Belief Dynamics

ILLC Dissertation Series 2001-08

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

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

Belief Dynamics

(Epistemo)logical Investigations

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. J.J.M. Franse ten overstaan
van een door het college voor promoties
ingestelde commissie, in het openbaar
te verdedigen in de Aula der Universiteit
op dinsdag 13 november 2001, te 10.00 uur

door

Allard Martijn Tamminga

geboren te Nijmegen

Promotor: Prof. dr. F.J.M.M. Veltman
Co-promotor: Prof. dr. H. Rott
Afdeling Wijsbegeerte
Faculteit der Geesteswetenschappen
Universiteit van Amsterdam
Nieuwe Doelenstraat 15
1012 CP Amsterdam

Cover design by [lowresmedia.com]
Cover photos by Roesmann/Essamba Arts
Printed and bound by Febodruk, Enschede

ISBN: 90–5776–075–4

Notre tête est ronde pour permettre à la pensée de changer de direction.

Francis Picabia

Contents

Acknowledgments

First and foremost, thanks are due to my promotor, Frank Veltman, and to my co-promotor, Hans Rott. I thank them for the freedom of research they granted me, for their continuing encouragement, and for their constructive criticisms.

Second, I wish to thank Johan van Benthem, Wiebe van der Hoek, Erik Krabbe, Michiel van Lambalgen, and Martin Stokhof, for their willingness to participate in the reading committee.

Third, I thank Koji Tanaka for being the co-author of the Appendix. Moreover, I thank Walter van der Star for translating the second and third chapters of my dissertation into English.

Lastly, the following people discussed parts of the thesis with me and thus helped me to improve on the final text: Paul Dekker, Eduardo Fermé, Jelle Gerbrandy, Carin Gotlieb, Paul Harrenstein, Wolfram Hinzen, Else de Jonge, Theo Kuipers, Karen Kwast, Rosja Mastop, Breanndán Ó Nualláin, Jeanne Peijnenburg, Graham Priest, Harry Stein, Anne Troelstra, and Sjoerd Zwart. The conversations with them has made the writing of this thesis an enjoyment.

Parts of my thesis have been published as articles in journals. I would like to thank the journals for their kind permission to include these articles in the present thesis. Chapter 2 appeared in Dutch as 'Waarheid als limiet van onderzoek: de ontoereikendheid van Charles Sanders Peirce' waarheidsopvatting' in the *Algemeen Nederlands Tijdschrift voor Wijsbegeerte* **93**, 2001, pages 73–92. The Appendix, the result of joint work with Koji Tanaka, appeared in 2001 as 'A Natural Deduction System for First Degree Entailment' in the *Notre Dame Journal of Formal Logic* **40**, 1999, pages 258–272.

Amsterdam AMT
September, 2001.

Chapter 1

Introduction

1.1 General Overview

This thesis deals with the dynamics of belief. In three interconnected, though independent studies, a viable alternative to foundational approaches in epistemology is explored. Rather than accept the research agenda dictated by traditional epistemology, with its stress on the pursuit of conditions under which our beliefs are justified or even true, the present work follows the lead of American pragmatism, and focusses on the articulation and defense of criteria according to which a change of mind may be judged legitimate.

The second and third chapters discuss the attempts of two pragmatist philosophers, Charles Sanders Peirce and Isaac Levi, to formulate and to formalize the *belief-doubt-belief* model, a model that sets out to set forth systematically the optimal strategies according to which to change our minds.

In the course of this undertaking, more and more logical considerations come to the fore, culminating in a discussion and an evaluation of formal systems for belief change. It is argued that belief change systems, despite the naturalistic setting in which they were first conceived, have outgrown their pragmatist origin and have metamorphosed into a branch of philosophical logic, where empirical considerations have become obsolete. To restore the connection between logical theory and epistemological practice, a case is made for reformulations of belief change systems that are cleared from elements that obstruct practical applications and empirical tests.

In the fourth chapter, a belief change system is presented that (1) uses finite representations of epistemic states, (2) can deal with inconsistencies adequately, (3) has finite operations of change, (4) can do without, but does not rule out, extra-logical elements, and (5) only licenses consistent beliefs. An Appendix, where the underlying logic of the belief change system of the fourth chapter is studied proof-theoretically, concludes the thesis.

1.2 Summary

In order to pave the way for more fruitful logical and epistemological investigations within the pragmatist tradition in philosophy, the second chapter clears the ground from metaphysical entanglements that result from the wish to construct theories of truth and reality, a wish that has, by and large, dominated the interests of pragmatist philosophers (and not only them) until the present day. A focus on the necessary and sufficient conditions for a statement to be true or on a conception of reality which is adequate to a preferred reading of the truth predicate still obfuscate research in contemporary epistemology, as they tend to restrict philosophical research to questions which have resisted plausible answers for ages.

Hence, the second chapter purports to show that Peirce's metaphysical aspirations for definitions of truth and reality in terms of a theory of inquiry cannot coherently be met by a pragmatist theory of knowledge. In order to come to grips with Peirce's definition of truth as the ideal limit of inquiry, a succinct exposition of Peirce's theory of inquiry and his philosophical logic is given. Attention is paid to deductive, inductive, and abductive reasoning and their interrelations, so as to be in a position to sketch Peirce's early ideas concerning the pragmatist *belief-doubt-belief* model, a model that will be the guiding line of this thesis. With his investigations into that model, Peirce aimed to articulate the 'method of science', a method that enables us to adapt our present pattern of expectations for the better in case of surprising or even contrary experience. Peirce maintained that if we would consciously follow the method of science, we would ultimately reach a pattern of expectations that would never be thwarted by any future experience. Since such a pattern of expectations would lead to the very same practical consequences as a *true* 'theory of the world', to borrow a phrase of Quine's, Peirce's pragmatist stand on meaning forbids him to withhold the predicate 'true' from that pattern of expectations. Thus Peirce accounts for the notion of truth.

In the remainder of the second chapter, the arguments of the foremost representative of a group of contemporary apologists of Peirce's account of truth, Cheryl Misak, are subjected to scrutiny and found to be insufficient: the principle of bivalence is defended improperly and her main argumentation for a Peircean account of truth turns out to be incompatible with a pragmatist epistemology. Hence, Peircean accounts of truth are found to be inadequate. Rather than accept the problem setting of traditional philosophy, with its focus on truth and reality, and plead, in the face of the failure of Peircean accounts of truth, for some alternative view on truth, the chapter ends with a plea for a more modest approach. It is proposed to retain as much as possible of Peirce's dynamic theory of scientific method, that is, his *belief-doubt-belief* model, and to focus, unencumbered by the traditional philosophical quest for truth, on the technicalities of contemporary accounts of belief change that are the proper heirs of Peirce's *belief-doubt-belief* model.

The third chapter critically studies the logical epistemology of Isaac Levi, the American pragmatist philosopher who took the lead in defining the field in philosophical logic which has become known under the name *belief change*. Levi sets out to formalize Peirce's and Dewey's *belief-doubt-belief* model, using classical logic, probability theory, and decision theory.

As an introduction, Levi's fundamental tenets are placed within the context of American pragmatism. As we saw in the second chapter, Peirce kept the 19th century English school of logic's theoretical interest in non-deductive forms of reasoning alive, and, apart from that, was interested in a systematical theory with which rational changes of belief could be captured. Levi shares Peirce's predilections, but he is, unlike Peirce, not interested in furnishing an account of truth based on his pragmatist epistemology. Alternatively, Levi sets out to formulate and to defend criteria under which a change of belief must be judged an improvement.

In order to do so, Levi first needs a device to represent epistemic states. He opts for *corpora*, that is, sets of sentences closed under logical deduction. Two historical reasons and one pragmatical reason for this choice are set forth. Since Levi has chosen corpora to represent epistemic states, he is able to interpret changes of epistemic states in terms of transitions from one corpus to another. One of the central claims of Levi's work on belief change is that all transitions of corpora can be reduced to two basic types of transition: expansion and contraction.

In *expansion*, a sentence is added to an agent's present corpus. Levi defends, using arguments from decision theory, conditions under which such an addition is legitimate. His criterion stipulates whether a proposed expansion is legitimate by weighing the informational value of the sentence to be added against its plausibility. Moreover, Levi discusses the logic of expanding a corpus by a sentence, once the expansion is judged legitimate. Next, an integrated perspective on *contraction* is developed. In contraction, a sentence is skipped from an agent's present corpus. Parallel to his account of expansion, Levi proffers criteria under which a contraction is legitimate or even, in a special case, required. Unlike the logic of expansion, the corpus that results from a contraction is not immediately given, once it has been judged legitimate to contract our current corpus with some sentence. Several corpora, all of them subsets of the corpus to be contracted, fulfill the logical requirements for the envisaged contraction. Levi advises to adopt the corpus that not only fulfills these requirements, but minimizes the loss of informational value as well. An application of Levi's theory of belief change to two central problems in philosophical logic, to wit, modal and conditional statements, completes the exposition of his theory of knowledge.

In the second part of the third chapter, Levi's logical epistemology is critically evaluated. First, we discuss Levi's claim that an epistemic state can be represented by a corpus. It turns out to be impracticable to decide which sentences are part of the corpus and which are not. Second, Levi's assumption that a system of parameters has been given at the outset is examined. As both his account of expansion and his account of contraction are based on a previous assessment of the informational value parameter, we focus on the question of how such an assessment is to be put into practice. A similar strategy is used in the criticism of the measure of boldness, a parameter that plays a role in judging envisaged expansions. Surprisingly, the desired resulting corpus is the clue to assess the values of the parameters.

