2, 29]. The information \( \beta \) does not give
the agent any reason to doubt $\alpha$, that is, to doubt the state of the window. Since the agent already considers $\{\alpha, \beta\}$ to possibly model the real world, the agent now believes that both the window and the door are open. That is, $P_3$ is changed to $P_4$, where

$$P_4(\{\alpha, \beta\}) = 1$$
$$P_4(\{\overline{\alpha}, \beta\}) = 0$$
$$P_4(\{\alpha, \overline{\beta}\}) = 0$$
$$P_4(\{\overline{\alpha}, \overline{\beta}\}) = 0.$$

This change is easily modelled by Bayesian conditionalization.

$$P_4 = P_3(\cdot|\beta).$$ (C)

Belief expansion is one of the belief change operations defined in [2]. Conditionalization is generally suggested as an intuitive model for expansion. However it has been argued that it is more appropriate to view belief expansion and conditionalization as different kinds of belief change [94]. But we do not discuss this issue here, since it is outside the scope of this paper. The pros and cons of conditionalizing the agent’s probability function based on the evidence is discussed in [50]. Jeffrey presents an argument to suggest that conditionalization is not best suited to change the probability function in all contexts even when the evidence or newly obtained piece of information is consistent with existing beliefs. Jeffrey states that conditionalization can be applied only at special occasions. We quote [50] (pg. 165):

The process of conditionalization, of the agent’s changing his subjective probability assignment from $\text{prob}$ to $\text{prob}_E$ upon learning that the evidence-proposition $E$ is true, is thus justified when it is applicable. However, there are cases in which a change in the probability assignment is clearly called for, but where the device of conditionalization cannot be applied because the change is not occasioned simply by learning of the truth of some proposition $E$. In particular, the change might be occasioned by an observation, but there
might be no proposition $E$ in the agent’s preference ranking of which it can
correctly be said that what the agent learned from his observation is that $E$ is
true.

When there is no proposition in the agent’s language that can exactly represent the ob-
servation, Jeffrey argues that the conditionalization should be based on the results of the
observation rather than by the observation itself.

**Definition 3.3:** Let $P$ denote the initial probability function, $\xi$ the evidence and $a$ the
probability of $\xi$ after observation. Then the final probability function is given by,

$$P_{\xi,a}(\cdot) = P(\cdot|\xi) \cdot a + P(\cdot|\neg\xi) \cdot (1 - a)$$

*(J)*

where $P(\xi)$ is neither 0 nor 1.

Equation J is generally termed as *Jeffrey conditionalization*. We illustrate this with the
following example.

**Example 3.4:** Over a period of time the agent realizes that the neighbour’s eye-sight has
deteriorated and is not completely reliable. The agent attributes 80% reliability to the
neighbour’s eye-sight. Let the agent’s initial belief state be given by $P_3$. Now when the
neighbour informs the agent that the door to the room is open, the agent does not accept
the neighbour’s information at its face value.

Let us suppose that the agent changes its belief state $P_3$ to $P_5$. Based on the neigh-
bour’s information, there is 80% chance that the door is open and a 20% chance that the
door is closed, that is, $P_5$ should be such that $P_5(\beta) = 0.8$ and $P_5(\neg\beta) = 0.2$. Here
the observation made or the information received cannot be represented by any particular
proposition having a probability of 1. No Bayesian conditionalization on $P_3$ can give us
the required $P_5$ because there is no proposition in the agent’s language that can represent the observation made. Example 3.4 represents the sort of situation highlighted by Jeffrey and therefore the required change from $P_3$ to $P_5$ can be modelled with the help of Definition 3.3. Jeffrey conditionalization based on the input that the final probability of $\beta$ is 0.8 gives the required $P_5$.

$$P_5(\{\alpha, \beta\}) = 0.8$$

$$P_5(\{\overline{\alpha}, \beta\}) = 0$$

$$P_5(\{\alpha, \overline{\beta}\}) = 0.2$$

$$P_5(\{\overline{\alpha}, \overline{\beta}\}) = 0.$$

Similar to conditionalization, imaging is also not capable of handling an observation which does not point to the truth of any proposition. A Jeffrey-style generalization is needed in case of imaging. In Example 3.2 we presented a situation where the robot is not completely reliable. Only some of the instructions are carried out by the robot. The agent knows that the robot is not reliable and does not completely expect the robot to open the door and hence does not believe that the door is open. The observation or the evidence in this example does not point to the truth of any sentence in the language. The change in the subjective probability function is modelled by partial imaging with respect to the result of the observation rather than the observation itself. Hence we believe partial imaging is the required generalization of imaging. Equation 2R highlights this property.

The two Equations 2R and J present a very simplified case where the observation talks about two sentences $\beta$ and $\overline{\beta}$ which form mutually exclusive and collectively exhaustive partitions of $\Omega$. Let us consider a more general case. Let $[\xi_i]$ where $0 \leq i \leq n - 1$ be mutually exclusive and collectively exhaustive sets of worlds and none of the $\xi_i$’s have an initial probability 0 or 1, that is, all $\xi_i$’s have positive prior probability. Consider $a_i$ where $0 \leq i \leq n - 1$ to be real numbers between 0 and 1 such that $\sum_{i=0}^{n-1} a_i = 1$. Suppose the agent makes an observation which dictates that the probability of each $\xi_i$ is $a_i$. Since all the partitions $\xi_i$’s have positive initial probability, the observation made is consistent.
with the initial belief state.\textsuperscript{4} Let \( I \) denote the set of all pairs \( \xi_i \) and \( a_i \). Extending Jeffrey’s probability kinematics to these \( \xi_i \)’s results in following final probability function.

\[
P^I_J(.) = \sum_{i=0}^{n-1} P(|\xi_i) \cdot a_i. \tag{3.4}
\]

Now, irrespective of the value of \( P(\xi_i) \) for any \( i \) between 0 and \( n - 1 \), partial imaging \( P \) with respect to \( I \) gives,

\[
P^p_I(.) = \sum_{i=0}^{n-1} P^\#(\xi_i) \cdot a_i. \tag{3.5}
\]

Thus it is clear that partial imaging mirrors Jeffrey-style generalization with respect to imaging. In this section we have explored one particular generalization of imaging which was brought about by a simple variation to the notion of imaging. There are many other possible variations and in the following section we study the implication of one of them. This variation leads to a different generalization of imaging, namely selective imaging. We show that it models a very special belief inclusion function, namely conditional update.

### 3.5 Selective Imaging

Suppose the robot receives an instruction of the form ‘\textit{In case the air conditioner fails, open the window}’. Instructions of this kind are based on the uncertainty associated with the antecedent, that is, the functioning of the air conditioner. We call instructions of this kind as \textit{conditional instructions}. When the agent learns of such an instruction being given to the robot, or when the agent gives such an instruction to the robot itself, it expects the robot to follow the same. Incorporating such conditional instructions in the belief state is not a simple task of belief update by the material implication: \textit{air conditioner fails \( \rightarrow \) the window is open}. According to the conditional instruction, the robot is instructed to

\textsuperscript{4}It should be noted that a sentence is said to be consistent with the belief state if and only if it has non-zero initial probability. Since each \( \xi_i \) has positive initial probability they all are consistent with the belief state.
open the window when the air conditioner fails; while the material implication suggests that the window is open when the air conditioner fails. The material implication assumes certain connection between the window being open and the air conditioner failing, which is not intended by the conditional instruction. Moreover, when the conditional instruction is received, the robot performs an action which could change the state of the world when the antecedent holds true in the initial state. Thus the antecedent and the consequent in a conditional instruction concern different states of the world. However, the material implication represents the sentence if the air conditioner fails then the window is open; the antecedent and the consequent here describe the same state of the world. Thus conditional instructions differ from conditional sentences, that latter being well studied in the literature [8, 93]. In this section we propose selective imaging, a generalization of imaging that can help model updating by conditional instructions.

As the name selective imaging suggests, there is some selection involved. When selective imaging a probability function $P$ by a sentence $\xi$, some possible worlds in $\Omega$ are selected and the probability mass associated with those worlds are then shifted to their corresponding $\#$-minimal $\xi$-worlds. When a selected world $\omega$ is a model of $\xi$, the probability associated with $\omega$ is transferred to its $\#$-minimal $\xi$-world, which is $\omega$ itself. Selective imaging uses a selection function to select the set of possible worlds which will lose the probability mass associated with them. We denote this selection function by $S$ and the probability distribution resulting from selective imaging $P$ with respect to $\xi$ by $P^S_{\xi}$.

**Definition 3.4:** A function $S$ from the set of all possible probability distributions of the form $S : \mathcal{P} \rightarrow 2^\Omega$ is called a selection function.

Let $\eta$ be a sentence in $\mathcal{L}$ such that $[\eta] = S(P)$. When selective imaging by $\xi$ the worlds in $S(P)$, that is, all the $\eta$-worlds, lose their probability to their corresponding $\#$-minimal
We define selective imaging formally as follows:

**Definition 3.5:** Given a selection function \( S \), the selective imaging of \( P \) with respect to a sentence \( \xi \) is defined as follows:

\[
\text{for every } \upsilon \in \Omega, \quad P^\xi_\upsilon(\upsilon) = \sum_{\Omega} P(\omega) \cdot \omega^\#(\upsilon) \cdot \omega(\eta) + P(\upsilon) \cdot \upsilon(\neg \eta)
\]

where sentence \( \eta \) in \( \mathcal{L} \) is such that \([\eta] = S(P)\).

The probability of any sentence \( \alpha \) as a result of selective imaging by \( \xi \) is,

\[
P^\xi_\upsilon(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega^\#(\alpha) \cdot \omega(\eta) + P(\alpha \land \neg \eta).
\]

(3.6)

The resultant probability of \( \xi \) is

\[
P^\xi_\upsilon(\xi) = P(\eta) + P(\xi \land \neg \eta) = P(\eta \lor \xi)
\]

(3.6a)

and the resultant probability of its negation is

\[
P^\xi_\upsilon(\neg \xi) = P(\neg \xi \land \neg \eta).
\]

(3.6b)

As in partial imaging, there are some interesting special cases that arise when considering different possibilities with respect to the selection function \( S \).

1. **Case \( P(\eta) = 1 \).** This happens when all the possible worlds with positive probability mass under \( P \) are selected by the selection function. In this case, selective imaging reduces to Lewis’s imaging. In other words, Imaging is a special case of selective imaging.

\[5\text{We use } S \text{ instead of } S(P) \text{ when the probability function being referred to is clear from the context.}\]
Observation 3.2: If \( P(\eta) = 1 \) then \( P^a_\xi = P^#_\xi \).

2. Case \( P(\eta) = 0 \). In this case, the probability mass associated with every selected possible world is zero. There is no shift of probability mass at all. Therefore there is no change in the initial probability function \( P \) upon selective imaging, that is, \( P^a_\xi = P \).

Observation 3.3: When \( P(\eta) = 0 \) then \( P^a_\xi = P \).

More special cases arise based on the initial probability of \( \xi \).

3. Case \( P(\xi) = 1 \). When the initial probability accorded to \( \xi \) is 1, it follows that all the worlds with initial non-zero probability mass are models of \( \xi \). It should be noted that the \#-minimal \( \xi \)-world of any \( \xi \)-world is itself. Hence when the selection function returns any set of worlds to lose their probability mass to associated \( \xi \)-worlds, there is no shift of probability mass. Selective imaging with respect to sentence \( \xi \) leaves the initial probability function unchanged. Therefore, \( P^a_\xi = P \).

4. Case \( P(\xi) = 0 \). There are two cases that can possibly arise here, (a) \( P(\eta) = 0 \) and (b) \( P(\eta) \neq 0 \). When \( P(\eta) = 0 \), as explained in case 2 above, we have \( P^a_\xi = P \). Alternatively, suppose we have \( P(\eta) \neq 0 \). It follows that there exists a model of \( \eta \) which is assigned an initial non-zero probability. Since this model is selected by the selection function, it loses its probability mass to its corresponding \#-minimal \( \xi \)-world. Moreover we have \( P(\xi) = 0 \), from which we note that \( P(\neg\xi) = 1 \). Therefore, this particular model of \( \eta \) with positive probability mass is also a model of \( \neg\xi \). Thus there is a shift of probability mass from at least one \( \neg\xi \)-world to a \( \xi \)-world, whereby the probability of \( \neg\xi \) is reduced from 1 to a value less than 1. Since \( \neg\xi \) is an initial belief of the agent, this case reflects disposal of \( \neg\xi \) from its belief set. We investigate the possibility of selective imaging modelling belief erasure function in Section 3.5.1.

Before we discuss the application of selective imaging in the study of belief change we
present the following two observations.

**Observation 3.4:** If $\eta \vdash \xi$ then $P^s_\xi = P$.

If the input sentence $\xi$ is such that the sentence $\eta$ entails $\xi$ then selective imaging does not change the probability function.

**Observation 3.5:** If $\eta \vdash \neg \xi$ then $P^s_\xi(\neg \eta) = 1$.

It can also be seen that if the sentence $\eta$ entails the negation of the input the result of selective imaging results in acquiring belief in $\neg \eta$.

### 3.5.1 Selective Imaging as a belief change function

Let us re-visit the case 4(b) discussed above. Selective imaging of $P$ by $\xi$ reduces the probability of $\neg \xi$. As a result, the probability of $\neg \xi$ becomes less than 1, and hence $\neg \xi$ is no longer a belief in the (resultant) belief state. The following theorem presents the relation between selective imaging and belief erasure.

**Theorem 3.2:** If the selection function $S$ is such that the associated sentence $\eta$ has a non-zero probability, then selective imaging satisfies the erasure postulates $E2$, $E3$ and $E4$.

Since selective imaging satisfies $E3$, the belief that needs to be removed is successfully withdrawn upon selective imaging. However selective imaging does not satisfy all the postulates of erasure. In particular, it does not satisfy the postulate $E1$ meaning that it is possible to acquire new beliefs via selective imaging. Therefore it is not appropriate to term selective imaging as a belief removal function. Since imaging is a special case
of selective imaging, selective imaging models a variation of belief update (discussed in case 1 in Section 3.5). We claim that selective imaging models a special update function which we term *conditional update*.

We re-visit the ‘agent-robot’ setup we used earlier in Examples 3.1 and 3.2. Consider the scenario where there is a set of rules pre-programmed into the robot’s system. Let these rules be conveniently listed in the user’s manual. Suppose the agent finds the following conditional instruction listed in the manual:

‘In case of gas leak, open the window’.

We call such instructions *conditional instructions* in order to differentiate from simpler instructions such as *open the window* or *close the curtains*. Suppose the sentence \( \eta \) denotes ‘there is gas leak in the room’. In Examples 3.1 and 3.2, we used \( \alpha \) to denote ‘window is open’. The conditional instruction or rule given above is denoted by \( \eta \rightarrow \alpha \). The symbol \( \rightarrow \) is not a logical connective and does not have a truth-value, while the sentences \( \eta \) and \( \alpha \) have truth-values. The conditional instruction states that when the condition (that is, existence of gas leak) is satisfied, the robot opens the window as instructed. We can roughly translate \( \eta \rightarrow \alpha \) as: if \( \eta \) is found to be the case (true), make \( \alpha \) the case (true).

When the agent believes that there is a gas leak in the room, the agent assigns a probability of 1 to \( \eta \); otherwise the probability assigned to \( \eta \) is something strictly less than 1. In particular, when the agent has not overruled the possibility of a gas leak, it assigns a positive probability to \( \eta \). The initial probability function, which represents the belief state of the agent, is updated by \( \eta \rightarrow \alpha \), and we represent the new belief state by \( P_{\eta \rightarrow \alpha} \). When updating by \( \eta \leftarrow \xi \), the required change takes place as follows. When \( \eta \) holds in the real world, according to the instruction given to the robot, the state of the world is modified unless the window is already open. Suppose the sentence \( \eta \) has a non-zero initial probability, that is, some models of \( \eta \) have non-zero probability. Then the change required to model update by \( \eta \leftarrow \xi \) demands that the probability associated with every model of \( \eta \) be transferred to the respective \#-minimal \( \xi \)-world.
Update by the sentence $\eta \leftrightarrow \xi$ is conditional upon the agent’s epistemic attitude towards $\eta$. Therefore we term update by a conditional instruction as *conditional update*. According to Definition 3.4, the selection function is independent of the input received. We alter Definition 3.4 in order to make it dependent on the input received, particularly the antecedent of the conditional instruction with which the state needs to be updated.

**Definition 3.6:** A function $S'$ from the set of all possible probability distributions and the set of all conditional instructions in the language, $S' : \mathcal{P} \times C \rightarrow 2^\Omega$, is called a context-sensitive selection function if and only if $S'(P, \eta \leftrightarrow \alpha) = [\eta]$.

It is worth noting that a context-sensitive selection function, as defined in Definition 3.6, is a variant of the selection function, as defined in Definition 3.4. Selective imaging in terms of Definition 3.6 will help in modelling conditional update. Given a conditional instruction $\eta \rightarrow \alpha$, context-sensitive selection function returns the set of all models of $\eta$. Thus, selective imaging with context-sensitive selection function would result in the probability mass of each $\eta$-world being transferred to its $\#$-minimal $\xi$-world. Therefore update by the conditional input $\eta \leftrightarrow \xi$ is given by selective imaging with respect to $\xi$ based on context-sensitive selection function $S'$, that is,

$$P_{\eta \rightarrow \xi}^\# = P_{\xi}^\# \quad \text{(CU)}$$

When the agent believes that there is a gas leak, the agent believes that the robot would have opened the window according to instructions given to it. Then the conditional input *in case of a gas leak, open the window* becomes an instruction to open the window. In such a case selective imaging by $\xi$ gives the same result as imaging as pointed out in Observation 3.2. Suppose instead that the agent believes that there is no gas leak, that is, $P(\neg \eta) = 1$. Then the agent does not believe that the robot will change the state of the window. Thus the agent does not change its belief state when updating by $\eta \leftrightarrow \alpha$. This
is shown in Observation 3.3. Apart from these boundary cases, this account appropriately handles conditional update when the antecedent is believed neither to be true nor to be false, that is, \(0 < P(\eta) < 1\).

Thus selective imaging, when coupled with a context-sensitive selection function, is appropriate for modelling conditional instructions. Partial imaging and selective imaging naturally lead to another variation of imaging that we will call *selective partial imaging*. In the following Section 3.6, we show that selective partial imaging models *conditional erasure*. Conditional erasure is the belief removal counterpart of conditional update.

### 3.6 Selective Partial Imaging

In Section 3.5 we observed that selective imaging is capable of modelling belief update based on conditional instructions. The agent updates its belief state by including the conditional instruction knowing that the robot will surely follow the instruction. However, when the agent does not believe all the instructions are being carried out and that there is a good chance this conditional instruction may not be complied with, it does not update its belief state by the conditional instruction. The belief state is changed by what we term a *conditional erasure*. We propose another variation of imaging, namely selective partial imaging, which will be useful in modelling this situation.

In the previous two sections, we have discussed two generalizations of imaging. In case of partial imaging only a share of the probability mass associated with each possible world is shifted to corresponding \(\xi\)-world. In selective imaging, the total probability mass associated with only some chosen possible worlds is shifted to the corresponding \(\xi\)-world. A combination of selective and partial imaging suggests a variation of imaging where *some* possible worlds lose a fixed *share* of their probability mass to the corresponding \(\xi\)-world. We call this variation *selective partial imaging*.

Suppose the observation made by the agent demands an enhancement of the probability of \(\xi\). Selective partial imaging uses a selection function, as described by Definition 3.4,
to choose a set of worlds from $\Omega$. Let the selection function be such that the set of worlds chosen are exactly the models of the sentence $\eta$. All the selected worlds lose a share of the associated probability mass to their respective $\#$-minimal $\xi$-worlds. We suggest that all the selected worlds lose uniformly the same proportion of the probability mass. The share of probability lost by the selected worlds is denoted by $a$, where $a$ is a real number between 0 and 1. For each world $\omega \in S(P)$, $\omega$ loses $a$-share of its initial probability, that is, $a \cdot P(\omega)$ and retains the rest of its initial probability, namely, $(1 - a) \cdot P(\omega)$. The $\#$-minimal $\xi$-world with respect to $\omega$ gains the lost probability, namely, $a \cdot P(\omega)$. We denote the result of such a change to the probability function $P$ by $P^{sp}_I$, where $I$ denotes the pair of sentence $\xi$ and a real number $a \in [0, 1]$.

**Definition 3.7:** Given a selection function $S$, selective partial imaging of $P$ with respect to sentence $\xi$ and $0 \leq a \leq 1$ is a probability distribution over $\Omega$ defined as follows: for every $\upsilon \in \Omega$,

$$P^{sp}_I(\upsilon) = a \cdot \sum_{\Omega} P(\omega) \cdot \omega^{\#}(\upsilon) \cdot \omega(\eta) + (1 - a) \cdot P(\upsilon) \cdot \upsilon(\eta) + P(\upsilon) \cdot \upsilon(\neg \eta)$$

The probability of any sentence $\alpha$ as a result of selective partial imaging (henceforth SPI) is given by,

$$P^{sp}_I(\alpha) = a \cdot \sum_{\Omega} P(\omega) \cdot \omega^{\#}(\alpha) \cdot \omega(\eta) + (1 - a) \cdot P(\alpha \land \eta) + P(\alpha \land \neg \eta). \quad (3.7)$$

From Equation 3.7 we can see that the resultant probability of $\neg \xi$ is

$$P^{sp}_I(\neg \xi) = P(\neg \xi) - a \cdot P(\eta \land \neg \xi) \quad (3.7a)$$

and the resultant probability of $\xi$ is

$$P^{sp}_I(\xi) = P(\xi) + a \cdot P(\eta \land \neg \xi) \quad (3.7b)$$
Just as in the cases of partial imaging and selective imaging there are a number of special cases that arise depending on the selection function $S$ and the value of the real number $a$.