As a consequence, contrary to Levi's repeated claim that his criteria for changes of belief are normative, his theory lacks normative force, at least as long as it remains to be seen how the initial values of the parameters figuring in his system can be assessed in a reliable and convincing way. Moreover, even if Levi's system is interpreted as a description of actual reasoning practices, as long as the parameters withstand an assessment which is independent of the result of a change of belief, Levi's theories cannot be put to an empirical test. In conclusion, a case is made for a theory of belief change that can do without extra-logical considerations which elude assessment, so as, on the one hand, to avoid the rationalistic mode of argument that has been typical of contemporary belief change literature, and, on the other hand, to pave the way for investigations into the empirical adequacy of proposed theories of belief change.

The fourth chapter expounds a belief change theory that, on the one hand, unlike standard systems for belief change, can do without extra-logical considerations, and, on the other hand, is able to cope with inconsistent information. Moreover, the beliefs that are licensed by the information offered always are consistent. Hence, the theory has two levels: a basic level dealing with the dynamics of possibly inconsistent information, and an upper level that extracts a set of consistent beliefs from the basic level.

On the basic level, first, a method is set forth, using first degree entailment as underlying logic, to represent possibly inconsistent information by a *finite state*, which consists of exactly all valuations that are minimally required to validate the information offered. An algorithm to find a formula's finite state is given. Next, an operation for expanding a finite state by a formula is defined. A representation theorem, showing that the expansion operation satisfies the intuitive requirements for expansion, provides a characterization of the expansion operation. A contraction operation, by which a formula can be deleted from a finite state, follows. Interestingly, the contraction operation does not presuppose – the presupposition is ubiquitous in theories of belief change that have been motivated by the works of Levi and of Alchourrón, Gärdenfors and Makinson – an extra-logical element such as a choice function or an ordering of (sets of) sentences. Moreover, the expansion and the contraction operation are each other's duals. A representation theorem characterizing the contraction operation in terms of a set of postulates completes the discussion of contraction. Remarkably, the postulates have been formulated in terms of the *information span* of a finite state, a property of finite states that has no counterpart in standard possible world semantics.

On the upper level, four *extractors*, operations retrieving a set of consistent 'plausible' beliefs from a finite state, are studied. In cases where the finite state is consistent, all extractors yield the same result. If, however, the finite state is inconsistent, the resulting sets of 'plausible' beliefs generally do not coincide.

Two basic extractors are discussed, the first a translation of a proposal of Restall and Slaney's to the present context of finite states, the second amounting to a contraction of the finite state to which the extractor was to be applied, with a formula indicating which literals behave inconsistently in that finite state. As the extractors give rise to non-monotonic inference operations, some proof-theoretical properties of the two basic extractors are discussed. Next, a selection function that chooses the most consistent elements from a finite state is set forth. This function can be used to preprocess the finite state, before applying one of the two basic extractors. Hence, four different extractors have been defined. Finally, the relative strength of the four extractors is assessed.

The Appendix, the result of a collaboration with Koji Tanaka from Macquarie University in Sidney, consists of an original proof-theoretical study of the paraconsistent and relevant logic first degree entailment, fde for short, which was the underlying logic of the system of belief change that was propounded in the fourth chapter. In a sense, this logic is a generalization of classical propositional logic (cpl), since it does not only consider total and consistent valuations as in cpl, but partial or inconsistent valuations as well. Hence, all inferences that are valid in fde are valid in cpl. The converse does not hold: for instance, fde does not have any tautologies and neither the disjunctive syllogism, nor the *ex falso quodlibet* rule of cpl hold true in fde.

The Appendix opens with a brief history of fde. Two logical semantics have been proposed for fde, a two-valued one by Routley and Routley and a four-valued one by Dunn. Here, we adopt the latter, four-valued semantics. Then, 'combined systems', proof-theories pioneered by Łukasiewicz, are discussed briefly. In combined systems both accepted and rejected formulas can be derived, distinguished in the derivation as they are preceded by a sign ('$\vdash$' for acceptance, '$\dashv$' for rejection) which indicates their status. This idea will be used in developing the natural deduction system for fde. Next, after presenting the language and the four-valued semantics of fde, a Gentzen-style natural deduction system for fde is presented. Soundness is proved more or less standardly. Completeness is proved via Henkin's method, though the construction used in the proof of the embedding lemma had to be adapted to meet the requirements of the subsequent model-existence lemma, where a four-valued model is required.

Chapter 2

Truth as Limit of Inquiry?

2.1 Introduction

In 1913, in his thesis *De kennisleer van het Anglo-Amerikaansch pragmatisme* (*The Epistemology of Anglo-American Pragmatism*), Tobias Muller puts forward a bold proposition that, although the arguments presented by Muller are now outdated, surely deserves to be reconsidered: "Pragmatism is essentially a reformation of logic". In this chapter I will set out which reforms of traditional logic were advocated by the American pragmatist Charles Sanders Peirce, how he attempted to explain the concept of truth by means of his logic, and why the latter undertaking is doomed to fail.

In the second half of the 19th century, both the foundations of traditional logic and those of the traditional ideal of knowledge, that had been laid down by Aristotle in his *Analytica priora* and his *Analytica posteriora*, were beginning to crack as a result of the constant erosion caused by new developments in the sciences. It became clear that modern mathematical reasoning could hardly be represented by the limited range of forms of reasoning provided by Aristotelian syllogistic, while the evidence postulate of Aristotle's ideal of knowledge was being undermined by such developments as the discovery of non-Euclidian geometries. The demise of the Aristotelian ideal of knowledge provoked a lively debate on the 'crisis of science', the 'crisis of certitudes', or even a 'crisis of culture', a discussion which was not restricted to a single philosophical school. To a large extent, early twentieth century philosophy consisted of an attempt to define its own position in the light of these crises in an intellectually sound (or unsound!) way.

In order to get some sort of overall view of the purpose and extent of the reform of traditional logic that Peirce had in mind, it is essential to make an excursion into the Aristotelian ideal of knowledge. Aristotle's own position is not adequately represented when it is understood as an "axiomatisation of scientific method", because it describes "the foundations of, and accordingly the requirements for, irrefutable knowledge" (*epistêmê*),[1] independently of the way in which this knowledge is acquired. It should be seen as an ideal of necessarily true and unshakable knowledge, an ideal that is compatible with different theories about the way in which that knowledge is attained. Evert Willem Beth gives the following description of that which he – somewhat misleadingly – calls "Aristotle's theory of science" [Beth, 1959, p. 31–32]:

> A *deductive science* is a system S of sentences, which satisfies the following postulates:
>
> (I) Any sentence belonging to S must refer to a specific domain of real entities;
>
> (II) Any sentence belonging to S must be true;
>
> (III) If certain sentences belong to S, any logical consequence of these sentences must belong to S;
>
> (IV) There are in S a (finite) number of terms, such that
>
> (a) the meaning of these terms is so obvious as to require no further explanation;
>
> (b) any other term occurring in S is definable by means of these terms;
>
> (V) There are in S a (finite) number of sentences, such that
>
> (a) the truth of these sentences is so obvious as to require no further proof;
>
> (b) the truth of any other sentence belonging to S may be established by logical inference starting from these sentences.
>
> The postulates (I), (II), and (III) will be called, respectively, the *reality*, the *truth*, and the *deductivity postulate*. The postulates (IV) and (V) together constitute the so-called *evidence postulate*; the fundamental terms and sentences, referred to in postulates (IV) and (V), are called the *principles* of the science under consideration.

For centuries, Euclidean geometry and Archimedean statics have prevailed as showpieces of the Aristotelian ideal of knowledge. In this connection we also have to mention Spinoza's *Ethica* and Newton's *Principia*. Only in the nineteenth century did developments in the different sciences – including the discovery of non-Euclidean geometries by Bolyai and Lobachevsky[2] – undermine the evidence postulate, as a consequence of which the Aristotelian ideal of knowledge itself finally fell into disfavour.

[1]See [De Rijk, 1988, p. 8].

[2]See [Beth, 1950, Chapter 1].

The success of the sciences had already forced philosophers to revise their understanding of science – since the beginning of the twentieth century, a dynamic conception of science in all its diversity has become generally accepted –, but only recently, the collapse of the Aristotelian ideal of knowledge gave rise to a new epistemology. Until now the revolution in epistemology has not been fully realized, a fact that can be partially explained by the aforementioned compatibility between a dynamic scientific practice, in which the evidence postulate no longer plays a part, and a static ideal of knowledge, in which necessity is still conceived as a characteristic of truth.