1. Case $P(\eta) = 1$. When the selection function chooses all the possible worlds in $\Omega$ or just chooses the set of possible worlds which have non-zero initial probability according to $P$ (that is, $S = \Omega$ or $S = ||P||$), the initial probability assigned to the sentence $\eta$ is 1. In such case each possible world loses $a$-share of its probability to its corresponding $\#$-minimal $\xi$-world. Thus $P_{I}^{sp} = P_{I}^{p}$. Partial imaging, therefore, is a special case of selective partial imaging.

Observation 3.6: When $P(\eta) = 1$, $P_{I}^{sp} = P_{I}^{p}$.

2. Case $P(\eta) = 0$. When the set of worlds chosen by the selection function is a null-set or when all the selected worlds have zero initial probability, the probability mass that is shifted to the corresponding $\xi$-worlds is nil. Therefore, there is no change to the initial probability function as a result of SPI. Hence we have $P_{I}^{sp} = P$.

3. Case $a = 0$. In this case the probability mass that is shifted from the selected worlds is zero. Hence there is no change to the initial probability function, which leads us to the same conclusion as in the previous case, that is, $P_{I}^{sp} = P_{I}^{s}$.

4. Case $a = 1$. When the real number $a$ takes the value 1, it denotes the case where the entire probability mass associated with every selected world is shifted to the corresponding $\#$-minimal $\xi$-world. Thus SPI is a generalization of selective imaging.

Observation 3.7: When $a = 1$, $P_{I}^{sp} = P_{\xi}^{s}$.

Cases 1 and 4 show that selective partial imaging is a generalization of both selective and partial imaging. Since both selective and partial imaging are generalizations of imaging
themselves, clearly selective partial imaging is a generalization of imaging.

5. Case $P(\eta) = 1$ and $a = 1$. In this case, all the worlds with non-zero probability are chosen by the selection function and the entire probability mass associated with each possible world is shifted to their corresponding $\xi$-world. Thus imaging is a special case of SPI.

Observation 3.8: If $P(\eta) = 1$ and $a = 1$, then $P_{\xi}^{sp} = P_{\xi}^\#$.

Two more special cases arise depending on the initial probability of $\xi$.

6. Case $P(\xi) = 1$. Any possible world $\omega$ which is a model of $\neg\xi$ has zero initial probability mass associated with it. Hence if it were chosen by the selection function, there will be no change in the probability mass associated with $\omega$ upon SPI with respect to $\xi$.

On the other hand, if $\omega$ is a model of $\xi$ and if it were selected by $S$, since $\omega_{\xi}^\# = \omega$, the world $\omega$ retains its initial probability. Therefore, there is no shift of probability mass from any possible world in $\Omega$. Hence $P_{\xi}^{sp} = P$.

7. Case $P(\xi) = 0$. If either $P(\eta) = 0$ or $a = 0$, from cases 2 and 3 above we realize that it leads to trivial cases of no change to the initial probability function. Hence in the following discussion we assume that $P(\eta) \neq 0$ and $a \neq 0$. Selective partial imaging with respect to $\xi$ in this case reduces the probability associated with the sentence $\neg\xi$. SPI with respect to $\xi$ reduces the probability of $\neg\xi$ from 1 to $(1 - a) \cdot P(\eta)$. Thus the sentence $\neg\xi$ which was initially a belief of the agent, loses its status as a belief. This shows that SPI provides an account of belief removal. We explore this property further in the following subsection.

3.6.1 Selective Partial Imaging as a belief change function

Let $\xi$ be a sentence in the language which has zero initial probability. Then its negation, $\neg\xi$, has the maximum initial probability ($P(\neg\xi) = 1$), that is, $\neg\xi$ is a belief of the agent.
Let the selection function $S$ be such that $P(\eta) \neq 0$ and the real number $a$ take a value from the set $(0, 1]$. From case 7 above, it is clear that SPI with respect to $\xi$ changes the initial probability function $P$ leading to discarding the belief $\neg \xi$. Thus SPI models removal of belief in $\neg \xi$. However it turns out that SPI falls short of satisfying the six postulates of belief erasure.

**Theorem 3.3:** Let the selection function $S$ be such that the associated sentence $\eta$ has a non-zero probability and $0 < a < 1$. Then the selective partial imaging function is a belief removal function that satisfies $E1, E2, E3$ and $E4$.

When the given probability function is changed by SPI with respect to $\xi$, no new beliefs are introduced in the belief state since SPI satisfies $E1$. Also the result of SPI ensures withdrawal of belief in $\neg \xi$ because SPI satisfies $E2$. Therefore SPI is purely a belief removal function. It must be noted that the variable $a$ is such that $0 < a < 1$, thus dismissing the possibility of $a = 1$. We ensure that $a \neq 1$ since, when $a = 1$, SPI reduces to selective imaging which does not satisfy $E1$. Nonetheless SPI does not model belief erasure, since it clearly does not satisfy postulate $E5$. Consider the case where $P(\neg \xi) = 1$, that is, $\neg \xi$ is a belief. It is quite possible that the sentence $\eta$, based on the selection function $S$, is such that $0 < P(\eta) < 1$. Now, the probability associated with the $\eta$-worlds are shifted to their corresponding $\#$-minimal $\xi$-worlds. Therefore the proportion among the $\neg \xi$-worlds based on their individual probabilities wrt to the distributions $P^\text{sp}_{\xi}$ and $P$ are different. Hence, $P^\text{sp}_{\xi}$ does not reduce to $P$ when conditionalized by $\neg \xi$. Thus SPI does not satisfy all the six postulates.

Let us assume that the agent believes the window is closed, that is, $\neg \alpha$. We also assume that the robot is ‘inefficient’ (as discussed in Example 3.2) and performs only 80% of its instruction set. Suppose the agent learns that the robot is instructed to open the window in the event of a gas leak, that is, $\eta \hookrightarrow \alpha$ where $\eta$ denotes *there is gas leak in the room*. Let us assume that the agent considers an $\eta$-world to be a possible representation
of the state of the real world. If the window is open in that world, then the instruction has no effect on the state of the real world. However, if the window is closed in it, there is 80% chance that the $\eta$-world would be changed by the robot by opening the window, and there is a 20% chance that the robot will leave the state of the real world as is. In terms of probability associated with the worlds, consider an $\eta$-world, $v$, with positive probability under $P$. Suppose the agent considers $v$ to be a possible representation of the real world, and it assigns a positive probability to it. Therefore, 80% of the probability associated with $v$ is transferred to $v'_{\eta}$ and the remaining 20% probability remains with $v$. If $\alpha$ is true in $v$ then $v'_{\alpha} = v$ and probability of $v$ is not reduced. On the other hand, if $\omega$ is a $\neg\eta$-world, even if $P(\omega) > 0$, the given conditional instruction does not provide any reason for the agent to re-consider its attitude towards $\omega$. When there is no gas-leak in the room, the robot performs no action to change the state of the world. Therefore the probability associated with $\omega$ is not reduced as a result of the information received.

We see that a conditional instruction $\eta \hookrightarrow \xi$ and the uncertainty associated with the action being performed, present an opportunity for the probability mass associated with some $\neg\xi$-worlds to be moved to their respective $\#-$minimal $\xi$-world. If there exists an $\eta$-world that is also a model of $\neg\xi$ and if that particular world has positive probability under $P$, then SPI ensures that a fraction of this probability is moved to $\xi$. Thus for any belief, $\neg\xi$, probability of $\neg\xi$ is possibly reduced from 1 to strictly less than 1, thus demoting it from a belief to a non-belief. The belief in $\neg\xi$ is erased conditional upon the agent’s attitude towards $\eta$, that is, when the agent considers $\eta$ to be probable, that is, $P(\eta) > 0$. We term this change as *conditional erasure* and we denote erasure of $\neg\xi$ conditional on the agent’s attitude towards $\eta$ by $P_{\eta \hookrightarrow \neg\xi}^{\ominus}$. A conditional erasure can be modelled by selective partial imaging coupled with a context-sensitive selection function. Given $\eta \hookrightarrow \xi \in C$, a context-sensitive selection function $S'$ selects all the models of $\eta$. All these models lose $a$-share of their probability to their respective $\#-$minimal $\xi$-models.
Thus erasure of \( \neg \xi \) conditional on \( \eta \) is given by:

\[
P_{\eta \rightarrow \neg \xi} = P_{\xi,\alpha}^{sp} \tag{CE}
\]

If the agent believes that there is no gas-leak (that is, \( P(\eta) = 0 \)), then according to the agent this conditional instruction refers to a hypothetical situation. Therefore \( \eta \rightarrow \alpha \) does not require any change to be made to the agent’s set of beliefs. However, if the agent believes that there is a gas-leak (that is, \( P(\eta) = 1 \)), then the agent believes that the robot will change the state of the world by opening the window. Updating by \( \eta \rightarrow \alpha \) in this case reduces to updating by \( \alpha \). However since the robot is only 80% reliable the required belief change dictates erasure by \( \neg \alpha \) as noted in Observation 3.6.

So far, we have provided three generalizations of imaging – partial imaging, selective imaging and selective partial imaging. Figure 3.3 graphically illustrates these three generalizations. It is clear that each of them is capable of modelling removal of a belief. Having described different ways of removing a belief, we present the following comparison of the three methods.

**Theorem 3.4:** Assume a probability function \( P \), a selection function \( S \) and the corresponding sentence \( \eta \) such that \( P(\eta) > 0 \). Suppose \( K^p_\xi, K^s_\xi \) and \( K^{sp}_\xi \) represent the belief sets corresponding to \( P^p_\xi, P^s_\xi \) and \( P^{sp}_\xi \), respectively. Then,

\[
K^p_\xi \subseteq K^{sp}_\xi \subseteq K^s_\xi
\]

where \( I = (\xi, a) \) and \( 0 < a < 1 \).

From the above theorem we infer that partial imaging is the strongest belief removal function compared to the other two. When partial imaging by \( \xi \), the set of beliefs that are retained are also retained when selective imaging or SPI with respect to \( \xi \). We also find that beliefs retained in SPI are also retained when selective imaging with respect to \( \xi \).
Figure 3.3: (a) depicts how the probability function $P$ is changed upon imaging by a particular sentence. The change to the probability function $P$ in terms of (b) partial imaging, (c) selective imaging and (d) selective partial imaging is as shown. Here the arrows point to the corresponding $\#$-minimal $\xi$-world and the self-loop denotes that the corresponding $\#$-minimal $\xi$-world is itself.

### 3.7 Discussion and Conclusion

Probabilistic belief change has been widely studied in the literature [29, 33, 34, 44, 56, 91]. However, a cogent account of incorporation and removal of beliefs in the probabilistic context is still lacking. Our main goal in this paper was to redress this weakness by developing an account of probabilistic belief change (erasure and update). We have done it by exploiting the efficacy of imaging in dealing with self-consistent evidential sentences that have a prior probability of zero, the sentences that fall in the critical zone [67]. In particular, we used a generalized version of imaging that we dub partial imaging. In the process we developed a generalized account of imaging similar to Jeffrey's generalization of Bayesian conditionalization. It is worth noting that this approach deals with probabilistic belief removal only in a dynamic environment, that is, it provides an account of probabilistic belief erasure. It does not handle belief removal in a static environment.
(belief contraction). We suspect a proper account of probabilistic belief contraction is far more complicated, and it is a problem we would like to tackle in our future work.

Apart from partial imaging, we presented two other generalizations of imaging, selective imaging, and selective partial imaging. These are unlike any known belief change devices encountered in the literature. We have suggested two novel forms of belief change, namely, conditional update and conditional erasure, that we believe are respectively modeled by these two operations. Unlike the conditionals generally studied in the literature [8, 93], these belief change functions are based on conditional sentences which represent conditional instructions.

We conclude this paper by noting an assumption that David Lewis [63] makes, namely that for any pair of possible world \( \omega \) and sentence \( \xi \), there exists one particular model of \( \xi \) which is more similar to \( \omega \) than all other models of \( \xi \), that is, the choice function \( \# \) returns a single world. This assumption is not shared by Gärdenfors [29]. Gärdenfors instead assumes that for any given pair \( \omega \) and \( \xi \), there exist a set of models of \( \xi \) that are most similar to \( \omega \) among the models of \( \xi \), that is, the choice function \( \# \) returns a set of worlds. For most part in this paper, we have worked with Lewis’s uniqueness assumption. It is evident that discarding this assumption regarding \( \omega \xi \# \) does not affect our discussion on partial imaging. It will be interesting to study in future if Selective Imaging and Selective Partial Imaging are insensitive to this assumption as well.
Chapter 4

Probabilistic belief contraction

*Only entropy comes easy.*

Anton Chekhov.

4.1 Introduction

Probability has been widely used to reflect a rational agent’s epistemic attitude. Different probabilistic representations of an epistemic state have been discussed in the literature, ranging from a set of probability functions representing the epistemic state [33, 34, 49] to a single probability function [28, 29]. While the former representation reflects imprecise probabilities that can often be associated with events, the latter is a simpler version which reflects precise probabilities associated with events. In this work, we follow the latter approach and represent the epistemic state by a single probability function. In our representation, a sentence is said to be a belief of the agent when it has probability 1 [60]; it signifies that the agent is certain about the sentence being true.\(^1\)

Often the epistemic state or belief state of the agent needs to undergo some change. This change may involve inclusion of a new belief, removal of an existing belief or replacement of an existing belief by a new one. In order to change the belief state, the probability function representing it needs to be changed. Inclusion of a belief involves

\(^1\)A sentence with probability 1 is termed *full belief* in [60].
changing the probability function so that the probability associated with the sentence to be included is made 1. The simplest account of belief inclusion is given by conditionalization, that is, the ratio rule\(^2\) [28, 29, 50]. Such a belief inclusion, that is, inclusion in a static environment, is termed belief expansion. Belief inclusion has also been modelled with the aid of Lewis’ imaging [63]. However such an account models inclusion in a dynamic environment, namely, belief update [53] and the corresponding account of probabilistic belief removal, namely, belief erasure has been studied in [81]. Our objective here, however, is to present an account of belief removal in a static environment, namely, belief contraction [2].

Belief contraction involves changing the probability function representing the current belief state such that a sentence (the belief being contracted) with prior probability 1 is assigned some probability less than 1. Therefore belief contraction requires an account of inverse conditionalization – that which reverses the effect of conditionalization. Study of probabilistic belief contraction has been relatively neglected with the probable exceptions of [28, 29, 74, 82]. Some guiding principles for probabilistic contraction have been proposed in [28, 29]. These principles are known as the AGM contraction postulates. Olsson [74] observes that these postulates do not sufficiently describe the result of contraction and proposes an additional constraint. But these postulates do not conclusively result in the construction of the contracted probability function.

Olsson’s postulate [74] compares the entropy of the initial and the contracted probability functions. Entropy measures have been often used in the study of belief change, especially belief inclusion [9, 35, 55, 56]. These approaches have been based on the principles of maximum entropy (commonly known as maxent) [14, 48, 75] and the principle of minimum cross entropy [14, 54, 90]. An account of inverse conditionalization based on entropy measures is presented in [82]. This is, to the best of our knowledge, the first account of a successful (that is, nontrivial) belief contraction in a probabilistic setting. In

\(^2\)Ratio rule: \(\text{prob}(h|e) = \frac{\text{prob}(h \land e)}{\text{prob}(e)}\) when \(\text{prob}(e) > 0\) and is not defined otherwise.
this account, the result of contraction is proposed as the probability function which has 
the maximum possible entropy value while satisfying the AGM postulates of contraction.
We term this contraction an \textit{indifferent contraction} and we note that it leads to excessive 
loss of information.

One of the guiding principles of belief change is the \textit{principle of informationa}
\textit{economy} [29, 87], which states that the loss of beliefs as a result of any belief change should 
be kept at a minimum. In [2, 29] three different contractions are discussed, namely, \textit{maxi-
choice contraction, partial meet contraction} and \textit{full meet contraction}. Among these 
three, full meet contraction results in a big change to the set of beliefs; while maxichoice 
contraction results in minimal loss of beliefs. When given a choice between two beliefs, 
under maxichoice contraction, one belief is retained and other is lost; while under full 
meet contraction, both are lost. Partial meet contraction strikes a middle ground between 
these two. It is based on the \textit{principle of preference and indifference} [87], wherein, given 
a choice between two beliefs, the preferred belief is retained; and when they are equally 
prefereed, they both are lost as a result of the change.

Indifferent contraction violates the principle of informational economy. We show that 
it, in fact, corresponds to a full meet contraction (Section 4.3). By that we mean, the 
indifferent contraction changes the belief state such that the change to the correpond-
ing set of beliefs is equivalent to the one caused by a full meet contraction acting on the 
belief set. Henceforth, in this paper, whenever we refer to a contraction as full meet or 
maxichoice or partial meet contraction, we mean that the contraction changes the belief 
state in such a way that the change to the corresponding set of beliefs is equivalent to 
the one caused by a full meet or maxichoice or partial meet contraction, respectively, act-
ing on the associated belief set. Our objective in this paper is to present an account of 
contraction that does not result in unnecessary loss of beliefs. In other words, we aim 
to present an account of partial meet contraction in a probabilistic setting. Towards this 
end, we present a variation of indifferent contraction which we term \textit{submaximal entropy}
contraction (Section 4.4), wherein we relax the condition that the contracted probability
function should have maximum possible uncertainty. However, as we show, even sub-
maximal entropy contraction corresponds to a full meet contraction. As an alternative, we
investigate an account of minimal contraction which results in a probability distribution
with the least possible uncertainty value (Section 4.4). It is observed that the minimal
contraction is very similar to a maxichoice contraction. Since partial meet contraction is
based on a notion of preference among the beliefs, we use a richer representation of the
belief state, where we represent the beliefs of the agent in terms of a probability distrib-
ution and the agent’s preferences among these beliefs in terms of a system of spheres
representation [32] (Section 4.5), to present an account of partial meet contraction. We
term the contraction operation in question preferential contraction.

Furthermore, different entropy measures have been studied in the literature, such as
Shannon entropy [89] and Hartley entropy [79]. The above mentioned contractions are
based on Shannon entropy measure. We finally present a discussion on the counterparts of
indifferent contraction, submaximal entropy contraction and minimal contraction based
on Hartley entropy measure (Section 4.6). We observe that an account of partial meet
contraction is easily obtainable from a submaximal entropy measure based on Hartley
entropy.

4.2 Preliminaries

In this paper we let \( L \) represent the propositional language generated from a finite set
of propositional atoms with the standard connectives \( \land, \lor, \neg \) and \( \rightarrow \). The sentences of
this language are denoted by lower case Greek letters (with or without any super- or
sub-scripts) such as \( \alpha \) and \( \beta \). The set of all possible worlds is denoted by \( \Omega \) whose
individual elements are denoted by \( \omega, \mu \) and \( \nu \) (with possible decorations). The set of
models of a sentence \( \alpha \) is the set of interpretations/possible worlds in which \( \alpha \) is true and
is represented by \([\alpha]\). We define \( \omega(\alpha) = 1 \) when \( \omega \in [\alpha] \) and \( \omega(\alpha) = 0 \) otherwise. A
classical consequence operator $Cn$ governs the background logic which is based on the classical consequence relation $\vdash$. The belief state is represented by a probability distribution over $\Omega$. We use $P$ and $Q$ (with and without decorations) to denote the probability distributions and the set of all probability distributions is denoted by $\mathcal{P}$. We denote the set of worlds which have nonzero probability under a distribution $P$ by $\|P\|$. It is also referred to as ‘support of $P$’. The probability associated with any sentence $\alpha$ is given by

$$P(\alpha) = \sum_{\omega \in \Omega} P(\omega) \cdot \omega(\alpha).$$

A sentence $\alpha$ is said to be consistent with a probability function $P$ if and only if $P(\alpha) > 0$, and the sentence is said to be inconsistent with $P$ otherwise. The set of beliefs associated with the belief state represented by $P$ is denoted by $\mathcal{K}_P$. The belief set $\mathcal{K}_P$ is given by the set of all sentences with probability 1,

$$\mathcal{K}_P = \{ \alpha \in \mathcal{L} : P(\alpha) = 1 \}. \quad (4.1)$$

When the context is clear we will avoid using subscripts for the belief set.

We use $P_{\alpha}^\circ$ to denote the probability distribution that results from performing the belief change operation of type $\circ$ on the belief state $P$ with respect to $\alpha$. The corresponding modified belief set is denoted by $\mathcal{K}_{\alpha}^\circ$. We use $+, -, *$ to denote belief expansion, contraction and revision respectively.

### 4.3 Inverse Conditionalization

Let us start with the simplest form of belief change, namely, belief expansion [2]. Expansion of a belief set occurs when new beliefs are added but no existing belief is retracted. Expansion preserves all the existing beliefs. When the sentence to be included is inconsistent with the existing beliefs, the result of expansion is a blow-up or explosion of the belief set to the set of all sentences in the language $\mathcal{L}$. However, when the sentence is
consistent with the existing beliefs, belief expansion results in a simple addition of the sentence to the existing set of beliefs along with its closure under the consequence operation. Similarly, Bayesian conditionalization by a sentence $\alpha$, which is consistent with the initial probability function, changes the probability function such that the resultant probability of $\alpha$ is 1. However, when $\alpha$ is not consistent with the probability function Bayesian conditionalization is not defined to handle this case. A trivial solution to conditionalization by $\alpha$ where $P(\alpha) = 0$ is that $P(\beta | \alpha) = 1$ for every sentence $\beta$,\(^3\) which implies that every sentence becomes a belief as a result of revising by a sentence with initial zero probability, resulting in an absurd belief state (where every sentence is a belief). Thus, in the context of probabilistic belief change, expansion is modelled by conditionalization which demonstrates similar characteristics [67].

On occasion an agent might need to remove a belief. Removal of a belief from a belief set is termed belief contraction [2, 29]. Since a belief set is closed under logical consequence, contracting a belief is not a simple issue. Contraction involves redistribution of the probabilities such that the belief being contracted will have a resultant probability less than 1. This re-distribution of the probability mass is vaguely guided by the postulates for belief contraction propounded in the AGM belief change theory [2, 29]. Equivalent postulates in a probabilistic setting, given by Gärdenfors [29], are as follows:

C1 $P^-\alpha \in \mathcal{P}$.

C2 $P^-\alpha(\alpha) = d$, where $d < 1$ iff $\not\models \alpha$.

C3 If $\vdash \alpha \leftrightarrow \beta$, then $P^-\alpha = P^-\beta$.

C4 If $P(\alpha) < 1$, then $P^-\alpha = P$.

C5 If $P(\alpha) = 1$, then $(P^-\alpha)^+ = P$.

C1 suggests that for all probability functions $P$ and all sentences $\alpha$, the result of contraction $P^-\alpha$ is also a probability function, that is, it satisfies Kolmogorov’s axioms. The postulate C2 states that as long as the belief being retracted is not a theorem ($\not\models \alpha$),

\(^3\)Such a probability function is termed a unit function. A more detailed discussion is given in [67].
contraction should be successful. Henceforth, in this work, the belief being removed is assumed not to be a theorem. Gärdenfors [29] identifies that the value of $P_\alpha^-$, that is $d$, which is not advocated to be a precise value in $C_2$, as an important parameter determining the result of contraction. $C_3$ suggests that the result of contraction is syntax independent. The postulate $C_4$ states that contraction is vacuous when the sentence being contracted is not a belief to begin with. Postulate $C_5$ is termed recovery. The recovery postulate given in [2] suggests that contraction by a belief $\alpha$ should be such that upon further expanding by $\alpha$, the agent should regain all the beliefs that were lost in contraction. However, a generalised form of the postulate suggests that the agent should be able to recover the initial belief state upon contraction followed by expansion. We call a function that satisfies the postulates $C_1$ to $C_5$ as an AGM-rational contraction function.

![Figure 4.1: Here each point denotes a possible world. The arc separates the $\alpha$-worlds from $\neg\alpha$-worlds. It is assumed that $P(\alpha) = 1$. The area enclosed by the dotted line denotes $\|P\|$, that is, the set of all worlds with non-zero probability under $P$.](image)

Figure 4.1 illustrates what the problem of probabilistic belief contraction involves. Contracting a belief $\alpha$ amounts to reducing the probability assigned to the $\alpha$-worlds by some non-zero quantity, and assigning non-zero probability to the $\neg\alpha$-worlds. In other words, the probability of some worlds in the left hand side of the arc, enclosed by the dotted line, is reduced and the probability assigned to some of worlds in the right hand side of the arc is increased. Contraction thus reduces to a problem of answering the following questions:

- **which $\alpha$-worlds have their probabilities reduced**, 

101
• by how much are these probabilities reduced, and
• which ¬\(\alpha\)-worlds gain the probability lost by the \(\alpha\)-worlds.

These questions are partially answered by the postulates of contraction. In order to satisfy
\(C_5\), we need an account of inverse conditionalization to shrink the probability associated
with all the \(\alpha\)-worlds proportionately. Thus all \(\alpha\)-worlds that have positive probability
lose some non-zero quantity of probability and their probabilities should be reduced by
the same factor. Furthermore, to satisfy \(C_1\), the sum of the probabilities over ¬\(\alpha\)-worlds
and the reduced probability of \(\alpha\) should add up to 1. The challenge in defining inverse con-
ditionalization comes in the form of determining the parameter identified by Gärdensfors
in postulate \(C_2\), that is, \(d\) and how the probability lost by \(\alpha\), that is, \(1 - d\), is re-distributed
among ¬\(\alpha\)-worlds. We will observe that the value of this parameter \(d\) will differ with
different contraction functions.

4.3.1 Contraction via Harper identity

presents an analysis of contraction in terms of revision which was translated in [29] to
the following and is now commonly known as Harper identity.

\[
\mathcal{K}_\alpha^- = \mathcal{K} \cap \mathcal{K}_{\neg\alpha}^*
\]

(Harper Identity)

The corresponding formalization of Harper identity in a probabilistic setting is as follows
[29]: when \(P(\alpha) = 1\),

\[
P_\alpha^- = PaP_{\neg\alpha}^*
\]  

(4.2)

where \(0 < a < 1\), with the understanding that \(PaQ = a \cdot P + (1 - a) \cdot Q\) for any two
probability distributions \(P, Q\). The probability distribution \(P_{\neg\alpha}^*\) in Equation 4.2 refers to
the result of revision of \(P\) by \(\neg\alpha\). Since \(P(\alpha) = 1\), we have \(P(\neg\alpha) = 0\). Revision by \(\neg\alpha\)
involves changing probability of \(\neg\alpha\) from zero to 1.
There are some alternative study of conditional probability, apart from Bayesian conditionalization, that fare well when having to update by a sentence with initial probability zero [24, 47, 78]. Among these, Hosiasson-Lindenbaum’s system of conditional probability [47] (henceforth, the HL system) has those characteristics that resemble a belief revision function [67]. When updating the given probability function by a sentence, as long as the sentence is not self-contradictory, the conditional probability, based on the HL-system, will result in a proper Kolmogorov function (that is, satisfies Kolmogorov’s axioms), irrespective of the initial probability of the sentence. Now, with belief revision being modelled in terms of HL-system, modelling belief contraction becomes straightforward. When $P^\star_{\alpha}$ results from the HL-system of conditional probability, Definition 4.2 gives a regular probability distribution $P_{\alpha}^-$ which satisfies the postulates of contraction. However, the HL-system is non-constructive and does not give a clear picture of how the revised probability distribution $P^\star_{\alpha}$ can be obtained. Hence we do not have a construction of probabilistic belief contraction based on this method.

### 4.3.2 Entropy based approach

In [74], Olsson argues that the postulates given by Gärdenfors in [29] are not sufficient to capture the idea of belief contraction. Belief contraction results in loss of information from the belief state. This may be viewed as increasing the uncertainty in the belief state. Uncertainty in a probability function is given in terms of an entropy measure. The Shannon entropy associated with a probability distribution, $P$, represented by $H(P)$, is given by the following [89]:

$$H(P) = -\sum_{\omega \in \Omega} P(\omega) \log P(\omega). \quad (4.3)$$

Olsson suggests that the contracted probability function should have a higher entropy value compared to the prior probability function [74]. Thus we have another posulate.

\[\text{In this work, we use logarithms to the base } 2.\]
Ramer provides an account of inverse conditionalization in [82] based on entropy measures. This can be summarized as follows. Let the prior probability function $P$ be such that $P(\alpha) = 1$. When required to retract the belief in a sentence $\alpha$, a new probability function $Q$ is chosen such that $Q(\alpha) < 1$. The recovery postulate $C5$ requires that $Q$ upon conditionalization by $\alpha$ reduces to the initial probability function $P$. Olsson’s postulate suggests that the entropy measure associated with $Q$ should be higher than the entropy measure associated with $P$. There will be many probability distributions that satisfy these constraints on $Q$. Among these distributions the one with highest uncertainty is chosen as the required probability function. Thus, inverse conditionalization by $\alpha$ is taken to be the probability function that has maximum possible entropy among all the functions that upon conditionalization by $\alpha$ reduces to $P$.

Since $\mathcal{L}$ is finitary, the set of possible worlds be finite. Let the cardinality of $\Omega$ be $n$, and the cardinality of the set of all $\neg\alpha$-worlds be $m$. Let $h = 2^{H(P)}$. Inverse conditionalization of probability distribution $P$ by $\alpha$, as proposed in [82], results in a probability distribution $Q$ which is as follows:

$$Q(\alpha) = \frac{h}{h+m}, \quad Q(\omega) = \frac{1}{h+m}, \text{ for any } \omega \in \neg\alpha$$

The probability function $Q$ shrinks the probability associated with each $\alpha$-world such that for any $\alpha$-world, $\omega$, only $\frac{h}{h+m}$-fraction of its initial probability is retained. In other words, for any $\alpha$-world, $\omega$, $Q(\omega) = \frac{h}{h+m}P(\omega)$. In this way $Q$ preserves the structure of $P$ within the set $[\alpha]$. Moreover, the probability mass collectively lost by the $\alpha$-worlds, that is, $\frac{m}{h+m}$, is uniformly distributed over $[-\alpha]$. Thus $Q(\alpha) < 1$, and $\alpha$ is dropped from the belief set. We term this contraction as indifferent contraction.

**Definition 4.1:** The indifferent contraction $P_\alpha^-$ of $P$ by a belief $\alpha$ is given by the distribution which has the maximum possible entropy among all the probability distributions.
that reduce to $P$ upon conditionalization by $\alpha$.

We call it indifferent contraction since it assigns equal non-zero probability to all the $\neg\alpha$-worlds, reminiscent of the principle of indifference due to Laplace.\textsuperscript{5} Let the initial probability distribution be $P$ where $\alpha$ is a belief, that is, $P(\alpha) = 1$. Let the cardinality of $[\alpha]$ be $l$ and the cardinality of $[\neg\alpha]$ be $m$ (such that $l + m = n$). We assume that the individual elements in $\Omega$ are indexed in some order, with the $l$ number of $\alpha$-worlds preceding the $m$ number of $\neg\alpha$-worlds. We represent $P$ as a list of probabilities associated with individual worlds.

$$P = \langle p_1, \ldots, p_l, 0, \ldots, 0 \rangle,$$

where $p_i$ is the probability associated with $P(\omega_i)$ for $\omega_i \in [\alpha]$. It should be noted that $p_i$ could be zero for some $1 \leq i \leq l$. The result of indifferent contraction, $P^-\alpha$, given by Equation 4.4, is as follows:

$$P^-\alpha = \langle dp_1, \ldots, dp_l, \frac{1}{h+m}, \ldots, \frac{1}{h+m} \rangle,$$

where $d = \frac{h}{h+m}$ and $h = 2^{H(P)}$.

The resultant probability of $\alpha$, $P^-\alpha(\alpha)$, is $d$ which is less than 1 and the resultant probability of $\neg\alpha$ is $1 - d$, that is, $\frac{m}{h+m}$. Also the entropy of the contracted probability distribution is $\log(h+m)$ which is greater than $\log h$ ($= H(P)$). The pre-requisite of $\alpha$ being a belief ensures this account also satisfies C4. Hence indifferent contraction satisfies postulates C1 to C5 and also CO. We show this result in the following theorem.

**Theorem 4.1:** Indifferent contraction satisfies all the postulates of contraction C1 to C5 and CO.

The proof of every result is provided in Appendix C.

\textsuperscript{5}This principle is also referred to as the principle of insufficient reason.
Gärdenfors [29] shows that if a probabilistic contraction on $P$ satisfies the postulates $C1$ to $C5$, then the corresponding contraction on the associated belief set satisfies the six basic AGM postulates of contraction [2, 29]. Moreover, a contraction function on a belief set is said to be a partial meet contraction if and only if it satisfies the basic AGM postulates of contraction [2, 29]. Thus when a probabilistic contraction on $P$ satisfies the postulates $C1$ to $C5$, the corresponding contraction on the belief set associated with $P$ is a partial meet contraction. We borrow the same terminology and refer to such a probabilistic contraction as a partial meet contraction, with the understanding that the corresponding contraction on the associated belief set is a partial meet contraction.

Two special cases of partial meet contraction are the maxichoice contraction and the full meet contraction [2, 29]. A maxichoice contraction retains a maximal subset of the initial belief set which does not entail the belief being contracted; while a full meet contraction retains only those beliefs that are common to all such maximal sets. Again we borrow the terminology and refer to a probabilistic contraction on the belief state, $P$, as a maxichoice contraction (or full meet contraction), when the corresponding contraction on the belief set associated with $P$ is a maxichoice contraction (or full meet contraction).

Since a belief set is considered to be closed under logical consequence retracting a belief from a belief set involves losing more than just the belief under question. Consider the case where $\alpha, \beta \rightarrow \alpha$ and $\beta$ all belong to the belief set $K$. When contracting $K$ by $\alpha$, it is necessary to remove either $\beta \rightarrow \alpha$ or $\beta$ or both, since $\alpha$ is deducible from the conjunction of the two. Thus there is a choice involved. Under maxichoice contraction, the agent chooses to retain either $\beta$ or $\beta \rightarrow \alpha$; while under full meet contraction, the agent loses both. However, when the contraction is partial meet, the agent makes this choice depending on its preference. If it prefers to retain $\beta$ then $\beta \rightarrow \alpha$ is lost and vice-versa. When the agent prefers both equally or is equally indifferent towards both of them, then both these beliefs are retracted. This is commonly referred to as the principle of preference and indifference [87].
The result of a full meet contraction on a belief set $K$ is such that the set of models of the contracted belief set $K_\alpha^-$ is the union of the models of $K$ and the models of $\neg \alpha$ [29]. In other words, $[K_\alpha^-] = [K] \cup [\neg \alpha]$. We observe that the indifferent contraction on $P$ has a similar effect on the associated belief set as well. Theorem 4.1 shows that an indifferent contraction is a partial meet contraction, while in fact it is a full meet contraction, as shown by the following theorem.

**Theorem 4.2:** An indifferent contraction is a full meet contraction.

As a result of indifferent contraction, the support of $P^-\alpha$ contains all $\neg \alpha$-worlds. Therefore, $\|P^-\alpha\| = \|P\| \cup [\neg \alpha]$. A sentence has probability 1 if and only if it contains all the worlds in the support. Thus the contracted set of beliefs is given by the following:

$$K_\alpha^- = \{ \beta \in K \mid [\neg \alpha] \subseteq [\beta] \}.$$

In other words, indifferent contraction only retains those beliefs that are logically deducible from $\neg \alpha$.

Our aim is to present an account of contraction that retains more of the initial set of beliefs, that is, an account of maxichoice contraction or a general account of partial meet contraction. In the following section, we proceed towards this goal and present a variation of indifferent contraction, namely, submaximal entropy contraction.

## 4.4 Submaximal Entropy Contraction

In this section we introduce another contraction, namely *submaximal entropy contraction*. This contraction function will not return a belief state that has maximum possible uncertainty as a result of contraction. Let $l$ be the cardinality of $[\alpha]$. Let the initial probability function $P$ be as follows:
\[ P = \langle p_1, \ldots, p_l, 0, \ldots, 0 \rangle. \]

The postulates of contraction \textbf{C1} to \textbf{C5} require that the probabilities \( p_1, \ldots, p_l \) be scaled by some factor, say \( d \), and the probability thus “shaved off” be distributed among the worlds in \( \neg \alpha \). That is \( Q = P^-_{\alpha} \) has the following form:

\[ Q = P^-_{\alpha} = \langle dp_1, \ldots, dp_l, q_{l+1}, \ldots, q_n \rangle, \]

where \( 0 < d < 1 \). There are many candidates that satisfy this criteria. Postulate \textbf{CO} helps in filtering out some of these possibilities.

Let us assume that the desired entropy value of \( P^-_{\alpha} \) is specified in advance to be \( t = H(P^-_{\alpha}) \). In this contraction, we look for a probability distribution that satisfies the postulates \textbf{C1} to \textbf{C5} and \textbf{CO} and have the specified entropy value \( t \). When the value of \( t \) is \( \log(h + m) \), where \( h = 2^{H(P)} \), there exists a unique distribution that satisfies these constraints and it is given by Equation 4.4. Contraction, in this case, is same as indifferent contraction. Clearly, \( t \) cannot be greater than \( \log(h + m) \) and to satisfy the postulate \textbf{CO} we need \( t \geq H(P) \). Therefore, we define a specified entropy value as \textit{acceptable} if it is greater than or equal to the entropy value of the initial distribution, that is, \( H(P) \), and less than the maximum entropy value that is possible as a result of contraction, that is, \( \log(h + m) \). In other words, \( t \) is acceptable if and only if \( H(P) \leq t < \log(h + m) \). Since we specify the entropy of the contracted state to be less than the maximum possible value, we term this contraction as \textit{submaximal entropy contraction}.

There are numerous probability distributions which have the specified entropy value \( t \). In order to choose one particular distribution among the many possible candidates, we use the principle of minimization of the relative entropy [90]. The relative entropy is also referred to as \textit{cross entropy}, \textit{divergence}, \textit{discrimination information}, \textit{Kullback-Leibler number}, and \textit{directed divergence} [31]. Given a probability distribution \( Q \), the divergence of another distribution \( P \) from \( Q \), denoted by \( D(P\|Q) \), is defined as follows:

\[
D(P\|Q) = \sum_{\omega \in \Omega} P(\omega) \log \frac{P(\omega)}{Q(\omega)}.
\]
We define the submaximal entropy contraction as follows:

**Definition 4.2:** Let $t$ be the desired entropy value of the resultant belief state when $P$ is contracted by $\alpha$. The contraction $P_\alpha^-$ is said to be a submaximal entropy contraction when (a) $P_\alpha^-$ reduces to $P$ upon conditionalization by $\alpha$, (b) $H(P_\alpha^-) = t$, and (c) $D(P \parallel P_\alpha^-) \leq D(P \parallel Q)$, for any distribution $Q$ that satisfies (a) and (b).

It should be noted that we use $D(P \parallel P_\alpha^-)$ instead of $D(P_\alpha^- \parallel P)$. The function $D(P \parallel Q)$ is defined only if $Q(\omega) > 0$ when $P(\omega) > 0$ for every $\omega \in \Omega$. When we compare $P_\alpha^-$ and $P$, there exists some $\neg \alpha$-world which has non-zero probability under $P_\alpha^-$. Since $\alpha$ is a belief, we have $P(\alpha) = 1$ and therefore, every $\neg \alpha$-world has zero probability under $P$. Thus we have some $\neg \alpha$-world which has non-zero probability wrt $P_\alpha^-$ and zero probability wrt $P$. Hence $D(P_\alpha^- \parallel P)$ is not defined.

It can be easily shown that submaximal entropy contraction satisfies the postulates of contraction C1 to C5. This optimization problem, that is, submaximal entropy contraction which involves minimizing $D(P \parallel P_\alpha^-)$, results in a non-linear equation on $d$, the factor by which the probability associated with the $\alpha$-worlds is scaled. Furthermore, as a result of submaximal entropy contraction, the mass $1 - d$ is uniformly distributed over the $\neg \alpha$-worlds. Therefore all the $\neg \alpha$-worlds have non-zero probability. Since the support of the contracted distribution, $\parallel P_\alpha^-$, still contains all the models of $\neg \alpha$, submaximal entropy contraction is also a full meet contraction, that is, the corresponding contraction on the belief set associated with $P$ is a full meet contraction.

**Theorem 4.3:** A submaximal entropy contraction is a full meet contraction.

Next, we present a somewhat special case of submaximal entropy contraction. When the specified value of entropy, $t$, is considered to be less than the entropy of the initial probability distribution, we get interesting results.
It is clear that the value of $t$ guides submaximal entropy contraction. Consider the case where the value of $t$ is pushed down as close as possible to 0. First we note that, in such a case, we cannot expect the contraction to satisfy the postulate CO. The entropy of a probability distribution $Q$ is 0 if and only if $Q(\omega) = 1$ for some $\omega \in \Omega$, and 0 for other worlds. If we specify the entropy of the contracted probability distribution to be 0, then $P_\alpha^−$ assigns 0 to every world in $\Omega$ except one and thus does not preserve the initial proportions among the $\alpha$-worlds. Conditionalizing $P_\alpha^−$ by $\alpha$ may not result in $P$ and hence C5 is not satisfiable. Hence, while an exact value 0 is not possible to achieve, we can come arbitrarily close to 0.