A second reason that can be put forward is the circumstance that the Aristotelian ideal of knowledge has long persevered in philosophy as the methodological backbone of so-called 'philosophical disciplines' (such as *philosophical logic*, *philosophical theory of space*, and *philosophy of nature* in the first half of the twentieth century[3]), thus preserving the scientific ring of metaphysics.[4]

> However, the illusive character of this preservation immediately comes to light when we examine what is actually being offered under the lofty heading of 'philosophical discipline'. It then becomes clear that all it involves is a collection of notions that have been discarded by the 'sciences'. [Beth, 1964, p. 33]

A third argument for the slow propagation of a new epistemology is the fact that modern (post-Fregean, 'mathematical') logic, an instrument that played a crucial part in the development of analytical philosophy, originated from the study of *mathematical* reasoning. The logical investigations of Frege and Russell, who criticized traditional logic on the grounds of its failure to describe mathematical reasoning adequately and who consequently developed new logics that were more suited for the task, produced logical systems that were concerned with the modelling of reasoning within a relatively small and atypical field of scientific knowledge, inasmuch as it was cumulative and deductive, namely mathematics. The triumph of this mathematically oriented logic led to the idea that 'mathematical logic' could serve as an instrument for the analysis of human knowledge as such (Russell's 'logical atomism' comes to mind). However, Hans Reichenbach justly writes:

> The way to a consistent empiricism is open only to those who are ready to interpret empirical knowledge in its own right, to abandon the prejudice that mathematics is the prototype of all knowledge. [Reichenbach 1948, p. 142]

[3] See [Beth, 1964, p. 38]. Contemporary philosophical logic bears hardly any similarity to its prewar namesake.

[4] Remnants of the Aristotelian ideal of knowledge are also found in the attempts at *Letztbegründung* by early twentieth-century analytic philosophers, with the exception of Otto Neurath and Kazimierz Ajdukiewicz.

Fourthly, the clear inadequacy of 'mathematical logic' in the fields of knowledge and communication outside mathematics has given rise to a vast number of esoteric logics for a broad range of large and small tasks, and that in turn has favoured the independence of contemporary philosophical logic, which has almost entirely freed itself from epistemology.

In his inquiries into the conditions of human cognition during his naturalistic period (from circa 1870 to the 1880s), Charles Peirce combined a systematic interest in non-deductive logics with an unfortunately not altogether consistent rejection of the Aristotelian ideal of knowledge. It is notable that precisely Peirce's attempts to define the traditional epistemological concepts of 'truth' and 'reality' have flourished in contemporary epistemology, while almost everyone has turned a blind eye to the new epistemological perspectives opened up by his pragmatism.[5]

In the following section, Peirce's philosophy of logic is summarized so that we can get a grip on his attempts to conceive truth as the limit of scientific inquiry. The final section is a refutation of the arguments of a contemporary advocate of a Peircean account of truth.

2.2 C.S. Peirce's Theory of Scientific Inquiry

All the familiar handbooks on the history of logic, whose authors unanimously restrict the domain of logical inquiry to mathematical reasoning, traditionally present Charles Sanders Peirce (1839–1914) as co-founder of the theory of relations and as a researcher in the field of Boolean algebra: the logical operator 'neither... nor...' with its remarkable properties had already been discovered by Peirce round 1880,[6] while Peirce's notation of logical quantifiers (Π for the universal and Σ for the existential quantifier), which was implemented by his pupil O.H. Mitchell in 1883, independently of Frege, quickly became current among algebrists.[7] The game theoretical interpretation of logical quantifiers, propagated by Paul Lorenzen and Jaakko Hintikka in the second half of the twentieth century, was also pioneered by Peirce.[8]

Nevertheless, there is something to be said against the image presented by historians of Peirce's philosophy of logic. Recent history of logic has mainly highlighted the historical run up to succesful developments in mathematical logic and, as a consequence, divergent views have been eclipsed.

From the beginning Peirce, who like Frege and Russell was a critic of traditional logic, followed in the footsteps of the English school of logic, that pursued

[5] Isaac Levi's pragmatist epistemology is an important exception.

[6] Since 1913, H.M. Sheffer has been wrongly credited with the discovery of the fact that the operators of classical propositional logic can be defined with only one operator. The *Sheffer stroke* 'not both... and...' is the dual of Peirce's operator. See [Kneale and Kneale, 1962, p. 423].

[7] [Putnam, 1982].

[8] [Hilpinen, 1982].

"a connected view of the principles of evidence and the methods of scientific investigation", and had a broader perspective of human reasoning than his mathematically oriented colleagues: the justification of scientific knowledge – a central problem in Kant's *Kritik der reinen Vernunft* – became the focal point of his philosophical interests.

In the first article of the lectures series *Illustrations of the Logic of Science*, published in 1877 and 1878, in which Peirce's aim is to describe "the method of scientific investigation", he submits that before we can understand human reasoning, we must acquaint ourselves with the fact "that there are such states of mind as doubt and belief – that a passage from one to the other is possible, the object of thought remaining the same, and that this transition is subject to some rules which all minds are alike bound by" (*CW* 3, 246).[9] Following this observation, Peirce, in accordance with Alexander Bains's definition of *belief* as "that upon which a man is prepared to act", considers the practical difference between *doubt* and *belief*:

> Our beliefs guide our desires and shape our actions. [...] So it is with every belief, according to its degree. The feeling of believing is a more or less sure indication of there being established in our nature some habit which will determine our actions. Doubt never has such an effect.
>
> [...] Doubt is an uneasy and dissatisfied state from which we struggle to free ourselves and pass into the state of belief; while the latter is a calm and satisfactory state which we do not wish to avoid, or to change belief into something else (*CW* 3, 247).[10]

Doubt can only exist thanks to a background of established beliefs, which together form a pattern of expectations.[11] An unforeseen experience – an infringement of the pattern of expectations – breeds true doubt: "A true doubt is [...] a doubt which really interferes with the smooth working of the belief-habit" (*CP* 5.510). Our beliefs are shaken by this unforeseen experience, which after all in some way

[9] '*CW m, n*' refers to page *n* of volume *m* of the *Writings of Charles S. Peirce*; '*CP m.n*' refers to section *n* of volume *m* of the *Collected Papers of Charles Sanders Peirce*; and '*NE m, n*' refers to page *n* of volume *m* of Peirce's *The New Elements of Mathematics.*

[10] "And what, then, is belief? [...] We have seen that it has just three properties: first, it is something that we are aware of; second, it appeases the irritation of doubt; and, third, it involves the establishment in our nature of a rule of action, or, say for short, a *habit*." (*CW* 3, 263.)

[11] "[T]here is but one state of mind from which you can 'set out,' namely, the very state of mind in which you actually find yourself at the time you do 'set out' – a state in which you are laden with an immense mass of cognition already formed, of which you cannot divest yourself if you would [...] Do you call it *doubting* to write down on a piece of paper that you doubt? If so, doubt has nothing to do with any serious business. [...] [T]here is much that you do not doubt, in the least. Now that which you do not at all doubt, you must and do regard as infallible, absolute truth." (*CP* 5.416.)

must be accounted for – an undesirable situation that can only be remedied by logic, which, according to Peirce, "may be defined as the science of the laws of the stable establishment of beliefs" (*CP* 3.429).

Reluctantly, Peirce chooses the term *inquiry* for the "struggle to attain a state of belief", a struggle that is provoked by the stimulus of doubt, and he states: "The sole object of inquiry is the settlement of opinion" (*CW* 3, 248). Most importantly, Peirce finds the question whether the belief that has thus been reached is *true* irrelevant for the study of *inquiry*, "for as soon as a firm belief is reached we are entirely satisfied, whether the belief be true or false" (*CW* 3, 248). This judgement clears the path for Peirce's pragmatist account of truth: a definition of truth in terms of *inquiry*.

Subsequently, four different methods with which our opinions can be settled are individually reviewed: 'the method of tenacity', 'the method of authority', 'the *a priori* method' (in which logical deduction already plays a part), and last but not least, 'the method of science'. Since the first three methods, albeit in different ways, are all subject to the whims and fancies of man, and, as opposed to the method of science, only strengthen a preconceived opinion and therefore cannot guarantee the reliability of their results, Peirce rejects them and then favours the method of science, that "tends to unsettle opinions at first, to change them and to confirm a certain opinion which depends only on the nature of investigation itself" (*CW* 3, 17):

> To satisfy our doubts, therefore, it is necessary that a method should be found by which our beliefs may be caused by nothing human, but by some external permanency – by something upon which our thinking has no effect. [...] It must be something which affects, or might affect, every man. And, though these affections are necessarily as various as are individual conditions, yet the method must be such that the ultimate conclusion of every man shall be the same. Such is the method of science. Its fundamental hypothesis, restated in more familiar language, is this: There are real things, whose characters are entirely independent of our opinions about them; those realities affect our senses according to regular laws, and, though our sensations are as different as our relations to the objects, yet, by taking advantage of the laws of perception, we can ascertain by reasoning how things really are, and any man, if he have sufficient experience and reason enough about it, will be led to the one true conclusion. (*CW* 3, 253–254)[12]

Reality, even though it only shows its true face once 'the final opinion' has been reached, and scientific method are the bulwarks that enable us to systematize the

[12]Cf. "[W]e may define the real as that whose characters are independent of what anybody may think them to be." (*CW* 3, 271.)

surging sea of subjective experience.[13] As reality only becomes known once our permanently ongoing inquiries have been completed, a description of that reality should play no part in the investigation of scientific method.

2.2.1 Analytic and Synthetic Reasoning

Peirce, who is rarely consistent in his terminology, distinguishes two main classes of inferences in scientific method: on the one hand there are explicative, analytic, or deductive inferences,[14] and on the other hand, ampliative, synthetic, or inductive inferences (later, in the 1878 essay 'Deduction, Induction, and Hypothesis', Peirce subdivides the latter class into (1) the class of inductive inferences and (2) the class of abductive inferences, also termed 'hypothesis'). Although in mathematics, of which probability theory is a part, it is possible to limit the study of proofs to the examination of deductive reasoning, "the only inferences which increase our real knowledge" are synthetic inferences, as "the facts summed up in the conclusion are not among those stated in the premises" (*CW* 3, 297).[15] The Kantian Peirce considers finding an answer to the question as to how synthetic inferences are possible to be "the lock upon the door of philosophy" (*CW* 2, 268).

Peirce's essay 'The Probability of Induction' is devoted to the following key question: can synthetic inferences be reduced to analytic inferences? Despite the fact that, in an earlier essay, Peirce commits himself to a different point of

[13]Quine goes even further and claims that our methods and techniques are also eligible for revision: "The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the picture, total science is like a field of force whose boundary conditions are experience." [Quine, 1953, p. 42.] It would be a misconception, based on a sloppy interpretation of Peirce's views on quantitative induction, to state, as Ernest Nagel did, that Peirce "maintained that the logical canons and methods of scientific inquiry have themselves been obtained in the course of inquiry, and that they are 'self-corrective' – both in the sense that hypotheses about statistical or other properties of populations can be improved by continued use of those logical methods, as well as in the sense that limitations and defects in those methods can be discovered in the very process of using them, and can be remedied by supplementing established methods with new ones suggested by problems encountered in inquiry." [Nagel, 1982, p. 308.] According to Peirce, reality and logic lie outside the domain of change. For a critique of Nagel's interpretation, see [Misak, 1991, Chapter 3, notably p. 111–119].

[14]As said before, deduction already plays a part in the '*a priori* method', with which Peirce particularly refers to philosophical systems that seem to comply with the Aristotelian ideal of knowledge discussed in our introduction. According to Peirce, it is in fact the evidence postulate that plays tricks on the method we use to settle our opinions: "The most perfect example of it [the *a priori* method; AMT] is to be found in the history of metaphysical philosophy. Systems of this sort have not usually rested upon any observed facts, at least not in any degree. They have chiefly been adopted because their fundamental propositions seemed 'agreeable to reason'" (*CW* 3, 252). Among others, Peirce's logic can be seen as an attempt to eradicate the subjectivist appeal to evidence or insight in epistemology and to locate the actual underlying rules of these faculties.

[15]Making a diagnosis is a good example of ampliative reasoning.

view,[16] deductive reasoning certainly does not lay down the standard for the judgement of all reasoning: "[I]t is only an absurd attempt to reduce synthetic to analytic reason" (*CW* 3, 301). Peirce arrives at this negative conclusion after discussing as clearly and as forcibly as possible the most eligible candidate for an eventual reduction, the 'frequency theory of probability',[17] in which, in accordance with Venn, probability is interpreted "as the proportion of times in which an occurence of one kind is accompanied by an occurence of another kind" (*CW* 3, 291). Nevertheless, there are clearly unbridgeable differences between analytic and synthetic reasoning,[18] differences that Peirce articulates as follows:

> When we draw a deductive or analytic conclusion, our rule of inference is that facts of a certain general character are either invariably or in a certain proportion of cases accompanied by facts of another general character. Then our premise being a fact of the former class, we infer with certainty or with the appropriate degree of probability the existence of a fact of the second class. But the rule for synthetic inference is of a different kind. [...] [S]ynthetic inference is founded upon a classification of facts, not according to their characters, but according to the manner of obtaining them. Its rule is, that a number of facts obtained in a given way will in general more or less resemble other facts obtained in the same way; or, *experiences whose conditions are the same will have the same general characters.*
>
> [...] [I]n the case of analytic inference we know the probability of our conclusion (if the premises are true), but in the case of synthetic inferences we only know the trustworthiness of our proceeding. As all knowledge comes from synthetic inference, we must equally infer that all human certainty consists merely in our knowing that the processes by which our knowledge has been derived are such as must generally have led to true conclusions. (*CW* 3, 305)[19]

---

[16]In 'The Doctrine of Chances', Peirce verges on a reductionism which he then fervently refutes in the next essay. He writes: "The theory of probabilities is simply the science of logic quantitatively treated" (*CW* 3, 278), and on top of that he adds (after all, the first quote certainly does not rule out a *qualitative* logic): "The general problem of probabilities is, for a given state of facts, to determine the numerical probability of a possible fact. This is the same as to inquire how much the given facts are worth, considered as evidence to prove the possible fact. Thus the problem of probabilities is simply the general problem of logic." (*CW* 3, 278.)

[17]The term is from [Kneale, 1949, p. 150].

[18]See also Peirce's explanation in *CW* 3, 324–325.

[19]The reader's attention should be drawn to the *epistemic component* of synthetic inferences, which after all are based on "a classification of facts [...] according to the manner of obtaining them." See also Hookway, who writes about synthetic inferences: "According to Peirce, their ampliative character is reflected in the fact that whether an inductive or hypothetical argument is a good one depends on the *non-existence* of some other knowledge". [Hookway, 1985, p. 31–32.]

Now that the attempt to base inductive inference solidly on probability theory has run aground, Peirce reassures the reader: after all, the trustworthiness of inductive reasoning can be substantiated with the concept of limits: "Though a synthetic inference cannot by any means be reduced to deduction, yet that the rule of induction will hold good in the long run may be deduced from the principle that reality is only the object of the final opinion to which sufficient investigation would lead" (*CW* 3, 305). Nevertheless, knowledge acquired through the 'logic of science' is especially susceptible to error because of its essentially synthetic character:

> That we ever do discover the precise nature of things, that any induction whatever is absolutely without exception, is what we have no right to assume. (*CW* 3, 317)

2.2.2 Deduction, Induction, and Abduction

The special nature of Peirce's logical investigations into the conditions of induction and abduction can be better understood by setting these investigations off against for example Russell's ideas concerning the relation between reason and intuition. Russell, who like Peirce is convinced of the untenability of the Aristotelian ideal of knowledge, in which reliance on intuition is essential in order to justify the fundamental principles, writes in his well-known essay 'Mysticism and Logic' of 1914:

> Instinct, intuition, or insight is what first leads to the beliefs which subsequent reason confirms or confutes; but the confirmation, where it is possible, consists, in the last analysis, of agreement with other beliefs no less instinctive. Reason is a harmonising, controlling force rather than a creative one. Even in the most purely logical realm, it is insight that first arrives at what is new. [Russell, 1917, p. 13]

Russell banishes insight from the realm of reason, so that any investigation into the conditions on the basis of which insight is reached is out of the question. Only once insight has been found, its tenability is tested by rational means.[20]

In his 1868 essay 'Questions Concerning Certain Faculties Claimed for Man', Peirce targets different claims concerning intuition and denies that we have an "intuitive faculty of distinguishing intuitive from mediate cognitions" (*CW* 2, 200). He points out that many supposed intuitions imperceptibly are still the result of reasoning,[21] and this encourages him to include these cognitive phenomena

[20] Reichenbach's distinction between 'context of discovery' and 'context of justification' also should be understood in this context.

[21] Peirce's thesis of the omnipresence of reasoning in the acquisition of knowledge is exemplified in the following quotes: "[W]hen the reasoning is easy and natural to us, however complex may be the premises, they sink into insignificance and oblivion proportionally to the satisfactoriness

in the domain of logic and consequently to champion the systematic investigation of non-deductive types of inference: according to Peirce, logic is "the doctrine of truth, its nature and the manner in which it is to be discovered" (*CW* 3, 14).

In the 1866 *Lowell Lectures*, Peirce already distinguishes between two types of ampliative inference, induction and hypothesis, a distinction that is made within the framework of Aristotelian syllogistic. Even in 1878, Peirce maintained – falsely, for that matter – that "all inference may be reduced in some way to *Barbara*" (*CW* 3, 323), but, with the considerations of the previous subsection in mind, he keeps his options open: "[A]s long as the *is* [the copula in *Barbara*; AMT] is taken literally, no inductive reasoning can be put into this form" (*CW* 3, 324). Nevertheless, the logical form to which inductive reasoning as well as abductive reasoning can be reduced is still linked to the form of the syllogism. After all, Peirce's examples only include propositions in subject-predicate form, and each inference invariably consists of two premises and one conclusion. Consequently, combinatorics show us that this straitjacket not only suits *Barbara*, the patron saint of deductive reasoning, but that it also fits the two 'new' types of reasoning, induction and abduction. Therefore, the strength of Peirce's distinctions does not necessarily reside in the recognition of the latter two types of reasoning, but rather in the independent status that Peirce gives them, and as a result, in the incitement to investigate under which conditions these types of reasoning are sound.

As an introduction to the definition of inductive and abductive inference, Peirce first discusses deductive reasoning, which is characterized by "the application of general rules to particular cases", and comes up with the following example (which is not, strictly speaking, of the form *Barbara*):

DEDUCTION

	Rule.	—	All the beans from this bag are white.
	Case.	—	These beans are from this bag.
∴	*Result.*	—	These beans are white.

In an earlier article from the same lectures series, in which induction is instead examined in a more statistical context, Peirce, priding himself on originality, defines an inductive inference as follows:

> *The inference that a previously designated character has nearly the same frequency of occurence in the whole of a class that it has in a sample drawn at random out of that class is induction.* (*CW* 3, 313)

Such an inference "only has its full force when the character concerned has been designated before examining the sample" (*CW* 3, 316), a reference by Peirce to

of the theory based upon them" (*CW* 2, 199), and "there is no judgment of pure observation without reasoning" (*CW* 3, 300).

the difference in forcefulness between the following considerations: (1) predicting that half the beans from a given bag of beans are white, then seeing that approximately half the beans from a handful of beans from the given bag are white, and concluding that half the beans from the bag concerned are white; and (2) seeing that approximately half the beans from a handful of beans from a given bag of beans are white and concluding that half the beans from the bag concerned are white. The latter reasoning is obviously the weakest.