Since we require that the contracted probability distribution reduce to $P$ upon conditionalization by $\alpha$, the entropy of such a probability distribution is greater than $dH(P)$, where $d$ is the factor by which the probability of $\alpha$-worlds is scaled. If we were to make $d$ larger, then the quantity $dH(P)$ becomes greater and hence the entropy of the resultant probability distribution becomes high. On the other hand, if we were to make the value $d$ very small, say close to 0, the quantity $dH(P)$ would be very small as well. The contribution of the $\alpha$-worlds towards the entropy of the resultant probability distribution would be small. However, when the probability mass $1 - d$, which is close to 1, is uniformly distributed over $\neg\alpha$-worlds, the contribution of $\neg\alpha$-worlds towards the entropy of the resultant distribution would be close to $\log m$, where $m$ is the cardinality of the set $[\neg\alpha]$. Thus to get a very small entropy value for the contracted probability distribution, we need to set $d$ very close to 0 and the probability mass $1 - d$ should be concentrated on a single arbitrary $\neg\alpha$-world. We term such a contraction as minimal contraction.

**Definition 4.3:** A contraction $P_\alpha^−$ of $P$ by a belief $\alpha$ is a minimal contraction when $P_\alpha^−$ reduces to $P$ upon conditionalization by $\alpha$ and has the minimal non-zero entropy.

6The $\alpha$-worlds would contribute a value of $dH(P) - d \log d$ towards the entropy of the resultant probability distribution. Since $\log d$ is negative ($d < 1$), $dH(P) - d \log d$ is greater than $dH(P)$. 

110
Minimal contraction results in a family of contracted probability functions depending on which model of \( \neg \alpha \) is assigned positive probability. We have already observed that minimal contraction does not satisfy Olsson’s postulate, \( \text{CO} \). It does not satisfy the postulate \( \text{C3} \) as well. Definition 4.3 does not specify which \( \neg \alpha \)-world receives the non-zero probability. Since when contracting \( \alpha \), the \( \neg \alpha \)-world which is assigned non-zero probability is chosen arbitrarily, it is not necessary that the same world is chosen upon contraction by \( \beta \), even when the two sets \( \neg \alpha \) and \( \neg \beta \) are the same. Minimal contraction will satisfy \( \text{C3} \) only when there is some structure to how a \( \neg \alpha \)-world is chosen and assigned positive probability. However, we do not investigate this issue any further here. Regardless of its failure to satisfy \( \text{CO} \) and \( \text{C3} \), minimal contraction satisfies the other postulates. We are, therefore, tempted to call it a “syntax-sensitive” partial meet contraction.

The support of the probability distribution resulting from a minimal contraction is such that \( \| P_\alpha^- \| = \| P \| \cup \{ \omega \} \), where \( \omega \) is an arbitrary \( \neg \alpha \)-world. Hence, it changes the support of the initial probability distribution minimally. Therefore, it is clear that minimal contraction retains a maximal subset of the initial set of beliefs. In [2, 29] such contraction is termed as maxichoice contraction. The following result highlights the relation between a minimal contraction and a maxichoice contraction.

**Theorem 4.4:** A minimal contraction is a “syntax-sensitive” maxichoice contraction.

The result of minimal contraction is such that \( P_\alpha^- (\alpha) = d \), where \( d \) is very small, close to 0 and \( P_\alpha^- (\neg \alpha) = 1 - d \) is very close to 1. Minimal contraction results in a big change with respect to the probabilities assigned to individual worlds; while making minimal change to the associated set of beliefs.

Till now, we have discussed two cases of submaximal entropy contraction:

1. the general case which results in a full meet contraction, and
2. the special case of minimal contraction which results in a “syntax-sensitive” maxi-
choice contraction.

We investigate this notion further to present an account of partial meet contraction.

A generalized account of contraction would be one where an arbitrary subset of \([-\alpha]\) is assigned non-zero probability as a result of contraction by \(\alpha\). We begin by noting that indifferent contraction and submaximal entropy contraction (given by Definition 4.2) assign positive probability to every world in \([-\alpha]\); while minimal contraction assigns positive probability to one arbitrary model of \(-\alpha\).

Recalling Equation 4.4, the result of indifferent contraction \(P^-\) of \(P\) by a belief \(\alpha\) is such that:

\[
P^-(\alpha) = \frac{h}{h + m}; \quad P^-(\omega) = \frac{1}{h + m}, \quad \omega \in [-\alpha],
\]

where \(h = 2^{H(P)}\) and \(m\) is the cardinality of \([-\alpha]\). The probability distribution \(Q = P^-\) has the maximum possible entropy value among all the distributions that reduce to \(P\) upon conditionalization by \(\alpha\), with \(H(Q) = \log(2^{H(P)} + m)\) or \(2^H(Q) = 2^{H(P)} + m\). By replacing \(m\), in Equation 4.4, with \(m'\), where \(m' \leq m\), we get a distribution \(Q'\):

\[
Q' = \left\langle \frac{h}{h + m'} p_1, \ldots, \frac{h}{h + m'} p_l, \frac{1}{h + m'}, \ldots, \frac{1}{h + m'}, 0, \ldots, 0 \right\rangle.
\]

Let \(\Delta\) be an arbitrary subset of \([-\alpha]\) whose cardinality is \(m'\). The distribution \(Q'\) is such that it reduces to \(P\) upon conditionalization by \(\alpha\) (since probability associated with every \(\alpha\)-world is scaled by the factor \(\frac{h}{h + m'}\)) and has the maximum possible entropy value among all the distributions over \([\alpha] \cup \Delta\). Let us assume that \(m' = 1\). We understand that \(Q'\) has the maximum possible entropy value over the set \([\alpha] \cup \Delta\), for any \(\Delta \subseteq [-\alpha]\) with cardinality 1. Suppose the specified entropy value \(t\) is such that \(t > H(Q') = \log(2^{H(P)} + 1)\). For any distribution \(Q''\) to have entropy value \(t\) and reduce to \(P\) upon conditionalization, it is clear that \(Q''\) should assign non-zero probability to more than one \(-\alpha\)-world. Thus when the specified entropy value \(t\) is such that \(t > \log(2^{H(P)} + k)\), for \(0 \leq k < m\), it is clear that the probability distribution with the specified entropy value and satisfying the postulate \(C5\) should assign positive probability to some subset of
whose cardinality is greater than $k$. In this way, by specifying the desired entropy value for the contracted probability distribution we can dictate the minimum number of $\neg \alpha$-worlds that receive non-zero probability.

While we can specify the minimum number of worlds in $[\neg \alpha]$ that receive non-zero probability, we are unable to specify the maximum number of $\neg \alpha$-worlds that receive positive probability. Therefore, submaximal entropy contraction could still result in an uniform distribution over $[\neg \alpha]$. Moreover, finding the values of $t$ where the solution changes its phase (includes a different number of $\neg \alpha$-worlds) can be done only numerically. These numerical constraints appear to have no logical meaning; they come from a system of nonlinear equations.

A submaximal entropy contraction may not have given a full meet contraction had we chosen $P_\alpha^-$ which does not assign positive probability to every $\neg \alpha$-world. But such a choice can only be based on further information, when available. If the agent had certain information that determines which of the $\neg \alpha$-worlds should have non-zero probability as a result of contraction, then a choice of $P_\alpha^-$ can be made consistent with this information, instead of choosing an uniform distribution over $[\neg \alpha]$. Thus to avoid a full meet contraction as a result of submaximal entropy contraction, some extra information or constraint is required. Moreover, we believe this extra information cannot be of numerical nature.

As mentioned earlier, partial meet contraction is based on the principle of preference and indifference. Thus, in order to give an account of the same, we need the belief state to be able to represent a preference relation among the agent’s beliefs. In the next section, we use a system of spheres representation [32] of a preference relation over the worlds (indirectly over the beliefs) to give an account of partial meet contraction.

### 4.5 Partial meet contraction

It is clear that in order to present an account of partial meet contraction, we need a mechanism that represents the agent’s preferences over beliefs. But our representation of belief
state is not capable of representing the agent’s preferences in its current form, since all the beliefs have probability 1. We need to incorporate a mechanism for representing preferences among the beliefs in our current representation of the belief state. We begin by recalling some of the representations of belief state, discussed in the literature, which give a preference relation among the agent’s beliefs.

Two different representations of preferences of the agent are well known in the literature, an *epistemic entrenchment* representation [29] and a *system of spheres* representation [32]. Accounts of partial meet contraction have been studied based on both these representations [29]. In these two cases the preferences of the agent are explicitly specified in the form of certain relations. Epistemic entrenchment gives a relation \( \preceq \) among the sentences of the language depicting the agent’s preference. In this representation \( \alpha \preceq \beta \) is understood as \( \beta \) being at least as preferred (or epistemically entrenched) as \( \alpha \) and hence when given a choice between losing \( \alpha \) and \( \beta \), the agent is expected to give up \( \alpha \).

In the system of spheres representation [32], the belief state is given by a total preorder relation \( \sqsubseteq \) over the set of possible worlds \( \Omega \).\(^7\) The relation \( \omega \sqsubseteq \nu \) is read as: \( \omega \) is at least as plausible as a model of real world as \( \nu \). The strict and the symmetric parts of the relation \( \sqsubseteq \) are given by \( \sqsubset \) and \( \approx \) respectively. Since we deal with a probability distribution over \( \Omega \), we will use a total preorder relation over \( \Omega \) to represent the agent’s preferences over its beliefs. We briefly recall some important aspects of this representation that will be necessary for our work.

Given any set of worlds \( \Delta \subseteq \Omega \), by \( \text{min}_{\sqsubseteq}(\Delta) \) we will denote the set of \( \sqsubseteq \)-minimal worlds of \( \Delta \) that the epistemic agent considers most plausible among those in \( \Delta \). In other words, \( \text{min}_{\sqsubseteq}(\Delta) = \{ \omega \in \Delta \mid \omega \sqsubseteq \omega', \text{ for all } \omega' \in \Delta \} \). In particular, \( \text{min}_{\sqsubseteq}(\Omega) \) will represent the set of most plausible worlds among all possible worlds as viewed by the agent. The beliefs of the agent are those that are true in all these most plausible worlds.

\(^7\) A total preorder relation is a connected, reflexive and transitive relation. A total preorder relation over \( \Omega \) can be pictorially represented as a system of concentric spheres, where the innermost sphere houses the minimal worlds according to the total preorder relation and the enclosing spheres are built up hierarchically based on the relation [32].
Hence the set of beliefs extracted from a total preorder \( \sqsubseteq \) over \( \Omega \), denoted by \( \mathcal{K}_{\sqsubseteq} \), is

\[
\mathcal{K}_{\sqsubseteq} = \{ \alpha \in \mathcal{L} \mid \min_{\mathcal{E}}(\Omega) \subseteq [\alpha] \}.
\] (G1)

In fact, the relation \( \sqsubseteq \) indirectly gives an ordering among the beliefs of the agent. A belief \( \alpha \) is said to be at least as preferred as the belief \( \beta \) when \( \omega \sqsubseteq \nu \) where \( \omega \) is a \( \min_{\mathcal{E}}[\neg \beta] \)-world and \( \nu \) is a \( \min_{\mathcal{E}}[\neg \alpha] \)-world.

Let the belief state of the agent, represented by a probability distribution \( P \), be appended with a total preorder relation \( \sqsubseteq \) over \( \Omega \). By comparing Equations 4.1 and G1, we understand that \( \min_{\mathcal{E}}(\Omega) \) are the worlds with positive probability under \( P \), that is, \( \| P \| = \min_{\mathcal{E}}(\Omega) \). We add the following constraint to the definition of indifferent contraction (Definition 4.1): *every \( \neg \alpha \)-world which is not in \( \min_{\mathcal{E}}[\neg \alpha] \) should have zero probability after contraction*. In other words, as a result of contraction, among the \( \neg \alpha \)-worlds only the worlds in \( \min_{\mathcal{E}}[\neg \alpha] \) will receive non-zero probability, that is, \( \| P_\alpha^{-} \| = \| P \| \cup \min_{\mathcal{E}}[\neg \alpha] \).

**Definition 4.4:** Given a total preorder relation \( \sqsubseteq \) on \( \Omega \), the preferential contraction \( P_\alpha^{-} \) of \( P \) by a belief \( \alpha \) is given by a probability distribution which has the maximum possible entropy value among all the distributions over the space \( [\alpha] \cup \min_{\mathcal{E}}[\neg \alpha] \) that reduce to \( P \) upon conditionalization by \( \alpha \).

This contraction is termed preferential contraction since the worlds in \( \min_{\mathcal{E}}[\neg \alpha] \) are given more preference than the rest of the \( \neg \alpha \)-worlds. Let the cardinality of \( [\neg \alpha] \) and \( \min_{\mathcal{E}}[\neg \alpha] \) be \( m \) and \( m' \) respectively. Preferential contraction of \( P \) by a belief \( \alpha \) is given by the following:

\[
P_\alpha^{-}(\alpha) = \frac{h}{h + m'}, \quad P_\alpha^{-}(\omega) = \frac{1}{h + m'}, \quad \omega \in \min_{\mathcal{E}}[\neg \alpha]
\] (4.5)

Let \( P \) be given as follows:

115
\[ P = \langle p_1, \ldots, p_t, 0, \ldots, 0 \rangle, \]

then the probability distribution resulting from preferential contraction is as follows:

\[ P^- = \langle dp_1, \ldots, dp_t, \frac{1}{h + m'}, \ldots, \frac{1}{h + m'}, 0, \ldots, 0 \rangle, \]

where \( d = \frac{h}{h + m'} \) and \( h = 2^H(P) \). Preferential contraction expands the distribution \( P \) over an arbitrary subset of \( \lnot \alpha \). Thus preferential contraction, employing a mixture of logic based constraints and numerical constraints, gives an account of partial meet contraction.

**Theorem 4.5:** A preferential contraction function is a partial meet contraction function.

Two special cases of preferential contraction need to be mentioned.

**Case 1:** Suppose the system of spheres is such that \( \min_{\sqsubseteq} \lnot \alpha = \lnot \alpha \). In this case, every model of \( \lnot \alpha \) receives non-zero probability as a result of contraction. Moreover, it is clear that Equation 4.5 reduces to Equation 4.4. Thus, in this case, the result of preferential contraction is same as that of indifferent contraction. This is not particularly surprising, since \( \min_{\sqsubseteq} \lnot \alpha = \lnot \alpha \) states that all the \( \lnot \alpha \)-worlds are equally plausible. When all the worlds are equally preferred and the agent is indifferent between them, preferential contraction gives a uniform distribution over \( \lnot \alpha \). Thus we have a full meet contraction.

**Case 2:** Suppose the system of spheres is such that \( \min_{\sqsubseteq} \lnot \alpha = \{ \omega \} \), for an arbitrary \( \omega \in \lnot \alpha \). In this case, only a single model of \( \lnot \alpha \) is picked to have non-zero probability as a result of contraction. By definition, preferential contraction results in a distribution with maximum possible entropy value over the space \( \|P\| \cup \{ \omega \} \); while minimal contraction, as given by Definition 4.3, results in a distribution with minimum possible (non-zero) entropy value over the same space. Minimal contraction assigns positive probability to an arbitrary single \( \lnot \alpha \)-world as a result of contraction. A total preorder relation could be used to choose this single \( \lnot \alpha \)-world.
The accounts of indifferent contraction (Section 4.3), submaximal entropy contraction, minimal contraction (Section 4.4) and preferential contraction (Section 4.5) are all based on the Shannon entropy measure. In the following section, we study the counterparts of these contraction functions based on Hartley entropy [79]. We observe that an account of partial meet contraction follows easily when we use the Hartley entropy.

4.6 A note on Hartley entropy based approaches

Hartley entropy measure [79], denoted by $H_0(P)$, is given by the following:

$$H_0(P) = \log \|P\|.$$  

Hartley’s measure only takes into account the support of the probability distribution. The characteristics of Hartley entropy are studied in [1] among others.

4.6.1 Indifferent contraction

Let $n$ be the cardinality of the set of all possible worlds $\Omega$. The maximum value that the Hartley entropy can take for any probability distribution over $\Omega$ is, evidently, $\log n$. Let $P$ be the probability distribution representing the agent’s belief state and $\alpha$ be a belief. Indifferent contraction based on Shannon entropy (Definition 4.1) results in a probability distribution that has the maximum possible entropy value among all distributions that reduce to $P$ upon conditionalization by $\alpha$. It results in an uniform distribution over the set $[\neg\alpha]$.

Similarly indifferent contraction, based on Hartley entropy, is expected to give a probability distribution which has maximum possible (Hartley) entropy and reduce to $P$ upon conditionalization by $\alpha$. Let $l'$ be the cardinality of the support of $P$. Then the Hartley entropy of the distribution $P$ is $\log l'$. The postulate C5 ensures that the $\alpha$-worlds which had zero initial probability will continue to have zero probability. Therefore the maximum possible entropy value for any contracted probability distribution is $\log(l' + m)$, where $m$
is the cardinality of $[\neg \alpha]$. Thus indifferent contraction is expected to give a probability distribution which assigns non-zero probability to every $\neg \alpha$-world. Let $P$ be given as follows:

$$P = (p_1, \ldots, p_l, 0, \ldots, 0),$$

where $l$ is the cardinality of $[\alpha]$. The result of indifferent contraction based on Hartley entropy is as follows:

$$Q = P^- = (dp_1, \ldots, dp_l, q_{l+1}, \ldots, q_{l+m}),$$

where $0 < d < 1$. It is evident that indifferent contraction satisfies the postulates $C1$ to $C5$. Moreover, we have $||P^-|| = ||P|| \cup [\alpha]$. Thus indifferent contraction based on Hartley entropy (henceforth, we write, indifferent Hartley contraction) is also a full meet contraction, similar to indifferent contraction based on Shannon entropy (henceforth, in short, indifferent Shannon contraction).

Indifferent Hartley contraction does not provide an unique contracted probability distribution. While indifferent Shannon contraction specifies the value of parameter $d$ to be $\frac{h}{h+m}$; indifferent Hartley contraction does not specify a particular value for $d$, except for the requirement that $0 < d < 1$. Moreover, how the probability mass $1 - d$ is distributed over the $\neg \alpha$-worlds is not specified as well. Without loss of generality, we can expect the contracted probability distribution to be uniform over $[\neg \alpha]$. Thus indifferent Hartley contraction results in,

$$P^-_\alpha(\alpha) = d; \quad P^-_\alpha(\omega) = \frac{1 - d}{m}, \quad \omega \in [\neg \alpha].$$

Indifferent Hartley contraction gives a family of contracted probability distributions depending on the value of $d$. The real number $d$ represents the probability of $\alpha$ in the contracted state and could be coupled with the information received by the agent which demands removal of the belief $\alpha$. 

118
4.6.2 Submaximal entropy contraction

Let $P$ represent the initial belief state and $\alpha$ be the belief being contracted. Submaximal entropy contraction based on Shannon entropy (Definition 4.2) can be summarised as follows: Let the desired (acceptable) entropy value of the contracted distribution be specified as $t$. The contracted probability distribution is chosen, from a number of distributions which have the desired entropy $t$ and reduce to the initial distribution upon conditionalization by $\alpha$, based on principle of minimum cross entropy. Such a contraction is uniform over $[\neg\alpha]$ and preserves the initial proportions among the worlds in $[\alpha]$.

Let the entropy value of the contracted state be specified explicitly, $t = H_0(P^-)$. Since Hartley entropy takes discrete values such as $\log k$, $1 \leq k \leq n$, we can either assume $t$ to be one of these discrete values or have a mechanism by which we choose a discrete entropy value $\log k$ that is closest to $t$. Without any loss of generality, we assume that $t = \log k$, for some $1 \leq k \leq n$. An acceptable Hartley entropy value $t$ would be such that $\log l' \leq t \leq \log (l' + m)$, where $l'$ is the cardinality of the support of $P$ and $m$ is the cardinality of $[\neg\alpha]$. The upper limit is evident, since $\log (l' + m)$ is the maximum possible Hartley entropy for the contracted state. For contraction to be successful, it is understood that at least one of the $\neg\alpha$-worlds should have non-zero probability. Furthermore, the restriction that $P^-_\alpha$ should reduce to $P$ upon conditionalization by $\alpha$ ensures that all the $\alpha$-worlds that had non-zero initial probability retain some non-zero probability after contraction. Thus the cardinality of the support of $P^-_\alpha$ should be greater than that of $P$ and hence $H_0(P^-_\alpha) > H_0(P)$, which forms the lower limit for an acceptable entropy value, $t$. It should be noted that, unlike the case where Shannon entropy is used (especially, the minimal contraction), it is not possible for the Hartley entropy value of the contracted distribution to be less than that of the initial distribution. This would imply that the cardinality of the support of $P^-_\alpha$ is less than the cardinality of the support of $P$, hence violating C5.

Let the initial distribution be as follows:
\[ P = \langle p_1, \ldots, p_l, 0, \ldots, 0 \rangle, \]

where \( l \) is the cardinality of \( [\alpha] \). It should be noted that some of the \( \alpha \)-worlds could have zero probability under \( P \). When an acceptable entropy value of \( P^-_\alpha \) is specified as \( t = \log k \), submaximal entropy contraction returns a probability distribution \( P^-_\alpha \) as follows:

\[ P^-_\alpha = \langle dp_1, \ldots, dp_l, \underbrace{q_{l+1}, \ldots, q_{l+m'}, 0, \ldots, 0}_{m'} \rangle, \]

where \( 0 < d < 1 \) and \( m' = k - l' \) and \( l' \) is the cardinality of the support of \( P \).