This striking distinction is no longer found in the essay 'Deduction, Induction, and Hypothesis', in which the definition of inductive reasoning is adapted to a syllogistic context: "[I]nduction is the inference of the *rule* from the *case* and *result*". A little further on, we read:

> Induction is where we generalize from a number of cases of which something is true, and infer that the same thing is true of a whole class. Or, where we find a certain thing to be true of a certain proportion of cases and infer that it is true of the same proportion of the whole class. (*CW* 3, 326)

This description serves to elucidate the following example:

Induction

	Case.	—	These beans are from this bag.
	Result.	—	These beans are white.
∴	*Rule.*	—	All the beans from this bag are white.

Abductive reasoning, here still termed 'hypothesis' by Peirce, actually consists of "the inference of a *case* from a *rule* and *result*" and is somewhat clarified by the following words:

> Hypothesis is where we find some very curious circumstance, which would be explained by the supposition that it was a case of a certain general rule, and thereupon adopt that supposition. Or, where we find that in certain respects two objects have a strong resemblance, and infer that they resemble one another strongly in other respects. (*CW* 3, 326)

Peirce illustrates his claims with the following reasoning:

Abduction

	Rule.	—	All the beans from this bag are white.
	Result.	—	These beans are white.
∴	*Case.*	—	These beans are from this bag.

Then Peirce comes up with a condition by which abductive inferences, which in current literature on argumentation theory are also discussed under the overall heading 'inferences to the best explanation', can be properly evaluated, even if they are only "a weak kind of argument":

> When we adopt a certain hypothesis, it is not alone because it will explain the observed facts, but also because the contrary hypothesis would probably lead to results contrary to those observed. (*CW* 3, 328)

According to Peirce, the following clearer rules should be adopted "in order that the process of making an hypothesis should lead to a probable result":

> 1. The hypothesis should be distinctly put as a question, before making the observations which are to test its truth. In other words, we must try to see what the result of predictions from the hypothesis will be.
> 2. The respect in regard to which the resemblances are noted must be taken at random. We must not take a particular kind of predictions for which the hypothesis is known to be good.
> 3. The failures as well as the successes of the predictions must be honestly noted. The whole proceeding must be fair and unbiased. (*CW* 3, 331)

However, there is also a condition that hypotheses have to meet before they can even be allowed to serve as a conclusion of an abductive inference: it must be possible to verify whether the abductive inference, of which the hypothesis concerned is the conclusion, complies with the rules mentioned above, that is, the hypothesis must make predictions that are verifiable. In short, it must comply with Peirce's pragmatic maxim: "Consider what effects, which might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object" (*CW* 3, 266).[22] Hypotheses without any practical effects cannot be evaluated correctly using the rules of abduction. The *admissibility* of an hypothesis is determined by Peirce's pragmatic maxim, the *tenability* of an hypothesis is (among other things) determined by the said three conditions of abduction.[23]

By way of conclusion, Peirce discusses at some length the similarities, but especially the differences between induction and abduction. With regard to their role in the formulation of scientific theories, Peirce claims (note the epistemic component in both cases):

[22] Unlike the verification criterion advocated by certain members of the Vienna Circle, hypotheses in the field of logic and mathematics also come under Peirce's pragmatic criterion. It would be getting too far off the subject to discuss Peirce's philosophy of mathematics here.

[23] Misak gives a list of Peirce's abduction criteria and their locations [Misak, 1991, p. 99] that would have relished Lakatos, but unfortunately she confuses the admissibility and the tenability of hypotheses.

> By induction, we conclude that facts, similar to observed facts, are true in cases not examined. By hypothesis, we conclude the existence of a fact quite different from anything observed, from which, according to known laws, something observed would necessarily result. The former, is reasoning from particulars to the general law; the latter, from effect to cause. The former classifies, the latter explains. (*CW* 3, 332)

Abductive reasoning allows us to conclude facts that are beyond direct observation,[24] which makes abduction an indispensable part of scientific method.

Our forever temporary beliefs determine the expectations on which our actions are based. An unforeseen experience undermines our beliefs, so that an *inquiry* is needed to re-establish the lost equilibrium and to overcome the doubt caused by the unforeseen experience. Such an inquiry follows certain specific rules that can be formulated and that are subdivided into three main types: abduction, deduction and induction. (Qualitative) induction provides us with classifications of our observations, expressed in 'empirical formulæ'. In turn, these often not fully accurate generalizations (our observations are often inaccurate) serve as a basis for an explanatory hypothesis that is tested by checking whether the predictions derived from this hypothesis add up. (Quantitative) induction plays a part in testing the hypothesis. Although essentially uncertain, the result of this *inquiry*, precisely due to the fact that it has undergone the procedure followed and thus warrants a 'maximum of expectation' and a 'minimum of surprise', supplies a sufficient basis for our thoughts and actions, until once again an unforeseen experience compels us to start a new *inquiry.* By open-mindedly and conscientiously following this procedure of 'belief–doubt–belief', we would eventually arrive at a set of beliefs that no experience can contradict, a pattern of expectations that will never have any surprises in store for us: the Truth.

2.3 Truth in Terms of Inquiry

Among all the current philosophical wrangling about truth, there are several authors calling themselves pragmatists, who maybe do not actually define, but at least provide an account of the concept of truth or of the somewhat broader 'warranted assertability' more or less in terms of a theory of scientific inquiry, and whose work is within the scope of the subject matter discussed in this chapter.[25]

[24]Cf. "The great difference between induction and hypothesis is, that the former infers the existence of phenomena such as we have observed in cases which are similar, while hypothesis supposes something of a different kind from what we have directly observed, and frequently something which it would be impossible for us to observe directly." (*CW* 3, 335–336.)

[25][Jardine, 1986, p. 21–35]; [Misak, 1991, *passim*] and [Misak, 1998]; [Putnam, 1981, p. 49–74] and [Putnam, 1990, p. 223] (in the preface to *Realism with a Human Face*, Putnam rejects interpretations of his definition of truth as "an *idealization* of rational acceptability" in terms

From this literature I have selected Cheryl Misak's study *Truth and the End of Inquiry: A Peircean Account of Truth* (1991), because she undertook not only to make a meticulous reconstruction of Peirce's doctrines, but also to defend them. In this section I will weigh her arguments on a gold platter to see if they have the right content.

As noted before, Peirce, who attempted to describe *inquiry* without reference to the concept of truth, insists on understanding truth (and falsehood) in terms of inquiry: "[T]he ideas of truth and falsehood, in their full development, appertain exclusively to the scientific method of settling opinion." (*CW* 3, 272) It will be clear – just a minimum of historical awareness will do – that the relation between truth and inquiry cannot simply be situated in the assumption that the results of scientific investigation are true. Peirce, who was familiar with the developments in the sciences of his time, seized the opportunity to argue for his explication of truth in terms of inquiry as follows:

> [A]ll the followers of science are fully persuaded that the processes of investigation, if only pushed far enough, will give one certain solution to every question to which they can be applied. [...] They may at first obtain different results, but, as each perfects his methods and his processes, the results will move steadily towards a destined centre. So with all scientific research. Different minds may set out with the most antagonistic views, but the progress of investigation carries them by a force outside of themselves to one and the same conclusion. This activity of thought by which we are carried, not where we wish, but to a foreordained goal, is like the operation of destiny. No modification of the point of view taken, no selection of other facts for study, no natural bent of mind even, can enable a man to escape the predestinate opinion. This great law is embodied in the conception of truth and reality. The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real. That is the way I would explain reality. (*CW* 3, 273)

This is a remarkable definition, not only because of the deterministic character attributed to inquiry, but also because of the fact that Peirce's explication of truth in terms of inquiry fully depends on an ostensibly mathematical concept of limits, a concept that is fundamental to crucial aspects of Peirce's work.[26]

of a Peircean ideal limit of inquiry); [Reynolds, 2000].

[26] In his article 'The Doctrine of Chances', after pointing out that human life is finite ("death makes the numbers of our risks, of our inferences, finite, and so makes their mean result uncertain" (*CW* 3, 283–284)), Peirce discusses the way in which, in the long run, a probability can be assigned to a probabilistic inference: "I can see but one solution of it. It seems to me that we are driven to this, that logicality inexorably requires that our interests shall *not* be limited. They must not stop at our own fate, but must embrace the whole community. This community,

In her *Truth and the End of Inquiry*, Misak supports a favourable interpretation of Peirce's explication of the concept of truth, an explication that, according to her, should not be interpreted as a *definition* of truth, but as a supplement to Tarski's correct but empty convention T.[27] In Misak's argument, well interspersed with quotations, Peirce's deliberations with regard to the concept of truth are eventually condensed into the following two subjunctive conditionals, that however do not have an equal status: the first, the so-called T–I conditional (from *truth* to *inquiry*) is a regulative assumption, the second, the I–T conditional (from *inquiry* to *truth*), a claim:

(T–I) *if* H is true *then* if inquiry relevant to H were pursued as far as it could fruitfully go, H would be believed. [Misak, 1991, p. 43]

(I–T) *if*, if inquiry relevant to H were to be pursued as far as it could fruitfully go, then H would be believed, *then* H is true. [Misak, 1991, p. 46]

2.3.1 The T–I Conditional versus Bivalence

Misak defends her refusal to assume the burden of proof for the T–I conditional by denying that it is a *claim*: according to her, we have to take into account the possibility that this conditional does not apply, a possibility of which Peirce was also aware. Namely, if indispensable material concerning a certain hypothesis has been lost once and for all, it is very unlikely that an endless series of investigations will finally provide definitive proof for or against the hypothesis concerned. Peirce writes:

---

again, must not be limited, but must extend to all races of beings with whom we can come into immediate or mediate intellectual relation. It must reach, however vaguely, beyond this geological epoch, beyond all bounds. He who would not sacrifice his own soul to save the whole world, is, as it seems to me, illogical in all his inferences, collectively. Logic is rooted in the social principle." (*CW* 3, 284.) A little further on he writes: "But all this requires a conceived identification of one's interests with those of an unlimited community." (*CW* 3, 285.) Induction, too, is eventually justified with a reference to the final opinion: "[T]hat the rule of induction will hold good in the long run may be deduced from the principle that reality is only the object of the final opinion to which sufficient investigation would lead." (*CW* 3, 305.)