We discuss some of the drawbacks of a submaximal entropy contraction based on Hartley entropy. The support of \( P^-_\alpha \) is such that \( \| P^-_\alpha \| = \| P \| \cup \Delta \), where \( \Delta \) is an arbitrary subset of \( [\neg \alpha] \) such that cardinality of \( \Delta \) is \( m' \). Even logic based constraints, such as described in Section 4.5 with the aid of a system of spheres, may not be helpful in this case, since the cardinality of \( \Delta \) is explicitly mentioned. Also, the value of \( d \) is not specified by the submaximal entropy contraction. Thus it gives a family of contractions depending on the value of \( d \). Moreover, by the principle of indifference, we can assume without loss of generality, that the probability mass \( 1 - d \) is uniformly distributed over the set \( \Delta \).

It is clear that, a submaximal entropy contraction based on Hartley entropy satisfies the postulates \( C1, C2, C4 \) and \( C5 \). Since, as mentioned above, the subset of \( [\neg \alpha] \) is arbitrarily chosen, unless some more structure is forced, this contraction will not satisfy \( C3 \). Thus it is a “syntax-sensitive” contraction function, similar to the minimal contraction defined in Section 4.4. Hence a submaximal entropy contraction based on Hartley entropy is a “syntax-sensitive” partial meet contraction.

### 4.7 Concluding Remarks

Study of probabilistic contraction has been relatively neglected. An elegant account of contraction is given in [82], which we term as *indifferent contraction*. This account is
based on the principle of maximum entropy [14, 90]. It does not make a lot of assumptions and results in a set of beliefs which is much smaller than the initial set of beliefs. In this paper, we show that indifferent contraction corresponds to a full meet contraction.

Our main goal in this paper has been to minimize the loss of beliefs as a result of contraction, by giving a generalized account of partial meet contraction. Towards this end, we proposed the submaximal entropy contraction. It turns out that a special case of submaximal entropy contraction is the minimal contraction. Minimal contraction retains a maximal subset of the initial set of beliefs as a result of contraction, that is, it is a probabilistic counterpart of maxichoice contraction. We also gave an account of partial meet contraction, namely, preferential contraction, with the aid of a belief state which is a combination of a probability distribution over $\Omega$ and a total preorder relation over $\Omega$. We do not here discuss the result of the contraction on the total preorder relation. We find such a discussion to be outside the scope of this work and retain that for future investigations.

The accounts described above are based on Shannon entropy. Subsequently, we gave counterparts of indifferent contraction and submaximal entropy contraction based on Hartley entropy. We found that a probabilistic account of partial meet contraction is fairly easily obtained based on Hartley entropy, albeit being “syntax-sensitive”.

Unfortunately, in both submaximal entropy contraction and minimal contraction, we fail to find a motivation to specify the entropy value or lower the entropy value to close to zero. Intuitively acceptable restrictions that lead to an economical probabilistic contractions remain an open problem.
Appendix A

Proofs for Chapter 2

Lemma 2.1  Let $\sqsubseteq$ be a consistent belief state and $\mathcal{K}$ its associated belief set. An AGM-rational state contraction function $\mathcal{C}$ satisfies $\mathcal{O}_\mathcal{C}[\alpha]$ for any sentence $\alpha$ iff for every sentence $\beta$ such that $\vdash \alpha \lor \beta$, $(\mathcal{K}_\alpha^-)_{\beta} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-$.  

Proof: Let $\mathcal{C}$ be a contraction function satisfying Eqn. G2, $\sqsubseteq$ a belief state and $\mathcal{K}$ given by Eqn. G1. Let $\alpha$ be any arbitrary sentence. We need to show that the contraction function $\mathcal{C}$ satisfies $\mathcal{O}_\mathcal{C}[\alpha]$ iff $(\mathcal{K}_\alpha^-)_{\beta} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-$ for every sentence $\beta$ such that $\vdash \alpha \lor \beta$.  

(Left to Right). Assume that $\mathcal{C}$ satisfies $\mathcal{O}_\mathcal{C}[\alpha]$. Let $\beta$ be an arbitrary sentence such that $\vdash \alpha \lor \beta$. Since $\vdash \alpha \lor \beta$ we have $[\neg \beta] \subseteq [\alpha]$. Therefore $min_{\mathcal{E}_\alpha} [\neg \beta] = min_{\mathcal{E}_\alpha} [\alpha \land \neg \beta]$. Since $\mathcal{C}$ satisfies $\mathcal{O}_\mathcal{C}[\alpha]$, we further have $min_{\mathcal{E}_\alpha} [\alpha \land \neg \beta] = min_{\mathcal{E}_\alpha} [\alpha \land \neg \beta]$. Therefore $min_{\mathcal{E}}(\Omega) \cup min_{\mathcal{E}} [\neg \alpha] \cup min_{\mathcal{E}_\alpha} [\neg \beta] = min_{\mathcal{E}}(\Omega) \cup min_{\mathcal{E}} [\neg \alpha] \cup min_{\mathcal{E}} [\alpha \land \neg \beta]$, that is, $(\mathcal{K}_\alpha^-)_{\beta} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-$.  

(Right to Left). Assume that $(\mathcal{K}_\alpha^-)_{\beta} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-$ for every sentence $\beta$ such that $\vdash \alpha \lor \beta$. Let $\omega, \omega'$ be two arbitrarily chosen worlds in $[\alpha]$. Let us assume that $\omega \subseteq \omega'$. There are two cases: (1) Suppose $\omega \in min_{\mathcal{E}}(\Omega)$. Then from Eqn. G2 we can deduce that $\omega \in min_{\mathcal{E}_\alpha}(\Omega)$, that is, $\omega \subseteq^{\mathcal{C}} \omega'$. (2) Suppose $\omega \notin min_{\mathcal{E}}(\Omega)$. Since $\mathcal{L}$ is finitely generated, there exists a sentence $\beta$ such that $[\neg \beta] = \{\omega, \omega'\}$. Since $[\neg \beta] \subseteq [\alpha]$ we have $\vdash \alpha \lor \beta$. Therefore $(\mathcal{K}_\alpha^-)_{\beta} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta}^-$, whereby, from Eqn. G2
\[
\min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq_a} [-\beta] = \min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq} [\alpha \land \neg \beta].
\]

However \( \omega \notin \min_{\subseteq} (\Omega) \), \( \omega \notin \min_{\subseteq} [-\alpha] \) and \( \omega \in \min_{\subseteq} [\alpha \land \neg \beta] \). Therefore \( \omega \in \min_{\subseteq_a} [-\beta] \), that is, \( \omega \subseteq_{a} \omega' \). Similarly assuming \( \omega \sqsubseteq \omega' \) yields \( \omega \subseteq_{a} \omega' \). Hence \( \omega \) satisfies \( \mathcal{OP}_{\alpha} [\alpha] \).

\[\square\]

**Lemma 2.2** Let \( \subseteq \) be a consistent belief state and \( \mathcal{K} \) its associated belief set. An AGM-rational state contraction function \( \mathcal{OP}_{\alpha} \) satisfies \( \mathcal{OP}_{\alpha} [-\alpha] \) for any sentence \( \alpha \) iff for every sentence \( \beta \) such that \( \vdash \alpha \rightarrow \beta \), \( (K_{\alpha}^{-})_{\beta} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \).

**Proof:** This proof is similar to that of Lemma 2.1. Let \( \vdash \) be a contraction function satisfying Eqn. G2, \( \subseteq \) a belief state and \( \mathcal{K} \) the corresponding belief set. Let \( \alpha \) be an arbitrary sentence. We need to show that \( \vdash \mathcal{OP}_{\alpha} [-\alpha] \) iff for every sentence \( \beta \) such that \( \vdash \alpha \rightarrow \beta \), \( (K_{\alpha}^{-})_{\beta} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \).

(Left to Right). Let \( \vdash \mathcal{OP}_{\alpha} [-\alpha] \). Let \( \beta \) be a sentence such that \( \vdash \alpha \rightarrow \beta \). Since \( \vdash \alpha \rightarrow \beta \) we have \([\alpha] \subseteq [\beta] \). Therefore \( \min_{\subseteq_a} [-\beta] \subseteq \min_{\subseteq_a} [-\beta \land \neg \alpha] \). Since \( \vdash \mathcal{OP}_{\alpha} [-\alpha] \) we further have \( \min_{\subseteq_a} [-\beta \land \neg \alpha] = \min_{\subseteq} [-\beta \land \neg \alpha] \). Therefore \( \min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq_a} [-\beta] = \min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq} [-\alpha \land \neg \beta] \), that is, \( (K_{\alpha}^{-})_{\beta} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \).

(Right to Left). Assume that \( (K_{\alpha}^{-})_{\beta} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \) for every \( \beta \) such that \( \vdash \alpha \rightarrow \beta \).

Let \( \omega, \omega' \) be two worlds in \([-\alpha] \). Assume further that \( \omega \subseteq \omega' \). Two cases arise here:

1. Suppose \( \omega \in \min_{\subseteq} (\Omega) \). Then \( \omega \in \min_{\subseteq_a} (\Omega) \), that is, \( \omega \subseteq_{a} \omega' \).
2. On the other hand, suppose \( \omega \notin \min_{\subseteq} (\Omega) \). Since \( \mathcal{L} \) is finitely generated there exists a sentence \( \beta \) such that \( [-\beta] = \{\omega, \omega'\} \). Since \( [-\beta] \subseteq [-\alpha] \) we have \( \vdash \alpha \rightarrow \beta \). Therefore \( (K_{\alpha}^{-})_{\beta} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \) which, via Eqn. G2, gives

\[
\min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq_a} [-\beta] = \min_{\subseteq} (\Omega) \cup \min_{\subseteq} [-\alpha] \cup \min_{\subseteq} [-\alpha \land \neg \beta].
\]

We have \( \omega \in \min_{\subseteq} [-\alpha \land \neg \beta] \) but \( \omega \notin \min_{\subseteq} (\Omega) \). Now, if \( \omega \notin \min_{\subseteq} [-\alpha] \) then from Eqn. G2 we have \( \omega \in \min_{\subseteq_a} (\Omega) \), that is, \( \omega \subseteq_{a} \omega' \). On the other hand, if \( \omega \notin \min_{\subseteq} [-\alpha] \), then \( \omega \in \min_{\subseteq_a} [-\beta] \), that is, \( \omega \subseteq_{a} \omega' \). Similarly assuming \( \omega \subseteq \omega' \) yields \( \omega \subseteq_{a} \omega' \). Hence –
satisfies $\mathcal{OP}_\alpha[\neg \alpha]$.

\textbf{Theorem 2.1} Every contraction function – satisfying Eqn. G2, $\mathcal{OP}_\alpha[\alpha]$ and $\mathcal{OP}_\alpha[\neg \alpha]$ satisfies the principled factored insertion (PFI).

\textbf{Proof:} Let $-\alpha$ be a contraction function that satisfies Eqn. G2, $\mathcal{OP}_\alpha[\alpha]$ and $\mathcal{OP}_\alpha[\neg \alpha]$. Let $\mathcal{K}$ be the belief set obtained from the belief state $\square$ by Eqn. G1. We need to show that for any arbitrary beliefs $\alpha$ and $\beta$ in $\mathcal{K}$ such that $\beta \in \mathcal{K}_\alpha$, $-\alpha$ satisfies PFI.

Suppose that $\alpha \lor \beta \in (\mathcal{K}_\alpha)_\beta^\neg$. Eqn. G2 gives that $min_{\mathcal{E}_\alpha} (-\alpha) = min_{\mathcal{E}_\alpha} (\Omega) \cup min_{\mathcal{E}_\alpha} [\neg \alpha]$.

Now applying Eqn. G2 again (for contraction by $\beta$) we get
\[ min_{(\mathcal{E}_{\alpha \beta})} (-\alpha) = min_{\mathcal{E}_\alpha} (-\alpha) \cup min_{\mathcal{E}_{\alpha \beta}} [\neg \alpha] . \]

From $\alpha \lor \beta \in (\mathcal{K}_\alpha)_\beta^\neg$ we get that
\[ min_{(\mathcal{E}_{\alpha \beta})} (-\alpha) \subseteq [\alpha \lor \beta], \text{ i.e., } min_{\mathcal{E}_{\alpha \beta}} [\neg \beta] \subseteq [\alpha] \text{ whereby } min_{\mathcal{E}_{\alpha \beta}} [\neg \beta] = min_{\mathcal{E}_\alpha} [\neg \beta \land \alpha]. \]

Since $-\alpha$ satisfies $\mathcal{OP}_\alpha[\alpha]$ we have $min_{\mathcal{E}_\alpha} [\neg \beta] = min_{\mathcal{E}_{\alpha \beta}} [\neg \beta \land \alpha] = min_{\mathcal{E}_{\alpha \beta}} [\neg \beta \land \alpha]$. Therefore Eqn. G2 gives $(\mathcal{K}_\alpha)_\beta^\neg = \mathcal{K}_\alpha \cap \mathcal{K}_{\alpha \rightarrow \beta}$ as desired.

Similarly, if $\alpha \rightarrow \beta \in (\mathcal{K}_\alpha)_\beta^\neg$ then we get $min_{(\mathcal{E}_{\alpha \beta})} (-\alpha) \subseteq [\alpha \rightarrow \beta]$, that is,
\[ min_{\mathcal{E}_{\alpha \beta}} [\neg \beta] \subseteq [\neg \alpha] \text{ whereby } min_{\mathcal{E}_{\alpha \beta}} [\neg \beta] = min_{\mathcal{E}_{\alpha \beta}} [\neg \beta \land \neg \alpha]. \]

Now since $-\alpha$ satisfies both $\mathcal{OP}_\alpha[\alpha]$ and $\mathcal{OP}_\alpha[\neg \alpha]$, we have $min_{\mathcal{E}_{\alpha \beta}} [\neg \beta \land \neg \alpha] = min_{\mathcal{E}_{\alpha \beta}} [\neg \beta \land \neg \alpha]$. Hence we get
\[ (\mathcal{K}_\alpha)_\beta^\neg = \mathcal{K}_\alpha \cap \mathcal{K}_{\alpha \rightarrow \beta} \cap \mathcal{K}_{\alpha \rightarrow \beta}. \] Thus $-\alpha$ satisfies PFI.

\textbf{Lemma 2.3} An AGM-rational contraction function – satisfies Eqn. 2.5 iff for every sentence $\alpha$, both: (a) $-\alpha$ satisfies $\mathcal{OP}_\alpha[\neg \alpha]$, and (b) $\omega \sqsubseteq_{\alpha} \omega'$ for every $\omega, \omega' \in \Omega$ such that $\omega \in [\neg \alpha]$ and $\omega' \in [\alpha] \setminus min_{\mathcal{E}_\alpha}[\alpha]$.

\textbf{Proof:} Let $-\alpha$ be a contraction function which satisfies Eqn. G2. Let $\square$ denote a consistent belief state and $\mathcal{K}$ the corresponding belief set.

\[ \]
(Left to Right). Assume that $\neg$ satisfies Eqn. 2.5. We need to show that when contracting by an arbitrary sentence $\alpha$, $\neg$ changes the belief state $\models$ to $\models \omega$ such that $\neg$ satisfies $\mathcal{OP}_\alpha[-\alpha]$ and for all $\omega, \omega' \in \Omega$ such that $\omega \in [-\alpha]$ and $\omega' \in [\alpha] \setminus \text{min}_C[\alpha]$, we have $\omega \models \omega'$.

Note that if $\models \alpha$, the set of worlds $[-\alpha]$ is empty and then $\neg$ trivially satisfies the requirements (a) and (b). We consider the non-trivial case of $\not\models \alpha$ in detail.

PART (a): First we show that $\neg$ satisfies $\mathcal{OP}_\alpha[-\alpha]$. Consider $\omega, \omega' \in [-\alpha]$. Suppose that $\omega \subseteq \omega'$. Since the language $\mathcal{L}$ is finitely generated, there exists a sentence $\beta' \in \mathcal{L}$ such that $[-\beta'] = \{\omega, \omega'\}$. Therefore we have $\models \alpha \rightarrow \beta'$ and $\models \alpha \lor \beta'$. Suppose $\omega \in \text{min}_C[\Omega]$ or $\omega \in \text{min}_C[-\alpha]$ then from Eqn. G2 we have $\omega \in \text{min}_C[\omega]$ and hence $\omega \models \omega'$. Alternatively, suppose that $\omega \notin \text{min}_C[-\alpha]$. We know that $\omega \in \text{min}_C[-\beta' \land -\alpha]$. Since $\neg$ satisfies Eqn. 2.5, $(\mathcal{K}_\alpha^-)_\beta^\lor = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \lor \beta'}^-$. Hence

$$\text{min}_{\text{min}_C[-\alpha]}[\omega] = \text{min}_C[\omega] \cup \text{min}_C[-\alpha] \cup \text{min}_C[-\beta' \land -\alpha].$$

Therefore from Eqn. G2, $\omega \in \text{min}_C[-\beta']$. This gives $\omega \models \omega'$. We can similarly show that if $\omega \subseteq \omega'$ then $\omega \models \omega'$. Hence $\neg$ satisfies $\mathcal{OP}_\alpha[-\alpha]$.

PART (b): Now we show that for all $\omega, \omega' \in \Omega$ such that $\omega \in [-\alpha]$ and $\omega' \in [\alpha] \setminus \text{min}_C[\alpha]$, we have $\omega \models \omega'$. Consider $\omega \in [-\alpha]$ and $\omega' \in [\alpha] \setminus \text{min}_C[\alpha]$. As the language $\mathcal{L}$ is finitely generated, there exists a sentence $\beta' \in \mathcal{L}$ such that $\omega$ and $\omega'$ are the only models of its negation. Two possibilities arise here:

Case (1): $\omega \in \text{min}_C[-\alpha]$. Then we will have $\omega \in \text{min}_C[-\alpha] \cap \text{min}_C[\omega]$ from which we conclude that $\omega \models \omega'$.

Case (2): $\omega \notin \text{min}_C[-\alpha]$. It is clear that $\beta'$ is such that $\models \alpha \lor \beta'$. Since $\neg$ satisfies Eqn. 2.5, $(\mathcal{K}_\alpha^-)_\beta^\lor = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \lor \beta'}^-$. Therefore with the aid of Eqn. G2 we get

$$\text{min}_{\text{min}_C[-\alpha]}[\omega] = \text{min}_C[\omega] \cup \text{min}_C[-\alpha] \cup \text{min}_C[-\beta' \land -\alpha].$$

We know that since $[-\beta' \land -\alpha]$ is a singleton set, $\omega \in \text{min}_C[-\beta' \land -\alpha]$. But we have assumed that $\omega \notin \text{min}_C[\Omega]$ and $\omega \notin \text{min}_C[-\alpha]$. This shows that $\omega \in \text{min}_C[-\beta']$. Also since $\omega' \notin \text{min}_C[-\beta' \land -\alpha]$ and $\omega' \notin \text{min}_C[-\beta']$, we have $\omega' \notin \text{min}_C[-\beta']$. This gives
\( \omega \subseteq_{\alpha} \omega' \), as desired.

(Right to Left). Let \(-\) be a contraction function that satisfies Eqn. G2. For every belief \( \alpha \), assume that \(-\) satisfies \( \mathcal{OP}_{\alpha}[-\alpha] \) and \( \omega \subseteq_{\alpha} \omega' \) for every \( \omega \in [-\alpha] \) and \( \omega' \in [\alpha] \setminus min_{\alpha}[\alpha] \). We need to show that \(-\) satisfies Eqn. 2.5 that is when \( \alpha \) is a belief of the agent and \( \beta \) is an arbitrary belief such that \( \forall \alpha \lor \beta \) and \( \beta \in K_{\alpha}^{-} \) then
\[
(K_{\alpha})_{\beta}^{-} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-}. \]

In fact we will prove a stronger result by showing that
\[
(K_{\alpha})_{\beta}^{-} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \text{ even when } \beta \notin K_{\alpha}. \]

Case 1: Consider \( \beta \notin K_{\alpha}^- \). Then we have \( min_{\in\alpha} [-\beta] \subseteq min_{\in\alpha} (\Omega) \). Therefore from Eqn. G2 we get \( min_{\in\alpha} (\Sigma_{\beta} \cap \Omega) = min_{\in\alpha} (\Omega) \). Since we have \( \beta \in K \) we can say that \( min_{\in\alpha} [-\beta] = min_{\in\alpha} [-\alpha] \cup min_{\in\alpha} [-\alpha \land -\beta] \). Therefore \( (K_{\alpha})_{\beta}^{-} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-} \).

Case 2: Consider \( \beta \in K_{\alpha}^- \). This gives \( min_{\in\alpha} (\Omega) \subseteq [\beta]. \) Note that \( \alpha \in K \), and hence \( min_{\in\alpha} [\alpha] = min_{\in}(\Omega) \). Therefore there exists no model in \( min_{\in\alpha}[\alpha] \) or in \( min_{\in\alpha}[-\alpha] \) which is also a model of \( -\beta \). Since \( \forall \alpha \lor \beta \) there is a world \( \omega \in [-\alpha] \) which models \( -\beta \). Therefore \( \omega \subseteq_{\alpha} \omega' \) for every \( \omega' \) which is a model of \( \alpha \land -\beta \).