[27] Misak remarks that "Tarski leaves open what it takes for a given sentence (or object) to satisfy the condition" [Misak, 1991, p. 127] and "the Tarski-style definition tells us nothing substantial about the property truth. Like the correspondence formula, it does not tell us what to expect of true hypotheses and it does not tell us how to go about inquiring into the truth of hypotheses. [...] In its efforts to refrain from invoking anything mysterious it fails to engage the truth with *anything* so as to give us a grasp of the predicate 'is true'" [Misak, 1991, p. 128]. Wright contests these statements in the first chapter of [Wright, 1992].

As a matter of fact, Tarski's convention T is not 'correct'; after all, it cannot handle the liar's paradox. Successful attempts to sidestep this problem with a fixed-point construction stem from [Kripke, 1975].

> But I may be asked what I have to say to all the minute facts of history, forgotten never to be recovered, to the lost books of the ancients, to the buried secrets. [...] Do these things not really exist because they are hopelessly beyond the reach of our knowledge? (*CW* 3, 274)

Here the principle of bivalence,[28] not to be confused with the principle of the excluded middle,[29] comes into play. After all, if we are unable to gather enough evidence in regard to a certain hypothesis H, H could, if we strictly apply the definition, be neither true nor false, which is inconsistent with the principle of bivalence that states that each hypothesis is either true or false.[30]

The solution for this problem that Nicholas Jardine puts forward in *The Fortunes of Inquiry* (1986) – "we must imaginatively escape from the spatio-temporal limitations on the evidence-gathering capacities of ourselves and other physically possible inquirers. The fiction of time-travel provides [...] precisely what is needed" [Jardine, 1986, p. 30] – is evidently rejected forthright by Misak.[31] Because Misak sees no other way out, the possibility is kept open that some questions may never be answered,[32] which implies that there is no need to look for arguments to substantiate the T–I conditional. Still, Peirce and Misak do not throw in the towel:

> Logic requires us, with reference to each question we have in hand, to hope some definite answer to it may be true. That *hope* with reference to *each* case as it comes up is, by a *saltus*, stated by logicians as a law concerning *all* cases, namely the law of the excluded middle. (*NE* IV, xiii)[33]

[28]In this connection, Crispin Wright uses the word *completeness*: "The requirement of completeness would be that, for each statement, *either* it *or* its negation must be justified under epistemically ideal circumstances." [Wright, 1992, p. 39.]

[29]The principle of the excluded middle, according to which each statement $H \vee \neg H$ is true, can be maintained in the absence of the principle of bivalence. For technical details, see Bas van Fraassen's study on supervaluations [Van Fraassen, 1966].

[30]Hookway remarks: "It is a *prima facie* implausible feature of the doctrine that it suggests that there is no fact of the matter whether such propositions are true." [Hookway, 1985, p. 38.] It may be clear that on the grounds of Peirce's conception of reality a correspondence theory of truth is not an option for Peirce: "[H]ow futile it was to imagine that we were to clear up the idea of *truth* by the more occult idea of *reality*!" (*CP* 1.578.)

[31]"A strategy involving counterfactual bravado abandons the pragmatist's commitment to say something about the relationship between truth and inquiry [...] and replaces it with a claim about what the relationship between truth and inquiry would be if inquiry were something it is not." [Misak, 1991, p. 153–154.]

[32]Peirce notes: "We cannot be quite sure that the community ever will settle down to an unalterable conclusion upon any given question. Even if they do so for the most part, we have no reason to think the unanimity will be quite complete, nor can we rationally presume any overwhelming *consensus* of opinion will be reached upon every question." (*CP* 6.610.)

[33]Cf. "[I]t is unphilosophical to suppose that, with regard to any given question (which has any clear meaning), investigation would not bring forth a solution of it, if it were carried

Misak promptly promotes this hope to the status of a regulative assumption – assumptions do not have to be defended – in inquiry, and even speaks of "the hope of bivalence" [Misak, 1991, p. 147], so that the principle of bivalence is secured:

> The inquirer must assume, for a hypothesis which is thought to be objective, that there is a chance that inquiry would eventually settle on its truth-value. [Misak, 1991, p. 141]

However, Misak does not blindly rely on the charity of the reader and presents him with a somewhat contrived consideration: "The fact that the regulative assumption is foolish with respect to a particular issue will coincide with the fact that no one will inquire into that issue." [Misak, 1991, p. 157.] Furthermore, Misak expresses the opinion that we have no right to claim that a certain hypothesis will never have any truth-value, because after all, the T–I conditional shows us that first the final opinion has to be reached before we can justify such a claim. "But since the antecedent of this conditional [that is, the consequent of the T–I conditional; AMT] is about what would be the case if inquiry were pushed indefinitely far, we can never assert it is fulfilled." [Misak, 1991, p. 156].

Can the principle of bivalence be defended more convincingly and without astute quiddities? In his 'What Price Bivalence?', Quine commends bivalence as a basic characteristic of our classical theories of nature:

> It has us positing a true-false dichotomy across all the statements that we can express in our theoretical vocabulary, irrespective of our knowing how to decide them. In keeping with our theories of nature we have viewed all such sentences as having factual content, however remote from observation. In this way simplicity of theory has been served. [Quine, 1981, p. 36]

The gist of the principle of bivalence is that we assume that even statements whose evidence has been lost once and for all are nevertheless either true or false:

> We declare that it is either true or false that there was an odd number of blades of grass in Harvard Yard at the dawn of Commencement Day, 1903. The matter is undecidable, but we maintain that there is a fact of the matter. [Quine, 1981, p. 32]

Quine does not feel inclined to defend bivalence by shielding it from possible criticism, as Misak did with her 'hope of bivalence'. Because Misak, as we shall see, rejects the epistemological holism on which Quine's defence of the principle of bivalence is based, it is not open to her to follow Quine's plausible strategy in order to strengthen her rather feeble position:

---

far enough." (*CW* 3, 274.) On other crucial points of argument in his epistemology, Peirce also addresses the hope that everything will turn out all right in the end. Thus "hope in the unlimited continuance of intellectual activity" (*CW* 3, 285) is Peirce's final resort for the defence of the possibility to attribute a probability to inferences.

> The question about the grass of 1903 hinged, one felt, on a robust matter of fact. Still, being clearly indecidable, the question makes empirical sense to us only by analogy and extrapolation. It makes sense because we often do count things, and are prepared even to count present blades of grass. We project these vivid notions into the inaccessible past as a matter of course, such is the organisation of our system of the world. [...] This undecidable question [...] makes empirical sense to us only by virtue of the devious connections between our systematic theory of the world and the various observations to which the system as a whole is answerable. [Quine, 1981, p. 35]

2.3.2 The I–T Conditional is Wide of the Mark

Let us now take a closer look at the I–T conditional and try to beat the pragmatists at their own game. Royce and Russell had already stated that the antecedent of the I–T conditional could not be a material implication, and in this chapter we will demonstrate that even a pragmatist cannot pull it off with a counterfactual.

It may be useful to start with Misak's argumentation for her pragmatist account of truth. The main argument for the I–T conditional runs as follows: suppose we have exhaustively investigated an hypothesis H and have come to the conclusion that H is the case. Then it could not be that a possible experience, whatever its origin may be, could still undermine this conclusion, since this would only mean that our inquiries have not been fully exhaustive. Peirce's pragmatic principle ("there is no distinction of meaning so fine as to consist in anything but a possible difference of practice" (*CW* 3, 265)) prohibits us to withhold the predicate 'true' from H.[34] Peirce writes:

> To believe the absolute truth would be to have such a belief that under no circumstance, such as would actually occur, should we find ourselves surprised. (MS 693, 166) [Misak, 1991, p. 87]

As stated, Royce and Russell make clear that the implication in the antecedent of our conditional *cannot* be a material implication:[35] Royce and Russell ask whether beliefs are in fact true as long as the possibility still exists that the final opinion may never be reached.[36] Their argument can be strengthened as follows:

[34]Misak rhetorically asks: "When we have beliefs that would forever withstand the tests of experience and argument, what is the point of refusing to confer upon them the title 'true'?" [Misak, 1991, p. 47.]

[35]See also Hookway 1985, 38–39.