This gives \( min_{\in\alpha} [-\beta] \subseteq [-\beta \land -\alpha] \). By our hypothesis the contraction function \(-\) satisfies \( \mathcal{OP}_{\alpha}[-\alpha] \). Therefore \( min_{\in\alpha} [-\beta] = min_{\in\alpha} [-\beta \land -\alpha] = min_{\in\alpha} [-\beta \land -\alpha] \).

Hence we have \( min_{\in\alpha} (\Sigma_{\beta} \cap \Omega) = min_{\in}(\Omega) \cup min_{\in\alpha} [-\beta] \cup min_{\in\alpha} [-\alpha \land -\beta] \) which gives
\[
(K_{\alpha})_{\beta}^{-} = K_{\alpha}^{-} \cap K_{\alpha \lor \beta}^{-}. \]

Therefore \(-\) satisfies Eqn. 2.5.

Lemma 2.4 For any arbitrary beliefs \( \alpha \) and \( \beta \) such that \( \beta \in K_{\alpha}^{-} \), an AGM-rational contraction function \(-\) satisfies Eqns. 2.1, 2.3 and 2.6 if and only if
\[
\forall \omega, \omega' \notin (min_{\in}(\Omega) \cup min_{\in}[-\alpha]), \omega \subseteq_{\alpha} \omega' \iff \omega \subseteq \omega'. \]

Proof: Let \(-\) be a contraction function that satisfies Eqn. G2. Let \( K \) be the belief set associated with a consistent belief state \( \subseteq \). Let \( \alpha, \beta \) be two arbitrary beliefs such that \( \beta \in K_{\alpha}^- \). We need to show that \(-\) satisfies Eqns. 2.1, 2.3 and 2.6 if and only if \( \forall \omega, \omega' \in \Omega \) such that \( \omega, \omega' \notin (min_{\in}(\Omega) \cup min_{\in}[-\alpha]) \) we have \( \omega \subseteq_{\alpha} \omega' \iff \omega \subseteq \omega' \).

(Left to Right). Let \(-\) satisfy Eqns. 2.1, 2.3 and 2.6. Let \( \omega, \omega' \in \Omega \) be such that
\( \omega, \omega' \notin \text{min}_\subseteq (\Omega) \) and \( \omega, \omega' \notin \text{min}_\subseteq [\neg \alpha] \). Since \( \mathcal{L} \) is finitely generated, there is a sentence \( \beta' \in \mathcal{L} \) such that \( \omega \) and \( \omega' \) are the only models of its negation. By our assumption of \( \omega, \omega' \) we get that \( \beta' \in \mathcal{K}_\alpha^- \).

Case 1: Suppose \( \omega \models \alpha \) and \( \omega' \models \alpha \). Then \( \vdash \alpha \lor \beta' \). Therefore \( \alpha \lor \beta' \in \mathcal{K}_\beta^- \) and hence from Eqn 2.1 we get \( (\mathcal{K}_\alpha^-)_{\beta'} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta'} \). From Eqn. G2, we can write this as \( \text{min}_{\subseteq-}(\Omega) \cup \text{min}_{\subseteq-}[\neg \beta'] = \text{min}_{\subseteq}(\Omega) \cup \text{min}_{\subseteq}[\neg \alpha] \cup \text{min}_{\subseteq}[\neg \beta' \land \alpha] \). However, since \( \beta' \in \mathcal{K}_\alpha^- \) and \( \vdash \alpha \lor \beta' \) we have \( \text{min}_{\subseteq-}[\neg \beta'] = \text{min}_{\subseteq}[\neg \beta' \land \alpha] \). Therefore \( \omega \subseteq \omega' \) iff \( \omega \subseteq \omega' \).

Case 2: Suppose \( \omega \models \neg \alpha \) and \( \omega' \models \neg \alpha \). Along the same lines as presented above, since \( \neg \) satisfies Eqn. 2.3 we have \( \omega \subseteq \omega' \) iff \( \omega \subseteq \omega' \).

Case 3: Suppose \( \omega \models \alpha \), \( \omega' \models \neg \alpha \) and \( \omega \sqsubseteq \omega' \). Therefore \( \omega \in \text{min}_{\subseteq}[\neg \beta'] \). From Eqn. G2 \( \text{min}_{\subseteq-}(\Omega) = \text{min}_{\subseteq}(\Omega) \cup \text{min}_{\subseteq}[\neg \beta'] \). Since we know that \( \alpha \) is a belief in \( \mathcal{K} \), we have that \( \text{min}_{\subseteq-}(\Omega) \subseteq [\alpha] \). From this we can derive that \( \alpha \lor \beta' \in \mathcal{K}_\beta^- \). The contraction function \( \vdash \) satisfies Eqn. 2.1. Therefore \( (\mathcal{K}_\alpha^-)_{\beta'} = \mathcal{K}_\alpha^- \cap \mathcal{K}_{\alpha \rightarrow \beta'} \). From Eqn. G2 we can write this as \( \text{min}_{\subseteq-}(\Omega) \cup \text{min}_{\subseteq-}[\neg \beta'] = \text{min}_{\subseteq}(\Omega) \cup \text{min}_{\subseteq}[\neg \alpha] \cup \text{min}_{\subseteq}[\neg \beta' \land \alpha] \). Since \( \beta' \in \mathcal{K}_\alpha^- \) we get \( \text{min}_{\subseteq-}[\neg \beta'] = \text{min}_{\subseteq}[\neg \beta' \land \alpha] \). This gives that \( \omega \subseteq \omega' \).

Case 4: Suppose \( \omega \models \alpha \), \( \omega' \models \neg \alpha \) and \( \omega' \sqsubseteq \omega \). Along the same lines as in case 3 we can show that, since \( \vdash \) satisfies Eqn. 2.3 we have \( \omega' \subseteq \omega \).

Case 5: Suppose \( \omega \models \alpha \) and \( \omega' \models \neg \alpha \). Also let \( \omega \approx \omega' \). This case too follows in similar lines as in case 3 and 4. In this case neither \( \alpha \lor \beta' \) nor \( \alpha \rightarrow \beta' \) belong to \( \mathcal{K}_{\beta'}^- \). Since \( \vdash \) satisfies Eqn. 2.6 it follows that \( \omega \approx \omega' \).

(Right to Left). Let \( \vdash \) be a contraction function that satisfies Eqn. G2. Given a belief state \( \subseteq \), upon contraction by an arbitrary belief \( \alpha \), let \( \vdash \) change the belief state to \( \subseteq^{-}_\alpha \) where for every \( \omega, \omega' \in \Omega \) such that \( \omega, \omega' \notin (\text{min}_{\subseteq}(\Omega) \cup \text{min}_{\subseteq}[\neg \alpha]) \) we have \( \omega \subseteq \omega' \) iff \( \omega \subseteq \omega' \). Let \( \beta \) be an arbitrary belief such that \( \beta \in \mathcal{K}_\alpha^- \).

Case 1: \( \vdash \beta \). Hence both \( \alpha \lor \beta \) and \( \alpha \rightarrow \beta \), being theorems, belong to \( \mathcal{K}_{\beta}^- \). Therefore this case satisfies the pre-conditions of both Eqns. 2.1 and 2.3, but not of Eqn. 2.6. Thus
Eqn. 2.6 is trivially satisfied. As to Eqns. 2.1 and 2.3, note that $[\beta] = [\alpha \lor \beta] = [\alpha \rightarrow \beta] = \Omega$, and hence $\min_{\subseteq}[-\beta] = \min_{\subseteq}[-\alpha \land -\beta] = \min_{\subseteq}[\alpha \land -\beta] = \emptyset$. The desired result easily follows from it.

Case 2: $\vdash \alpha$ and $\not\vdash \beta$. It follows that $\alpha \lor \beta$ is a theorem, and hence belongs to $K_{\beta}$. Now $\alpha \rightarrow \beta$ is logically equivalent to $\beta$; hence $K_{\beta} = K_{\alpha \rightarrow \beta}$. Furthermore, since $\vdash \alpha$, we get $\subseteq = \subseteq$, whereby, $K_{\alpha} = \subseteq$. Hence $[(K_{\alpha})_{\beta}] = \min_{\subseteq}[-\beta] = \min_{\subseteq}[\alpha \land -\beta]$ wherefrom the desired result $(K_{\alpha})_{\beta} = K_{\alpha \rightarrow \beta} = K_{\alpha} \cap K_{\alpha \rightarrow \beta}$ follows.

Case 3: Let $\not\vdash \alpha$ and $\not\vdash \beta$. Three subcases arise here:

Case (3a): Suppose $\alpha$ and $\beta$ are such that $\alpha \lor \beta \in K_{\beta}$. From Eqn. G2 we can write $\min_{\subseteq}(\Omega) \subseteq [\alpha \lor \beta]$. Therefore $\min_{\subseteq}[-\beta] \subseteq [\alpha]$, i.e. $\min_{\subseteq}[-\beta] = \min_{\subseteq}[-\beta \land \alpha]$. Since $\beta \in K_{\alpha}$, if $\omega$ is a model of $-\beta$ then $\omega \not\in (\min_{\subseteq}(\Omega) \cup \min_{\subseteq}[-\alpha])$. By our hypothesis $-\vdash$ preserves ordering within $[-\beta]$, that is, $-\vdash$ satisfies $\bot \alpha [\not\vdash \beta]$. Hence $\min_{\subseteq}[-\beta] = \min_{\subseteq}[-\beta]$. This gives $\min_{\subseteq}[-\beta] = \min_{\subseteq}[-\beta \land \alpha]$. Therefore when $\alpha \lor \beta \in K_{\beta}$ we have $(K_{\alpha})_{\beta} = K_{\alpha} \cap K_{\alpha \rightarrow \beta}$, that is, $-\vdash$ satisfies Eqn. 2.1.

Case (3b): Suppose $\alpha$ and $\beta$ are such that $\alpha \rightarrow \beta \in K_{\beta}$. Following along the same lines as presented for Case (3a) we can show that $-\vdash$ satisfies Eqn. 2.3.

Case (3c): Suppose $\alpha$ and $\beta$ are such that $\alpha \lor \beta \not\in K_{\beta}$ and $\alpha \rightarrow \beta \not\in K_{\beta}$. With similar arguments as in Case (3a) we can show that $-\vdash$ satisfies Eqn. 2.6.

\[\blacksquare\]

**Observation 2.1** Let $-\vdash$ be an AGM-rational contraction function, and the degree of belief function $d$ is appropriately related with the presumed belief state $\subseteq$. Then:

(a) Equation 2.1 is satisfied iff Equation 2.7 is,

(b) Equation 2.3 is satisfied iff Equation 2.8 is, and

(c) Equation 2.6 is satisfied iff Equation 2.9 is.

**Proof:**

Let us make the general assumption that $-\vdash$ is an AGM-rational contraction function...
and \( \sqsubseteq \), \( \mathcal{K} \) denote a consistent belief state and the corresponding belief set. Also assume that \( d \) is appropriately related to \( \sqsubseteq \). We provide the proofs of the three different parts of this observation separately.

**Part (a):**

Consider two arbitrary beliefs \( \alpha \) and \( \beta \). It will be sufficient to show that \( \alpha \lor \beta \in \mathcal{K}_{\beta} \) is equivalent to \( d(\alpha \lor \beta) > d(\alpha \rightarrow \beta) \).

*(Left to Right).* Let \( \alpha \lor \beta \in \mathcal{K}_{\beta} \). This implies that \( \min_{\sqsubseteq_{\beta}}(\Omega) \subseteq [\alpha \lor \beta] \), whereby, \( \min_{\sqsubseteq}[-\beta] \subseteq [\alpha] \). Therefore for any world \( \omega' \) in \([\neg\beta \land \neg\alpha]\) and \( \omega \in \min_{\sqsubseteq}[-\beta \land \alpha] \) we have \( \omega \sqsubseteq \omega' \). In terms of degrees of belief we have \( d(\alpha \lor \beta) > d(\alpha \rightarrow \beta) \).

*(Right to Left).* For the reverse, we just need to trace our step backwards in the proof for the \((\Rightarrow)\) part.

**Part (b):**

For the proof for the \((\Rightarrow)\) part, we begin by assuming that \( \alpha \rightarrow \beta \in \mathcal{K}_{\beta} \). This implies \( \min_{\sqsubseteq}[-\beta] \subseteq [-\alpha] \) which leads us to conclude \( d(\alpha \rightarrow \beta) > d(\alpha \lor \beta) \). The proof for the \((\Leftarrow)\) part can be obtained by simply tracing our steps in the \((\Rightarrow)\) part backwards.

**Part (c):**

For the proof for the \((\Rightarrow)\) part, we begin by assuming that neither \( \alpha \rightarrow \beta \) nor \( \alpha \lor \beta \) are retained in \( \mathcal{K}_{\beta} \). This implies \( \min_{\sqsubseteq}[-\beta] = \min_{\sqsubseteq}[-\beta \land \alpha] \cup \min_{\sqsubseteq}[-\beta \land \neg\alpha] \) and from the definition of degrees of belief we deduce that \( d(\alpha \rightarrow \beta) = d(\alpha \lor \beta) \). Again the proof for the \((\Leftarrow)\) part can be obtained by simply tracing the above steps backwards.

**Lemma 2.5** Any AGM-rational contraction function \( - \) that satisfies Equation 2.13 also satisfies Equation 2.2.

**Proof:** Let \( - \) be an AGM-rational contraction function that satisfies Eqn. 2.13. Let \( \sqsubseteq \) be the presumed belief state and \( \mathcal{K} \) be its associated belief set that is consistent. We need to show that for any two beliefs \( \alpha \) and \( \beta \) such that \( \vdash \alpha \lor \beta \) we have \( (\mathcal{K}_{\alpha})_{\beta} = \mathcal{K}_{\alpha} \cap \mathcal{K}_{\alpha \rightarrow \beta} \).
Case 1: Suppose $\alpha$ is a sentence such that $\vdash \alpha$. Since $-$ is an AGM-rational contraction function we have $(K\alpha^-)\beta = K\beta^- = K\alpha^- \cap K\alpha^\rightarrow \beta$. Hence $-$ satisfies Eqn. 2.2 trivially.

Case 2: Suppose $\vdash \beta$. Based on similar arguments as presented in case 1, we can say that $-$ trivially satisfies Eqn. 2.2.

Case 3: Suppose $\forall \alpha, \forall \beta$ but $\vdash \alpha \lor \beta$. By Property (1) of conditional degrees of belief (given in Section 2.3.1) we have $d(\alpha \lor \beta | \neg \alpha) = \infty$. However $d(\alpha \rightarrow \beta | \alpha) < \infty$ since, the alternative, $d(\alpha \rightarrow \beta | \alpha) = \infty$ would yield $\vdash \alpha \rightarrow \beta$ which is not possible given our assumptions $\vdash \alpha \lor \beta$ and $\nvdash \beta$. Therefore we have $d(\alpha \lor \beta | \neg \alpha) > d(\alpha \rightarrow \beta | \alpha)$.

From Eqn. 2.13 we get $(K\alpha^-)\beta = K\alpha^- \cap K\alpha^\rightarrow \beta$. Hence $-$ satisfies Eqn. 2.2.

Lemma 2.6 Any AGM-rational contraction function $-$ that satisfies Equation 2.12 also satisfies Equation 2.4.

Proof: Let $-$ be an AGM-rational contraction function that satisfies Eqn. 2.12. Let $\mathcal{K}$ be the underlying belief state and $\mathcal{K}$ its associated, consistent belief set, and $d$ the relevant degree of belief function. Assume two beliefs $\alpha$ and $\beta$ such that $\vdash \alpha \rightarrow \beta$. We need to show that $(K\alpha^-)\beta = K\alpha^- \cap K\alpha^\rightarrow \beta$.

Case 1: Assume $\vdash \beta$. Since $-$ is AGM-rational, we have $(K\alpha^-)\beta = K\alpha^-$. Furthermore, since $\vdash \alpha \lor \beta$, we have $K\alpha^\lor \beta = K$. Hence we get the desired result, $(K\alpha^-)\beta = K\alpha^- \cap K\alpha^\rightarrow \beta$.

Case 2: Assume $\vdash \alpha$. Since by assumption $\vdash \alpha \rightarrow \beta$, it follows that $\vdash \beta$. Thus it reduces to Case 1.

Case 3: Suppose $\forall \alpha, \forall \beta$ but $\vdash \alpha \rightarrow \beta$. Now, by Property (1) of conditional degrees of belief (given in Section 2.3.1), we have $d(\alpha \rightarrow \beta | \alpha) = \infty$. However, $d(\alpha \lor \beta | \neg \alpha) < \infty$ because $d(\alpha \lor \beta | \neg \alpha) = \infty$ would give $d(\alpha \lor \beta) = \infty$. That would lead to $\vdash \beta$. Therefore we have $d(\alpha \rightarrow \beta | \alpha) > d(\alpha \lor \beta | \neg \alpha)$. From Eqn. 2.12 we get $(K\alpha^-)\beta = K\alpha^- \cap K\alpha^\rightarrow \beta$, as desired.
**Theorem 2.2** An AGM-rational contraction function is a moderate contraction function iff it satisfies Equation 2.2 and 2.5.

**Proof:** Let $-\alpha$ be an AGM-rational contraction function. Consider the belief set $\mathcal{K}$ which is obtained from a given consistent belief state $\sqsubseteq_{\alpha}$ by Eqn. G1. We need to show that the contraction function $-\alpha$ is a moderate contraction function iff it satisfies Eqns. 2.2 and 2.5 for any two arbitrary beliefs $\alpha, \beta$ in $\mathcal{K}$.

*(Left to Right).* Let us assume that $-\alpha$ is a moderate contraction function. By definition of a moderate contraction function we know that $-\alpha$ satisfies $OP_{\alpha}[\alpha]$ for any arbitrary belief $\alpha$. Hence from Lemma 2.1 it is clear that $-\alpha$ satisfies Eqn. 2.2.

Furthermore, consider any two $\omega$ and $\omega'$ such that $\omega \in \neg[\alpha]$ and $\omega' \in [\alpha] \setminus \min_{\subseteq}(\Omega)$. Since $\alpha$ is a belief in $\mathcal{K}$, we have $\min_{\subseteq}(\Omega) = \min_{\subseteq}[\alpha]$. Furthermore, since $\omega \in \neg[\alpha]$ and $\omega' \in [\alpha] \setminus \min_{\subseteq}[\alpha]$, we get $\omega \subseteq_{\alpha} \omega'$. Since we also have that $-\alpha$ satisfies $OP_{\alpha}[\neg\alpha]$, Lemma 2.3 leads us to conclude that $-\alpha$ satisfies Eqn. 2.5.

*(Right to Left).* Let $-\alpha$ be an AGM-rational contraction function that satisfies Eqns. 2.2 and 2.5 when successively contracting two arbitrary beliefs $\alpha$ and $\beta$. We need to show that $-\alpha$ is a moderate contraction function.

Since $-\alpha$ satisfies Eqn. 2.2, we can conclude from Lemma 2.1 that $-\alpha$ satisfies $OP_{\alpha}[\alpha]$. By our assumption that $-\alpha$ satisfies Eqn. 2.5 and from Lemma 2.3, we also have that $-\alpha$ satisfies $OP_{\alpha}[\neg\alpha]$. We can also conclude that for any two worlds $\omega$ and $\omega'$ such that $\omega \in \neg[\alpha]$ and $\omega' \in [\alpha] \setminus \min_{\subseteq}[\alpha]$, we have $\omega \subseteq_{\alpha} \omega'$. Since $\alpha$ is a belief of the agent, we also have $\min_{\subseteq}[\alpha] = \min_{\subseteq}(\Omega)$. Therefore we have that $\omega \subseteq_{\alpha} \omega'$ when $\omega \in \neg[\alpha]$ and $\omega' \in [\alpha] \setminus \min_{\subseteq}(\Omega)$. Hence $-\alpha$ is a moderate contraction function, satisfying conditions MC1 - MC4 as presented in the Section 2.2.1.

**Theorem 2.3** An AGM-rational contraction function is a natural contraction function iff it satisfies Equations 2.1, 2.3 and 2.6.

**Proof:** The proof of this theorem follows directly from Lemma 2.4. ■

132
Theorem 2.4  An AGM-rational contraction function is a lexicographic contraction function iff it satisfies Equations 2.12, 2.13 and 2.14

Proof:  Let $\Box$ be an AGM-rational contraction function. Let $\sqsubseteq$ represent a consistent belief state and $\mathcal{K}$ denote the corresponding belief set. Let $\alpha$ and $\beta$ be any arbitrary beliefs. We need to show that the contraction function $\Box$ is a lexicographic contraction function iff it satisfies Eqs. 2.12, 2.13 and 2.14.

(Left to Right). Let $\Box$ be a lexicographic contraction function. We need to show that for any arbitrary beliefs $\alpha$, $\beta$, $\Box$ satisfies Equations 2.12, 2.13 and 2.14. We do that by analysing four possible cases.

Case 1: Let both $\vdash \alpha$ and $\vdash \beta$. Then $\mathcal{K} = \mathcal{K}_{\alpha}^{\neg} = (\mathcal{K}_{\alpha}^{\neg})_{\beta} = \mathcal{K}_{\alpha \lor \beta} = \mathcal{K}_{\alpha \rightarrow \beta}$. Hence the Equations 2.12, 2.13 and 2.14 are trivially satisfied.