[36]William Kneale takes it one step further: "It is surely false, however, that the possibility of rational action in the circumstances we are considering depends on the prospects of survival of the human race. And even if we were sure that the human race would survive for ever and were animated by the most devoted altruism, we could attach no meaning to the promise of an advantage which was to be realized only at the *end* of infinite duration." [Kneale, 1949, p. 166.]

suppose that our investigations relevant to H will never be completed. Then, of course, the fact also applies that our investigations relevant to $\neg H$ will never be completed. In that case, the statements "inquiry relevant to H has been pursued as far as it could fruitfully go" and "inquiry relevant to $\neg H$ has been pursued as far as it could fruitfully go" are both false. Therefore, the material implications "if inquiry relevant to H has been pursued as far as it could fruitfully go, then H is believed" and "if inquiry relevant to $\neg H$ has been pursued as far as it could fruitfully go, then $\neg H$ is believed" are both true. Now, suppose that the I–T conditional is correct. Then we have, firstly "*if*, if inquiry relevant to H has been pursued as far as it could fruitfully go, then H is believed, *then* H is true" and secondly, "*if*, if inquiry relevant to $\neg H$ has been pursued as far as it could fruitfully go, then $\neg H$ is believed, *then* $\neg H$ is true". With *modus ponens*, it now follows that "H is true" and "$\neg H$ is true", which obviously was not intended. So understanding the implication in the antecedent of the I–T conditional as a material implication just will not wash; after all, it cannot be argued that the final opinion will always be reached.

Misak repeatedly claims that the problem mentioned above disappears like snow in summer if the antecedent of the I–T conditional is a subjunctive conditional. Though he does not mention Misak, Crispin Wright (his critique is aimed primarily at Peirce's account of truth) attacks this claim in his *Truth and Objectivity* (1992). Wright sidesteps Quine's second dogma – the dogma of reductionism – by advocating a holistic approach to evidence: "[A]ny piece of information may, in the context of an appropriate epistemic background, be relevant to any particular belief." [Wright, 1992, p. 45.][37] This means, according to Wright, that we can only speak of "epistemically ideal circumstances" if we have *all the empirical information* at our disposal.

> [I]t is hard to see how a subject who somehow accomplished a Peircean state of comprehensive empirical information, could have any intimation that she had done so. By what principle could a subject discount the idea that there was still more to learn? But that reflection sets up a tension within any account of truth of the Peircean sort. For the idea that what is true is what a subject meeting certain conditions, C, would be in a position to acknowledge directly requires that a subject who was actually in conditions C – a subject of whom it was *true* that she was in conditions C – would be in position to acknowledge the fact. If such an acknowledgment would be impossible, then the antecedents of the subjunctive conditional which, on a Peircean view, explicate

[37] Contrary to Misak, Christopher Hookway seems to regard Peirce's account of truth as holistic: "Peirce defends a substantive theory of truth, which sees as true those propositions that enter into some ideally coherent body of opinions." [Hookway, 1985, p. 70.] Because, as we shall see, it has no bearing on the purpose of our argument whether Peirce's position is holistic or not, the issue will not be pursued here.

> what it is for a thought to be true, are uniformly false on purely conceptual grounds. Since the status of subjunctive conditionals with conceptually impossible antecedents is, by and large, extremely moot, that is bad news for Peircean views of truth. [Wright, 1992, p. 46]

In her recent 'Deflating Truth: Pragmatism *vs.* Minimalism' (1998), Misak defends herself against this ostensibly devastating criticism by saying that "the pragmatist can and should stay away from the ideas of total evidence and epistemically ideal conditions. Inquiry [...] is not to be thought of as global, complete inquiry, where every question is decided, including the question of whether inquiry is complete." [Misak, 1998, p. 413.] In short, Misak rejects Wright's argumentation, because Wright postulates a holism in regard to relevance that Misak does not share. However, Misak leaves the reader in the dark about arguments against this holism (a hiatus that seriously weakens her argument), but aside from that, she overlooks the fact that Wright's pitfalls also function without the postulation of holism.

I will now demonstrate that Misak's defence is incomplete to say the least, while I suspect that it can only be supplemented coherently if pragmatism is renounced. Let us assume, like Misak, that we can never be sure whether or not our inquiry relevant to an hypothesis H has been exhaustive:

> The pragmatist is a fallibilist and will simply agree that a person could never know that inquiry into a given question (never mind inquiry *tout court*) had been pursued as far as it could fruitfully go. [Misak, 1998, p. 414][38]

The pragmatist programme prohibits us to attribute any truth to the statement "inquiry has been pursued as far as it could fruitfully go", because after all, "what is incapable of being known is [...] not real at all" (*CW* 3, 319) and "*cognizability* (in its widest sense) and *being* are not merely metaphysically the same, but are synonymous terms" (*CW* 2, 208).[39] This cripples Misak's pragmatist account of truth: the assumption with which the main argument for the I–T conditional starts, namely that our inquiries relevant to hypothesis H have been absolutely exhaustive, immediately leads to a contradiction, from which everything follows, at least in standard logics. Each counterfactual with that same antecedent is therefore true, and so the sentences "*if*, if inquiry relevant to H were to be pursued as far as it could fruitfully go, then H would be believed, *then* H is true" and "*if*,

[38]Cf. "[S]ince the antecedent of this conditional is about what would be the case if inquiry were pushed indefinitely far, we can never assert that it is fulfilled." [Misak, 1991, p. 156.]

[39]See also: "[I]f in respect to some question – say that of the freedom of the will – no matter how long the discussion goes on, no matter how scientific our methods may become, there never will be a time when we can fully satisfy ourselves either that the question has no meaning, or that one answer or the other explains the facts, then in regard to that question, there certainly is no *truth*." (*CP* 5.565.)

if inquiry relevant to H were to be pursued as far as it could fruitfully go, then H would be *dis*believed, *then* H is true" are equivalent under Misak's assumption that we can never know whether our inquiries relevant to a certain question have been exhaustive. In short, even without referring to some form of holism, Wright's argument works.[40]

Of course, this attack may be parried with the claim that it is in fact possible to know whether we have reached the final opinion relevant to a certain question. This is the path Crispin Wright takes following his critique of Peircean theories of truth so as to give the skeleton of his 'minimalist' programme some flesh by constructing at least one possible truth predicate. Wright, who apparently subscribes to Putnam's statement that "the justification conditions for sentences change as our total body of knowledge changes, and cannot be taken as fixed once and for all" [Putnam, 1983, p. 85], sidesteps this problem by coming up with his own candidate – *superassertibility* – for a truth predicate:

> A statement is superassertible [...] if and only if it is, or can be, warranted and some warrant for it would survive arbitrarily close scrutiny of its pedigree and arbitrarily extensive increments to or other forms of improvement of our information. [Wright, 1992, p. 48]

Wright claims to be using the concepts 'information state' and 'improvement' in a purely formal sense, but fails – and that is the long and the short of it – to frame those terms in a formal theory. Many often crucial questions therefore remain unanswered: first, if Wright wishes his truth predicate 'superassertibility' to give out more than just the aura of a gratuitous suggestion, he will be required to answer the following question: given an information state σ, in which a proposition ϕ is warranted, how can it be demonstrated (or refuted) that for each improvement (whatever it may be) σ' of σ the proposition ϕ is still warranted?

In Quine's view, however, such a proof can never be provided: "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system." [Quine, 1953, p. 43.] If Quine and Wright were to solve this conflict of opinions, it is *not* a feasible option to forsake Wright's minimalist programme and resort to the endorsement of an overall philosophical system – in that way, the adversaries would only entrench themselves –; it would be preferable to investigate whether our methods of inquiry actually can warrant the permanence that Wright desires and that Quine detests: this is a problem of logic.

Secondly: suppose that a proof exists for the permanence of a given proposition ϕ, can this proof be extended to a statement about superassertible proposi-

[40] A second, if somewhat insipid objection lies in the annoying circumstance that the truth (Misak disguises this problem by speaking of *correctness* [Misak, 1991, p. 69]) of the I–T conditional, like the truth of all other statements, can only be determined once our inquiries into I–T have been completed, regardless of any convincing argument that may have been put forward presently.

tions independently of Wright's theory of the improvement of information states and of propositions warranted by information states? (For example: exactly all tautologies are superassertible, as they are in *update semantics* [Veltman, 1996].) In this case, what is the rationale of Wright's theory, if we can establish the characteristic to be defined independently of his proposals? Without an elaborate and well-argued formal theory of the phenomena that have been discussed, we are unable to judge Wright's argument, and as a result his considerations remain sterile.

Furthermore, isn't Wright just attempting to give an interpretation of the traditional static concept of truth?[41] Should we still be taking such a notion of truth seriously when attempting to develop a consistently naturalistic epistemology?

2.4 Conclusion

As we have seen, the strong need that many philosophers feel to interpret the traditional, static notions of truth and reality that are linked with the Aristotelian ideal of knowledge, cannot be satisfied by Peirce's theory of scientific inquiry. However, in spite of his incoherent theorizing in this matter, Peirce's attempts to understand human cognition from a fundamentally dynamic perspective deserve every credit, not lastly because they offer us an alternative for the Aristotelian ideal of knowledge. Before tackling the tricky themes of theoretical philosophy and rushing into definitions of the concepts of 'truth' and 'reality', it is advisable to gain a thorough understanding of the dynamics of human cognition. Only when we have articulated the mechanisms of the acquisition and revision of knowledge with sufficient accuracy, will we be able, on the basis of these formal-logical investigations, to provide a more than tentative answer to well-defined and detailed epistemological questions concerning the relation between 'warranted assertability' and 'truth'.

[41] "The idea of a statement whose complete and final warrant is wholly available to the speaker himself *no matter what happens* – or of a speaker who neither needs nor can benefit from the data of others – is precisely the old notion of knowledge which is private and incorrigible." [Putnam, 1987, p. 53.]