Case 2: Let $\vdash \alpha$ and $\nvdash \beta$. By the properties of conditional degrees of belief, we have $d(\alpha \lor \beta | \neg \alpha) = \infty$ (since $\vdash \alpha \lor \beta$) and $d(\alpha \rightarrow \beta | \alpha) = d(\beta)$ which is less than $\infty$. Therefore $d(\alpha \lor \beta | \neg \alpha) > d(\alpha \rightarrow \beta | \alpha)$ which satisfies the pre-condition for Eqn. 2.13. Clearly then, the preconditions of Eqns. 2.12 and 2.14 are not satisfied, whereby Eqns. 2.12 and 2.14 are trivially satisfied. As to Eqn. 2.13, since $\Box$ is a lexicographic contraction function it satisfies $\mathcal{OP}_\alpha[\alpha]$ whereby $\sqsubseteq_{\alpha} \sqsubseteq = \sqsubseteq$. Therefore we have $\min_{\sqsubseteq_{\alpha}}[\neg \beta] = \min_{\sqsubseteq}[\neg \beta] = \min_{\sqsubseteq}[\neg \beta \land \alpha]$. Therefore from Eqn. G2 we get $(\mathcal{K}_{\alpha}^{\neg})_{\beta} = \mathcal{K}_{\beta} = \mathcal{K}_{\alpha} \cap \mathcal{K}_{\beta} = \mathcal{K}_{\alpha} \cap \mathcal{K}_{\alpha \rightarrow \beta}$, as desired.

Case 3: Let $\nvdash \alpha$ and $\vdash \beta$. Since $\Box$ satisfies Eqn. G2 we note that the second contraction in $(\mathcal{K}_{\alpha}^{\neg})_{\beta}$ is vacuous. From this we can see that $\Box$ trivially satisfies Eqns. 2.12, 2.13 and 2.14.

Case 4: Let both $\nvdash \alpha, \nvdash \beta$. We discuss two subcases here.

Sub-case (4a): Suppose $\beta \notin \mathcal{K}_{\alpha}^{\neg}$. From Eqn. G2 and since $\beta$ is a belief in $\mathcal{K}$, we see that there is a model of $\neg \beta$ in $\min_{\sqsubseteq}[\neg \alpha]$. Any complete chain of worlds in $[\neg \alpha]$ begin with some world in $\min_{\sqsubseteq}[\neg \alpha]$. Since there exists a model of $\neg \beta$ in $\min_{\sqsubseteq}[\neg \alpha]$ we have
\(d(\beta | \neg \alpha) = 0\), that is, \(d(\alpha \lor \beta | \neg \alpha) = 0\). From \(\beta \in \mathcal{K}\) we get that \(d(\beta | \alpha) > 0\), that is, \(d(\alpha \rightarrow \beta | \alpha) > 0\). Hence \(d(\alpha \rightarrow \beta | \alpha) > d(\alpha \lor \beta | \neg \alpha)\). Note hence that this satisfies the precondition of Eqn. 2.12, and hence Eqns. 2.13 and 2.14 are trivially satisfied. We need only to show that Eqn. 2.12 is satisfied.

Now, \(\min_{\overline{E_{\alpha}}} (\Omega) = \min_{\overline{E_{\alpha}}} (\Omega) \cup \min_{\overline{E_{\alpha}}} \neg \beta\). Since \(\beta \notin \mathcal{K}_{\alpha}\) we have

\(\min_{\overline{E_{\alpha}}} \neg \beta \subseteq \min_{\overline{E_{\alpha}}} (\Omega)\). Therefore \(\min_{\overline{E_{\alpha}}} (\Omega) = \min_{\overline{E_{\alpha}}} (\Omega)\), that is, \((\mathcal{K}_{\alpha})_{\beta} = \mathcal{K}_{\alpha}\).

Again from \(\beta \in \mathcal{K}\) and \(\beta \notin \mathcal{K}_{\alpha}\) we get \(\min_{\overline{E}} \neg \alpha \cap \neg \beta \neq \emptyset\). We can therefore write

\(\min_{\overline{E}} \neg \alpha = \min_{\overline{E}} \neg \alpha \cup \min_{\overline{E}} \neg \alpha \land \neg \beta\). From Eqn. G2 we get

\(\min_{\overline{E}} (\Omega) = \min_{\overline{E}} (\Omega) \cup \min_{\overline{E}} \neg \alpha = \min_{\overline{E}} (\Omega) \cup \min_{\overline{E}} \neg \alpha \cup \min_{\overline{E}} \neg \alpha \land \neg \beta\).

Therefore \((\mathcal{K}_{\alpha})_{\beta} = \mathcal{K}_{\alpha} \cap \mathcal{K}_{\alpha \lor \beta}\), as desired.

Sub-case (4b): Suppose \(\beta \in \mathcal{K}_{\alpha}\). We will show that Eqn. 2.13 is satisfied, and leave out Eqn. 2.12 and 2.14 since the proofs are analogous. Let us assume that

\(d(\alpha \lor \beta | \neg \alpha) > d(\alpha \rightarrow \beta | \alpha)\). We wish to first show that \(\min_{\overline{E_{\alpha}}} \neg \beta = \min_{\overline{E}} \neg \beta \land \alpha\) from which the desired result will easily follow. First the trivial case: if \(\vdash \alpha \lor \beta\), then clearly \(\neg \beta \subseteq [\alpha]\) whereby \(\min_{\overline{E_{\alpha}}} \neg \beta = \min_{\overline{E}} \neg \beta \land \alpha\). Now the non-trivial case:

assume that \(\not \vdash \alpha \lor \beta\), that is, \([\neg \alpha \land \neg \beta] \neq \emptyset\). Let \(\omega' \in \min_{\overline{E}} \neg \beta \land \neg \alpha\). Now, it is easily noted that \(\not \vdash \alpha \rightarrow \beta\), that is, \([\alpha \land \neg \beta] \neq \emptyset\). Let \(\omega \in \min_{\overline{E}} \neg \beta \land \neg \alpha\). Hence, by the properties of conditional degree of beliefs, there is a chain of worlds in \([\neg \alpha]\) ending in \(\omega'\) whose length is greater than any chain of worlds in \([\alpha]\) ending in \(\omega\). Hence we have \(\omega \subseteq \omega'\). Therefore, from \(\mathcal{OP}_{\alpha} [\alpha]\) we get \(\min_{\overline{E_{\alpha}}} \neg \beta = \min_{\overline{E}} \neg \beta \land \alpha\). From Eqn. G2 we get \((\mathcal{K}_{\alpha})_{\beta} = \mathcal{K}_{\alpha} \cap \mathcal{K}_{\alpha \rightarrow \beta}\). Hence we see that \(\neg\) satisfies Eqn. 2.13. Similarly we can show that \(\neg\) satisfies Eqn. 2.12 and 2.14.

(Right to Left). Let \(\neg\) be an AGM-rational contraction function which satisfies Eqns. 2.12, 2.13 and 2.14 when contracting two arbitrary beliefs \(\alpha\) and \(\beta\). We need to show that \(\neg\) is a lexicographic contraction function, that is, it satisfies \(\mathcal{OP}_{\alpha} [\alpha], \mathcal{OP}_{\alpha} [\neg \alpha]\) and LC3.

Lemma 2.5 states that any AGM contraction function that satisfies Eqn. 2.13 also

134
satisfies Eqn. 2.2. Now from Lemma 2.1 we know that any AGM contraction function that satisfies Eqn. 2.2 also satisfies $\mathcal{CP}_a[\alpha]$. Hence we see that $\neg \alpha$ satisfies $\mathcal{CP}_a[\neg \alpha]$.

Hence it will be sufficient to show that $\neg \alpha$ satisfies $\mathcal{LC}_3$.

Case 1: $\vdash \alpha$. We have shown that $\neg \alpha$ satisfies $\mathcal{OP}_a[\alpha]$. Hence, we have $\subseteq_a = \subseteq$.

Case 2: $\nvdash \alpha$. Consider two worlds $\omega$ and $\omega'$ such that $\omega$ is a model of $\alpha$ and $\omega'$ is a model of $\neg \alpha$. Let $n$ denote the minimum length of a complete chain of worlds in $[\alpha]$ that end in $\omega$ and $m$ denote the minimum length of a complete chain of worlds in $[\neg \alpha]$ that end in $\omega'$. Since $\mathcal{L}$ is finitely generated there exists a sentence $\beta$ in the language such that the only models of $\neg \beta$ are $\omega$ and $\omega'$.

**Subcase (2a):** Suppose $n = 0$ whereby, $n \leq m$. Then we have $\omega \in min_{\subseteq}[\alpha]$, that is, $\omega \in min_{\subseteq}(\Omega)$. Therefore $\omega \subseteq_a \subseteq \omega'$ when $n = 0$. However suppose $m = 0$, that is, $m \leq n$. Then $\omega' \in min_{\subseteq}[\neg \alpha]$. Since $\neg \alpha$ satisfies $\mathcal{OP}_a[\neg \alpha]$.

**Subcase (2b):** Suppose $n,m \neq 0$ and $n < m$. With similar reasoning as in case b above, together with the fact that $\neg \beta \in K_{\neg \alpha}$, following in similar lines as presented in case b, we can conclude that $\omega \approx_a \omega'$.

Thus $\neg \alpha$ satisfies $\mathcal{LC}_3$ and hence it is a lexicographic contraction function. $\blacksquare$
Appendix B

Proofs for Chapter 3

Observation 3.1 Partial imaging is homomorphic.

Proof: Let \( P, P_1 \) and \( P_2 \) be probability functions such that \( P = P_1 b P_2 \), for some real number \( b \) between 0 and 1. As given in Equation 3.2, for a sentence \( \xi \) in the language and some real number \( a \) between 0 and 1, \( P_\xi^a = P_\xi a P \) that is, \( P_\xi^a = (P_1 b P_2)_\xi^a (P_1 b P_2) \).

For readability, we skip mentioning \( a \) in the subscripts, it is assumed to be present where every partial imaging is mentioned. We need to show that \( P_\xi^a = P_1^a b P_2^a \). In [29] it is shown that imaging is homomorphic, that is, \( P^\# = P_1^\# b P_2^\# \). Therefore, \( P_\xi^a = (P_1^\# b P_2^\#) a (P_1 b P_2) \).

The right hand side can be re-written as follows:

\[
= a \cdot (b \cdot P_1^\# + (1 - b) \cdot P_2^\#) + (1 - a) \cdot (b \cdot P_1 + (1 - b) \cdot P_2) \\
= b \cdot (a \cdot P_1^\# + (1 - a) \cdot P_1) + (1 - b) \cdot (a \cdot P_2^\# + (1 - a) \cdot P_2) \\
= b \cdot P_1^p + (1 - b) \cdot P_2^p. 
\]

Thus we get \( P_\xi^p = P_1^p b P_2^p \), proving partial imaging to be homomorphic.

Theorem 3.1 Partial imaging is a belief erasure function when \( 0 < a < 1 \).

Proof: Let \( P \) be a probability function which represents the belief state of an agent. Let the variable \( a \) take a value in \( (0, 1) \) and \( I = (\xi, a) \). We need to show that \( P_\xi^p \) satisfies all the erasure postulates E1, . . . , E6. It is important to note that we associate partial imaging \( P \) by a sentence \( \xi \) with the erasure of a belief \( \neg \xi \).

Proof for E1: Let \( \alpha \) be a sentence such that \( P_\xi^p(\alpha) = 1 \). From Equation 3.3 we gather
that, $P^p_T(\alpha) = 1$ iff both $P^\#_\zeta(\alpha) = 1$ and $P(\alpha) = 1$. Therefore, $P^p_T(\alpha) = 1$ only if $P(\alpha) = 1$. This shows that partial imaging satisfies $E1$.

Proof for $E2$: Suppose $\xi$ is such that $P(\xi) = 1$. As discussed in Section 3.4, Case 3, we have $P^p_T = P$ and hence $E2$ is satisfied.

Proof for $E3$: Consider a sentence $\xi$ such that $\neg \xi$ is a belief of the agent. Also assume that $\not \models \neg \xi$, that is, there exists a $\omega \in \Omega$ such that $\omega \models \xi$. Then for every possible world there exists a most similar $\xi$-world. Thus $P^p_T(\neg \xi) = 1 - a$ (from Equation 3.3a). Therefore we have $P^p_T(\neg \xi) < 1$, satisfying $E3$.

Proof for $E4$: Suppose $\gamma$ is a sentence such that $\not \models \neg \xi \leftrightarrow \neg \gamma$. Then $[\gamma] = [\xi]$, that is, models of $\xi$ are also the models of $\gamma$ and vice-versa. Therefore, for every world $\omega$, $\omega^\#_{\xi} = \omega^\#_{\gamma}$. This gives us $P^\#_\xi = P^\#_{\gamma}$ and hence $P^p_{\xi, a} = P^p_{\gamma, a}$. Hence partial imaging satisfies $E4$.

Proof for $E5$: Since expansion in probabilistic setting is modelled by conditionalization, we need to show that conditionalization of $P^p_T$ by $\neg \xi$ results in $P$. Let $P_f$ denote the result of conditionalization of $P^p_T$ by $\neg \xi$. When $\neg \xi$ is initially a belief, we have $P(\neg \xi) = 1$ and therefore for any $\omega$ which is a model of $\xi$, $P(\omega) = 0$. By definition of conditionalization, we get $P_f(\neg \xi) = 1$. Therefore, for any world $\omega$ which is a model of $\xi$, we have $P_f(\omega) = 0 = P(\omega)$. However, if $\omega$ is a model of $\neg \xi$, $P_f(\omega) = \frac{P^p_T(\omega)}{P^p_T(\neg \xi)} = \frac{(1-a) P(\omega)}{(1-a)} = P(\omega)$. Therefore we have $P_f(\omega) = P(\omega)$ for every possible world $\omega$, and hence $P_f = P$. This shows that partial imaging satisfies $E5$.

Proof for $E6$: From Observation 3.1, it can be easily seen that partial imaging satisfies $E6$. ■

Observation 3.2 If $P(\eta) = 1$ then $P^s_{\xi} = P^\#_{\xi}$.

Proof: Let $P$ be the initial probability function and $S$ be a selection function on $\Omega$. Let $\eta$ be the sentence associated with the selection function $S$ such that $[\eta] = S(P)$. Suppose $P(\eta) = 1$. We need to show that $P^s_{\xi} = P^\#_{\xi}$, for any sentence $\xi$. We will show that $P^s_{\xi}(\alpha) = P^\#_{\xi}(\alpha)$ for every sentence $\alpha \in \mathcal{L}$. 138
Consider a sentence $\alpha \in \mathcal{L}$. The probability of $\alpha$ after selective imaging is given by Equation 3.6, $P^s_\xi(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) + P(\alpha \land \lnot \eta)$. Since $P(\eta) = 1$, we have $P(\lnot \eta) = 0$ and therefore $P(\alpha \land \lnot \eta) = 0$. When $P(\eta) = 1$, we have that all the possible worlds with non-zero initial probability are chosen by $S$. In other words, when $P(\omega) > 0$ we have $\omega \in S(P)$, that is, $\omega(\eta) = 1$ and $P(\omega) = 0$ whenever the world $\omega$ is not a model of $\eta$, that is, $\omega(\eta) = 0$. Therefore, $\sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) = \sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha)$. This gives, $P^s_\xi(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) = P^\#_\xi(\alpha)$.

**Observation 3.3** When $P(\eta) = 0$ then $P^s_\xi = P$.

**Proof:** Consider a probability function $P$ and a selection function $S$. Let $\eta$ be the sentence associated with the selection function, that is, $[\eta] = S(P)$. Suppose $P(\eta) = 0$. We need to show that for every sentence $\xi$, $P^s_\xi = P$. We will prove this by showing that $P^s_\xi(\alpha) = P(\alpha)$, for every sentence $\alpha \in \mathcal{L}$.

Equation 3.6 gives that $P^s_\xi(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) + P(\alpha \land \lnot \eta)$. Since $P(\eta) = 0$, we have $P(\lnot \eta) = 1$, that is, all the worlds with non-zero initial probability are models of $\lnot \eta$. Hence $P(\alpha \land \lnot \eta) = P(\alpha)$. Again since $P(\eta) = 0$, we have $P(\omega) = 0$ whenever $\omega(\eta) = 1$. Therefore we have $\sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) = 0$.

Hence $P^s_\xi(\alpha) = P(\alpha)$.

**Observation 3.4** If $\eta \vdash \xi$ then $P^s_\xi = P$.

**Proof:** Let $P$ be the initial probability function, $S$ the associated selection function and $\eta$ a sentence such that $[\eta] = S(P)$. Let $\xi$ be a sentence in $\mathcal{L}$ such that $\eta \vdash \xi$. We need to show that for every sentence $\alpha \in \mathcal{L}$, $P^s_\xi(\alpha) = P(\alpha)$.

Equation 3.6 gives that $P^s_\xi(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) + P(\alpha \land \lnot \eta)$. Since $\eta \vdash \xi$, we have $[\eta] \subseteq [\xi]$. Therefore whenever $\omega(\eta) = 1$, for any possible world $\omega$, we have $\omega(\xi) = 1$. This gives that when $\omega(\eta) = 1$, $\omega^\#_\xi = \omega$. Therefore for any sentence $\alpha$ and any world $\omega$, $\omega(\eta) \cdot \omega^\#_\xi(\alpha) = \omega(\eta) \cdot \omega(\alpha) = \omega(\eta \land \alpha)$. This gives us, $P^s_\xi(\alpha) = \sum_{\Omega} P(\omega) \cdot \omega(\eta \land \alpha) + P(\alpha \land \lnot \eta) = P(\alpha \land \eta) + P(\alpha \land \lnot \eta) = P(\alpha)$.

**Observation 3.5** If $\eta \vdash \lnot \xi$ then $P^s_\xi(\lnot \eta) = 1$.  

139
Thus selective imaging satisfies\(E_4\) has non-zero probability. Therefore\(P\)

\[
P(\omega \cdot \omega^\# (\eta)) = P(\eta).
\]

Since \(\eta \vdash \neg \xi\), we have \([\xi] \subseteq [\neg \eta]\). Therefore \(\omega^\# (\eta) = 1\) for every \(\omega \in \Omega\). Hence from Equation 3.6 we have \(P^* (\neg \eta) = \sum_{\alpha} P(\omega) \cdot P(\omega (\eta)) + P(\neg \eta) = P(\eta) + P(\neg \eta) = 1\).

**Theorem 3.2** If the selection function \(S\) is such that the associated sentence \(\eta\) has a non-zero probability, then selective imaging satisfies the erasure postulates \(E_2, E_3\) and \(E_4\).

**Proof:** Let \(P\) be a probability function that represents the belief state of the agent and \(S\) be a selection function. Since the language \(L\) is finitely generated, there exists a sentence \(\eta\) in \(L\) such that \([\eta] = S(P)\). Let the selection function be such that \(P(\eta) > 0\).

**Proof for \(E_2\):** Suppose that \(\xi\) is a sentence such that \(P(\xi) = 1\). For any sentence \(\alpha \in L\), \(P(\omega) \cdot \omega^\# (\alpha) \cdot P(\eta) = 0\) whenever \(\omega \models \neg \xi\). Since \(P(\eta) > 0\), it follows that, every \(\eta\)-world with positive probability is also a model of \(\xi\). Therefore, when \(P(\omega) > 0\) and \(\omega(\eta) = 1\), we have \(\omega(\xi) = 1\), that is, \(\omega^\# = \omega\). Hence, Equation 3.6 reduces to,
\[
P^* (\alpha) = \sum_{\alpha} P(\omega) \cdot P(\omega (\alpha) \cdot P(\eta)) + P(\alpha \land \neg \eta) = P(\alpha \land \eta) = P(\alpha),
\]
for every \(\alpha \in L\). This shows that selective imaging satisfies \(E_2\).

**Proof for \(E_3\):** Let us suppose that \(\xi\) is a sentence such that \(\neg \xi\) is a belief, that is, \(P(\neg \xi) = 1\). Also assume \(\not \models \neg \xi\). Since \(P(\eta) > 0\) there exists a model of \(\eta\), say \(\omega\), which has non-zero probability. Therefore \(P(\neg \xi \land \eta) > 0\). From Equation 3.6b we find that \(P^* (\neg \xi) = P(\neg \xi \land \neg \eta) < 1\). Hence selective imaging satisfies \(E_3\).

**Proof for \(E_4\):** Let \(\gamma\) be a sentence such that \(\vdash \neg \xi \iff \neg \gamma\). This gives \([\xi] = [\gamma]\), therefore for any possible world \(\omega, \omega^\# = \omega^\#\). Hence from Equation 3.6 we get,
\[
P^* (\alpha) = \sum_{\alpha} P(\omega) \cdot \omega^\# (\alpha) \cdot P(\eta) + P(\alpha \land \neg \eta)
\]
\[
= \sum_{\alpha} P(\omega) \cdot \omega^\# (\alpha) \cdot P(\eta) + P(\alpha \land \neg \eta) = P^* (\alpha),
\]
for every sentence \(\alpha \in L\).

Thus selective imaging satisfies \(E_4\).

**Observation 3.6** When \(P(\eta) = 1\), \(P^*_{\xi} = P^*_{\gamma}\).

**Proof:** Given a probability function \(P\) over \(L\), let \(S\) be a selection function. Suppose the
sentence $\eta$ associated with $S$ such that $P(\eta) = 1$. Also let $a$ be a real number between 0 and 1. We denote the pair of sentence $\xi$ and real number $a$ by $I = \langle \xi, a \rangle$. We need to show that given a sentence $\xi$, $P^{sp}_I(\alpha) = P^s_I(\alpha)$ for every sentence $\alpha \in L$.

From Equation 3.7 we have

$$P^{sp}_I(\alpha) = a \cdot \sum_{\omega} P(\omega) \cdot \omega^\#(\alpha) \cdot \omega(\eta) + (1-a) \cdot P(\alpha \land \eta) + P(\alpha \land \neg \eta).$$

When $P(\eta) = 1$, we have $P(\neg \eta) = 0$, which gives us $P(\alpha \land \neg \eta) = 0$ and $P(\alpha \land \eta) = P(\alpha)$. Therefore, Equation 3.7 reduces to,

$$P^{sp}_I(\alpha) = a \cdot \sum_{\omega} P(\omega) \cdot \omega^\#(\alpha) \cdot \omega(\eta) + (1-a) \cdot P(\alpha).$$

That is, $P^{sp}_I(\alpha) = a \cdot P^\#_I(\alpha) + (1-a) \cdot P(\alpha)$, that is, $P^{sp}_I(\alpha) = P^s_I(\alpha)$.

**Observation 3.7** When $a = 1$, $P^{sp}_I = P^s_I$.

**Proof:** Let $P$ be a probability function and $\xi$ be a sentence in $L$. Suppose that $S$ be a selection function and $\eta$ be the associated sentence. We need to show that when the variable $a$ takes the value 1, we get $P^{sp}_{\xi,1} = P^s_\xi$. We use $I$ to denote the pair of sentence and the real number $a = 1$, $I = \langle \xi, 1 \rangle$.

Equation 3.7 gives, for any sentence $\alpha \in L$,

$$P^{sp}_{\xi,a}(\alpha) = a \cdot \sum_{\omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) + (1-a) \cdot P(\alpha \land \eta) + P(\alpha \land \neg \eta).$$

Since $a = 1$, Equation 3.7 reduces to $P^{sp}_I(\alpha) = \sum_{\omega} P(\omega) \cdot \omega^\#_\xi(\alpha) \cdot \omega(\eta) + P(\alpha \land \neg \eta)$. From Equation 3.6, we get $P^{sp}_I(\alpha) = P^s_I(\alpha)$.

**Observation 3.8** If $P(\eta) = 1$ and $a = 1$, then $P^{sp}_I = P^\#_I$.

**Proof:** This result follows from Observations 3.2 and 3.7.

**Theorem 3.3** Let the selection function $S$ be such that the associated sentence $\eta$ has a non-zero probability and $0 < a < 1$. Then the selective partial imaging function is a belief removal function that satisfies **E1, E2, E3 and E4**.

**Proof:** Given a probability function $P$ which represents the belief state of the agent, let $S$ denote the selection function. There exists a sentence $\eta \in L$ such that $[\eta] = S(P)$. Let $a$ be a real number which takes a value from $(0, 1)$ and $I = \langle \xi, a \rangle$.

**Proof for E1:** Now suppose that $\xi$ is a sentence such that $P(\neg \xi) = 1$, that is, $\neg \xi$ is a belief of the agent. Let $\alpha$ be a sentence in $L$ such that $P^{sp}_I(\alpha) = 1$, that is, every possible
sets corresponding to $P_\xi$ be a sentence such that $\exists \xi, a$ sponding sentence $\eta$. Assume a probability function $\omega$ and $K$ a given probability function is given by $\alpha$ and $\omega$ for every world $\omega$, then $\omega \in [\alpha]$ and hence $P(\alpha) = 1$. This shows that SPI satisfies $E1$.

Proof for $E2$: Suppose $\xi$ be a sentence in $\mathcal{L}$ such that $P(\xi) = 1$. From our hypothesis, we get $P(\eta) > 0$. This gives us that every model of $\eta$ with positive probability is a model of $\xi$. Therefore when $\omega(\eta) = 1$, we have $\omega_\xi^\# = \omega$. Hence, if $P(\omega) \cdot \omega_\xi^\#(\alpha) \cdot \omega(\eta) \neq 0$ for any sentence $\alpha \in \mathcal{L}$ then we can re-write it as $P(\omega) \cdot \omega(\alpha \land \eta)$. Hence from Equation 3.7 we have $P_\xi(\alpha) = a \cdot P(\alpha \land \eta) + (1-a) \cdot P(\alpha \land \eta) = P(\alpha \land \eta)$, that is, $P_\xi(\alpha) = P(\alpha \land \eta) + P(\alpha \land \neg \eta) = P(\alpha)$, for every sentence $\alpha \in \mathcal{L}$. This shows SPI satisfies $E2$.

Proof for $E3$: Since $P(\neg \xi) = 1$ and $P(\eta) > 0$, every model of $\eta$ with positive probability is also a model of $\neg \xi$. Hence $P(\neg \xi \land \eta) > 0$, and from Equation 3.7a we get $P_\xi(\neg \xi) = P(\neg \xi) - a \cdot P(\eta \land \neg \xi)$. Therefore $P_\xi(\neg \xi) < 1$, showing that SPI satisfies $E3$.

Proof for $E4$: Now let $\gamma$ be a sentence such that $\vdash \neg \xi \leftrightarrow \neg \gamma$, then $[\xi] = [\gamma]$. Therefore for every world $\omega$, $\omega_{\xi}^\# = \omega_{\gamma}^\#$. It follows from Equation 3.7 that $P_{\xi,a}(\alpha) = P_{\gamma,a}(\alpha)$, for every $\alpha \in \mathcal{L}$. Hence $E4$ holds.

Theorem 3.4 Assume a probability function $P$, a selection function $S$ and the corresponding sentence $\eta$ such that $P(\eta) > 0$. Suppose $K_\xi$, $K_\xi^s$ and $K_\xi^{sp}$ represent the belief sets corresponding to $P_\xi$, $P_\xi^s$ and $P_\xi^{sp}$, respectively. Then, $K_\xi^p \subseteq K_\xi^{sp} \subseteq K_\xi^s$ where $\mathcal{I} = \langle \xi, a \rangle$ and $0 < a < 1$.

Proof: Let $P$ represent the belief state of the agent and $S$ denote the selection function. There exists a sentence $\eta$ such that $[\eta] = S(P)$. Suppose $P(\eta) > 0$ and $0 < a < 1$. Let $\xi$ be a sentence such that $P(\neg \xi) = 1$ and $\mathcal{I} = \langle \xi, a \rangle$. The belief set $K$ corresponding to a given probability function is given by $K = \{ \alpha \in \mathcal{L} \mid \|P\| \leq [\alpha] \}$, that is, $[K] = \|P\|$. 

142
Therefore, in order to show that $K^s_\zeta \subseteq K^{sp}_I \subseteq K^p_I$, it is sufficient to show that
\[ \|P^s_\zeta\| \subseteq \|P^{sp}_I\| \subseteq \|P^p_I\|. \]

Let $\mu \in \|P^s_\zeta\|$, then we have 2 cases: either (a) $\mu \in \|P\|$ or (b) $\mu = \omega^\#_\zeta$ for some $\omega \in \|P\|$ and $\omega \in S(P)$.

Case (a) Suppose that $\mu \in \|P\|$. Then either $\mu \models \neg \eta$ or $\mu \models \eta \land \xi$.

When $\mu \models \neg \eta$, Equation 3.7 gives us $P^{sp}_I(\mu) = a \cdot \sum \omega P(\omega) \cdot \omega^\#_\zeta(\mu) \cdot \omega(\eta) + (1 - a) \cdot P(\mu \land \eta) + P(\neg \eta \land \mu)$. Therefore $P^{sp}_I(\mu) = a \cdot \sum \omega P(\omega) \cdot \omega^\#_\zeta(\mu) \cdot \omega(\eta) + P(\mu)$, which gives $P^{sp}_I(\mu) > 0$, that is, $\mu \in \|P^{sp}_I\|$.

When $\mu \models \eta \land \xi$, $P^{sp}_I(\mu) = a \cdot \sum \omega P(\omega) \cdot \omega^\#_\zeta(\mu) \cdot \omega(\eta) + (1 - a) \cdot P(\mu \land \eta) + P(\neg \eta \land \mu)$, that is, $P^{sp}_I(\mu) \geq P(\mu)$. Therefore, $\mu \in \|P^{sp}_I\|$. Therefore $\mu \in \|P^p_I\|$.

Case (b) Suppose that $\mu = \omega^\#_\zeta$ for some $\omega \in \|P\|$. Then $P^{sp}_I(\mu) = \sum \omega \in S(P) \{a \cdot P(\omega) : \omega^\#_\zeta = \mu\}$, therefore $\mu \in \|P^{sp}_I\|$. Therefore $\|P^s_\zeta\| \subseteq \|P^{sp}_I\|$.

In a similar fashion we can show that $\|P^{sp}_I\| \subseteq \|P^p_I\|$.  

---

1Here we distort the meaning of $\omega^\#_\zeta(\alpha)$ to mean $\omega^\#_\zeta(\mu) = 1$ iff $\mu = \omega^\#_\zeta$. 

143
Appendix C

Proofs for Chapter 4

Theorem 4.1  Indifferent contraction satisfies all the postulates of contraction C1 to C5 and CO.

Proof: Let $P$ be the probability distribution that represents the agent’s initial belief state. Let $\alpha$ be a belief of the agent, that is, $P(\alpha) = 1$. We further assume that $\alpha$ is a non-trivial belief, that is, $\forall \alpha$. The result of indifferent contraction of $P$ by $\alpha$ is denoted by $P^-_{\alpha}$.

Let the cardinality of $\Omega$, $[\alpha]$, $[-\alpha]$ and the support of $P$ be $n$, $l$, $m$, and $l'$ respectively. Suppose the initial distribution is as follows:

$$P = \langle p_1, \ldots, p_l, 0, \ldots, 0 \rangle.$$  

The result of indifferent contraction is then given by the following:

$$P^-_{\alpha} = \langle dp_1, \ldots, dp_l, \frac{1}{h + m}, \ldots, \frac{1}{h + m} \rangle,$$

where $h = 2^H(P)$ and $d = \frac{n}{h + m}$. From Definition 4.1 it follows that indifferent contraction satisfies the postulate C1. Clearly $0 < d < 1$. For every $\omega \in [\alpha]$, we have $P^-_{\alpha}(\omega) = dP(\omega)$. The resultant probability of $\alpha$ is given by, $P^-_{\alpha}(\alpha) = \sum_{\omega \in [\alpha]} dP(\omega) = d$. Thus $P^-_{\alpha}(\alpha) < 1$, satisfying postulate C2.

When $\vdash \alpha \leftrightarrow \beta$, then we have $[\alpha] = [\beta]$ and $[-\alpha] = [-\beta]$. Indifferent contraction of $P$ by $\alpha$ is such that $P^-_{\alpha}(\omega) = dP(\omega)$, when $\omega \in [\alpha]$ and similarly, indifferent contraction of $P$ by $\beta$ is such that $P^-_{\beta}(\omega) = dP(\omega)$, when $\omega \in [\beta]$, where $d = \frac{h}{h + m}$, $m$ is the cardinality.
of $[-\alpha]$ as well as that of $[-\beta]$. Thus we have $P_\alpha^-(\omega) = P_\beta^-(\omega)$, when $\omega \in [\alpha]$ (or when $\omega \in [\beta]$). Also, $P_\alpha^- (\omega) = \frac{1}{h+m} = P_\beta^- (\omega)$, for $\omega \in [-\alpha]$ (or $\omega \in [-\beta]$). Thus we have $P_\alpha^- (\omega) = P_\beta^- (\omega)$, for every $\omega \in \Omega$, hence $P_\alpha^- = P_\beta^-$, satisfying the postulate C3.

The postulate C4 is vacuously satisfied, since indifferent contraction is defined only for a belief $\alpha$, that is, only when $P(\alpha) = 1$. When $P(\alpha) < 1$, then indifferent contraction does not change the initial distribution. Now let us consider C5. Suppose $P(\alpha) = 1$, by Definition 4.1, conditionalizing $P_\alpha^-$ by $\alpha$ gives $P$. Since expansion is modelled by Bayesian conditionalization, we have $(P_\alpha^-)_\alpha^+ = P$, indifferent contraction satisfies the postulate C5.

The entropy of $P_\alpha^-$, which is given by indifferent contraction, is $\log(2^{H(P)} + m)$. Since log is a monotonic function, $\log(2^{H(P)}) < \log(2^{H(P)} + m)$, where $m$ is the cardinality of $[-\alpha]$ and is a positive integer (since we have assumed that $\alpha$ is not a theorem). Thus $H(P_\alpha^-) \geq H(P)$, satisfying the postulate CO.

**Theorem 4.2** An indifferent contraction is a full meet contraction.

**Proof:** Let $P$ be the initial distribution representing the agent’s belief state and $\mathcal{K}$ be the corresponding set of beliefs, given by Equation 4.1. Let $\alpha$ be a belief of the agent, that is, $P(\alpha) = 1$. The result of indifferent contraction of $P$ by $\alpha$ is denoted by $P_\alpha^-$ and the corresponding belief set is denoted by $\mathcal{K}_\alpha^-$. A probabilistic contraction function is said to be a full meet contraction [29] when it satisfies the postulates C1 to C5 and the associated contracted belief set $\mathcal{K}_\alpha^-$ is such that $[\mathcal{K}_\alpha^-] = [\mathcal{K}] \cup [-\alpha]$.

Theorem 4.1 shows that indifferent contraction satisfies postulates C1 to C5. From Equation 4.1, we get that $\mathcal{K}_\alpha^- = \{ \beta \in \mathcal{L} : P_\alpha^-(\beta) = 1 \}$, that is, $\mathcal{K}_\alpha^- = \{ \beta \in \mathcal{L} : \| P_\alpha^- \| \subseteq [\beta] \}$. Since, as a result of indifferent contraction, every $-\alpha$-world is assigned positive probability, we have $[-\alpha] \subseteq \| P_\alpha^- \|$. Since $\| P_\alpha^- \| = \| P \| \cup [-\alpha]$, it follows that $[\mathcal{K}_\alpha^-] = [\mathcal{K}] \cup [-\alpha]$. Hence indifferent contraction is a full meet contraction.
**Theorem 4.3** A submaximal entropy contraction is a full meet contraction.

**Proof:** Let $P$ be the initial distribution representing the belief state of the agent and $\alpha$ be a belief, that is $P(\alpha) = 1$. Submaximal entropy contraction results in reducing the probability associated with the $\alpha$-worlds proportionally and all the $\neg\alpha$-worlds are assigned equal non-zero probability.

A probabilistic contraction function is said to be a full meet contraction [29] when it satisfies the postulates C1 to C5 and the associated contracted belief set $\mathcal{K}_\alpha^-$ is such that $[\mathcal{K}_\alpha^-] = [\mathcal{K}] \cup [\neg\alpha]$.

Submaximal entropy contraction satisfies postulates C1 and C5 by definition (Definition 4.2). The postulate C4 is also satisfied vacuously, since it is defined only for the case $P(\alpha) = 1$. The contracted probability distribution, $P^-_\alpha$ is a uniform distribution over the set $[\neg\alpha]$ and assigns non-zero probability to every $\neg\alpha$-world. Thus $P^-_\alpha(\alpha) < 1$, satisfying postulate C2.

When $\vdash \alpha \leftrightarrow \beta$, then $[\alpha] = [\beta]$ and $[\neg\alpha] = [\neg\beta]$. As a result of contraction by $\alpha$, probability associated with every $\alpha$-world is scaled by a non-zero value $d$, determined by the minimization problem and the mass $1 - d$ is spread uniformly over $[\neg\alpha]$. The same happens when contracting by $\beta$. Hence submaximal entropy contraction satisfies the postulate C3.

The support of contracted probability distribution $P^-_\alpha$ is such that $\|P^-_\alpha\| = \|P\| \cup [\neg\alpha]$. This is the same as in the case of indifferent contraction. Hence along same lines as Theorem 4.2, submaximal entropy contraction is a full meet contraction. ■

**Theorem 4.4** A minimal contraction is a “syntax-sensitive” maxichoice contraction.

**Proof:** Let $P$ be the initial distribution representing the agent’s belief state and $\mathcal{K}$ be the corresponding set of beliefs, given by Equation 4.1. Let $\alpha$ be a belief of the agent, that is, $P(\alpha) = 1$. The result of minimal contraction of $P$ by $\alpha$ is denoted by $P^-_\alpha$ and the corresponding belief set is denoted by $\mathcal{K}_\alpha^-$. A probabilistic contraction function is said to be a “syntax-sensitive” maxichoice con-
traction when it satisfies the postulates C1, C2, C4, C5 and the associated contracted belief set $\mathcal{K}_{\alpha}$ is such that $[\mathcal{K}_{\alpha}] = [\mathcal{K} \cup \{\omega\}$ for some $\omega \in [\neg \alpha]$.

Minimal contraction results in a probability distribution where the probabilities associated with the $\alpha$-worlds are reduced proportionally to a very small value. Minimal contraction, we assume, is equipped with some mechanism to choose a particular world from the set $[\neg \alpha]$. The chosen world is assigned non-zero probability which is close to 1. From Definition 4.3 it follows that minimal contraction satisfies the postulate C1. For every $\omega \in [\alpha]$, we have $P_\alpha^-(\omega) = dP(\omega)$, where $d$ is very small. The resultant probability of $\alpha$ is given by, $P_\alpha^-(\alpha) = \sum_{\omega \in [\alpha]} dP(\omega) = d$. Thus $P_\alpha^-(\alpha) < 1$, satisfying postulate C2.

The postulate C4 is vacuously satisfied, since minimal contraction is defined on when $P(\alpha) = 1$. By Definition 4.3, minimal contraction satisfies the postulate C5.

As a result of minimal contraction, we have $\|P_\alpha^-\| = \|P\| \cup \{\omega\}$, where $\omega \in [\neg \alpha]$. From Equation 4.1, it is clear that $[\mathcal{K}_\alpha^-] = [\mathcal{K} \cup \{\omega\}$ for some $\omega \in [\neg \alpha]$, as required. ■

**Theorem 4.5** A preferential contraction function is a partial meet contraction function.

**Proof:** Let $P$ be the initial distribution representing the agent’s belief state and $\mathcal{K}$ be the corresponding set of beliefs, given by Equation 4.1. Let $\alpha$ be a belief of the agent, that is, $P(\alpha) = 1$. The result of preferential contraction of $P$ by $\alpha$ is denoted by $P_\alpha^-$ and the corresponding belief set is denoted by $\mathcal{K}_\alpha^-$. Let $\sqsubseteq$ be a total preorder relation on $\Omega$. We need to show that preferential contraction satisfies the postulates C1 to C5.

Let the cardinality of $\Omega$, $[\alpha]$, $[\neg \alpha]$ and the support of $P$ be $n$, $l$, $m$, and $l'$ respectively. Suppose the initial distribution is as follows:

$$P = \langle p_1, \ldots, p_l, 0, \ldots, 0 \rangle^{m}.$$

Let the cardinality of $[\neg \alpha]$ and $\min_{\sqsubseteq}[\neg \alpha]$ be $m$ and $m'$ respectively. The probability distribution resulting from preferential contraction is as follows:

$$P_\alpha^- = \langle dp_1, \ldots, dp_l, \frac{1}{h + m'}, \ldots, \frac{1}{h + m'}, 0, \ldots, 0 \rangle^{m'}.$$  

where $h = 2^{H(P)}$ and $d = \frac{h}{h + m'}$. From Definition 4.4, it is clear that preferential contrac-
tion satisfies $C_1$. The resultant probability of $\alpha$ is $P^-_\alpha(\alpha) = dP(\alpha) = d$, which is less than 1. Hence preferential contraction satisfies $C_2$.

When $\vdash \alpha \leftrightarrow \beta$, then $[\alpha] = [\beta]$ and $[\neg \alpha] = [\neg \beta]$. Therefore we have $\min\subseteq[\neg \alpha] = \min\subseteq[\neg \beta]$. We have

$$P^-_\alpha(\omega) = P^-_\beta(\omega) = \begin{cases} 
    dP(\omega), & \text{if } \omega \in [\alpha] = [\beta] \\
    \frac{1}{h+m'}, & \text{if } \omega \in \min\subseteq[\neg \alpha] = \min\subseteq[\neg \beta] \\
    0, & \text{otherwise.}
\end{cases}$$

Thus preferential contraction satisfies $C_3$. The postulate $C_4$ is vacuously satisfied, since preferential contraction is only defined when $P(\alpha) = 1$ and postulate $C_5$ is satisfied evidently by definition. Thus, preferential contraction is a partial meet contraction.  

$\blacksquare$
Bibliography


[71] Abhaya C. Nayak, Randy Goebel, Mehmet A. Orgun, and Tam Pham. Iterated be-