Chapter 3

Isaac Levi's Epistemology: A Critique

3.1 Isaac Levi's Epistemology

Isaac Levi, professor at the Columbia University of New York since 1970, is first and foremost a philosopher for technicality buffs: using a rather idiosyncratic terminology based on logic and probability theory, he puts forward his epistemological tenets in a voluminous series of publications. Levi defends a radical new perspective on a number of traditional epistemological issues, while firmly re-establishing the bond between logic and epistemology, which has been increasingly loosening during the past forty years. Furthermore, Levi is the main pioneer and initiator of contemporary logical-philosophical research into *belief change*. Levi places himself explicitly within the tradition of American pragmatism. Accordingly, some of the keynotes of Levi's epistemology can only be understood properly when placed against the background of American pragmatism.

3.1.1 Non-Deductive Logic and the *Belief-Doubt-Belief* Model

The lawyers and scientists who in January 1872 founded the 'Metaphysical Club', the cradle of American pragmatism, set out to combine the work of the British psychologist Alexander Bain, who defined 'belief' as "that upon which a man is prepared to act", with considerations concerning the philosophy of law and the theory of evolution in order to arrive at an overall theory of human thinking. From the very beginning, the members of the club, among which William James (1842–1910) and, albeit much later, Charles Sanders Peirce (1839–1914) would gather most laurels, rejected the ubiquitous tenet that true knowledge is to be modelled after mathematics.[1]

[1]The lawyer Oliver Wendell Holmes Jr., one of the six key members of the Metaphysical Club, writes in his *The Common Law* from 1881: "The law embodies the story of a nation's

One of the results of this heresy, which no American pragmatist would ever repudiate, was that the field of logic, as opposed to the logical inquiries of Frege and the early Russell, came to include not only mathematical reasoning, but, in Peirce's words, 'the method of scientific investigation' as well. John Dewey (1859–1952) later explicitly advocated an even wider domain for logic: the very logic of inquiry which leads to scientific results was also the methodological backbone of 'common sense'.[2] Levi, whose *continuity thesis* owes much to the American pragmatist tradition, is of the opinion that from a methodological point of view, the same mechanisms underlie both 'scientific inquiry' and 'practical deliberation':

> The difference between theoretical inquiry and practical deliberation is a difference in goals and not a difference in the criteria for rational choice that regulate efforts to realize these goals. [Levi, 1980, p. 73] [3]

The first consequence of this wider interpretation of the field of methodological inquiry was that the pragmatists, beginning with Peirce, placed non-deductive reasoning, such as induction and abduction, at a central place of logic. However, attempts of early American pragmatists to characterize non-deductive inferences remained informal until the publication of Dewey's methodological *magnum opus* *Logic: The Theory of Inquiry* in 1938. By that time, the necessary groundwork had been done in Europe. Bertrand Russell and Rudolf Carnap grappled in vain with formal solutions to justify, among others, generalizations: Russell's *Our Knowledge of the External World* (1914) en Carnap's *Der logische Aufbau der Welt* (1928) showed the failure of the endeavour to uncover the mechanisms by which we acquire knowledge of the world on the basis of observations and making use of the then recent instruments of mathematical logic and set theory. It emerged that the conditions under which tentative extensions of our beliefs, such as for example inductive inferences, are justified, could not be tackled with mathematical logic and set theory. In 1950, Rudolf Carnap, who had been teaching at the University of Chicago since 1936, presented in his *Logical Foundations of Probability* an epoch-making treatment of non-deductive reasoning based on probability theory: inductive logic was born. Isaac Levi was one among many working within this

development through many centuries and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics. In order to know what it is we must know what it has been, and what it tends to become." Quoted in [Kuklick, 1977, p. 50–51].

[2]According to Arthur Danto, this continuity thesis in the field of methodology is characteristic for *naturalism*: "Naturalism [...] is a species of philosophical monism according to which whatever exists or happens is natural in the sense of being susceptible to explanation through methods which, although paradigmatically exemplified in the natural sciences, are continuous from domain to domain of objects and events." Quoted in [Keil and Schnädelbach, 2000, p. 20]. Levi for his part actually sees the continuity thesis as the pivot of pragmatism: "What is 'pragmatic' about pragmatism is the recognition of a common structure to practical deliberation and cognitive inquiry in spite of the diversity of aims and values that may be promoted in diverse deliberations and inquiries." [Levi, 1991, p. 78.]

[3]See also [Levi, 1980, p. 16–19 and 23].

philosophical research program which was especially strong during the seventies and eighties.

The rejection of the axiomatic ideal as a standard for all knowledge[4] in combination with the acceptance of an evolutionist perspective, also had a second consequence. While many epistemologists, including logical empiricists, focused their attention mainly on the justification of the *results* of the acquisition of knowledge, concentrating on the rational reduction of these results to their origins (basic principles, *Protokollsätze*, or sensory stimuli), the American pragmatists chose a different perspective. Mainly through their agency, the process of *belief change* became a respectable subject for epistemological study. In Levi's opinion, *pedigree epistemology* – Levi's condescending expression for the epistemological enterprise of justifying our beliefs by tracing them back to their origins by means of a rational reconstruction – has proven to be a dead end. As an alternative program, Levi suggests to investigate under which circumstances a change of our current state of knowledge is justified:

> Whatever its origins, human knowledge is subject to change. In scientific inquiry, men seek to change it for the better. Epistemologists ought to care for the improvement of knowledge rather than its pedigree. [Levi, 1980, p. 1]

Levi, an advocate and pioneer of a normative approach in which adjustments of epistemic states are investigated using a *logical* apparatus, aims at formulating criteria under which a *change* of knowledge is also an *improvement*. It is not Levi's aim to describe how our knowledge actually changes,[5] but how it reasonably *should* change:

> The central problem of epistemology ought to be [...] to provide a systematic account of criteria for the improvement of knowledge. Alternatively stated, the problem is to offer a systematic characterization of conditions under which alterations in a corpus of knowledge are legitimate or are justified. [Levi, 1976, p. 1]

Although Levi was not the first to put knowledge change on the logical agenda, he is certainly the main initiator of its systematic investigation. Levi's philosophical forebears Peirce and Dewey already propagated a dynamic approach to

[4]Levi writes: "Following the tradition of Peirce and Dewey, I reject the requirement of self-certified first premises and principles for justifications of belief." [Levi, 1991, p. 4.]

[5]Since Thomas Kuhn's *The Structure of Scientific Revolutions* (1962), in which he defends a "new historiography of science", it is accepted practice to explain belief changes in terms of 'revolutionary' paradigm shifts. Context independent rational factors are supposed to only play a marginal part in the explanation of paradigm shifts, since paradigms before and after a scientific revolution appear to be incommensurable. Levi underplays the significance of these "changes in conceptual framework" – "there are no revolutionary changes or, at any rate, there should not be" [Levi, 1980, p. 68] – and actually wants to investigate, *given* a conceptual framework, into the criteria on the basis of which our beliefs ought to be changed and improved. Within such a conceptual framework all epistemic states are commensurable [Levi, 1991, p. 65].

knowledge with their *belief-doubt-belief* model, which should be understood as the first attempt at a logical description of the process of knowledge change. This model can be broadly outlined as follows: our actual state of knowledge forms a pattern of expectations, on which we base our actions. As long as we have no reasonable cause to doubt (parts of) this current epistemic state, it makes no sense to feign some Cartesian doubt, since after all there is "much that you do not doubt, in the least. Now that which you do not at all doubt, you must and do regard as infallible, absolute truth." (Peirce, *CP* 5.416.) Nevertheless, change is sometimes required. An experience "which really interferes with the smooth working of the belief-habit" (Peirce, *CP* 5.510), and which therefore, unlike an academic doubt, causes *true* doubt, shakes our opinions so that an inquiry becomes necessary to re-establish the lost equilibrium and to overcome the doubt caused by this unforeseen experience. This inquiry will lead to a new state of belief which, since it is the result of a procedure executed in accordance with the rules of methodology, will constitute a firm and sufficient basis for our thoughts and deeds, until an unforeseen experience forces us again to an inquiry.[6] While Peirce and Dewey focussed on a meticulous investigation of the criteria for knowledge change,[7] some of Levi's other predecessors achieved an informal description at most.[8]

---

[6]According to Levi, this model is "the greatest insight in the pragmatist tradition" [Levi, 1991, p. 163].

[7]Dewey's activities in the field of logic cover his whole philosophical career and find their culmination in his work *Logic: The Theory of Inquiry* (1938). His earlier logical-philosophical investigations are embodied in the *Studies in Logical Theory* (1903), in his more pedagogically oriented study *How We Think* (1910), and in the *Essays in Experimental Logic* (1916). [Burke, 1994] gives a survey and a defence of Dewey's studies in logic.

[8]See [James, 1907, p. 34–35]: "The individual has a stock of old opinions already, but he meets a new experience that puts them to a strain. [...] The result is an inward trouble to which his mind till then has been a stranger, and from which he seeks to escape by modifying his previous mass of opinions. He saves as much of it as he can, for in this matter of belief we are all extreme conservatives. So he tries to change first this opinion, and then that (for they resist change very variously), until at last some new idea comes up which he can graft upon the ancient stock with a minimum of disturbance of the latter, some idea that mediates between the stock and the new experience and runs them into one another most felicitously and expediently.

This new idea is then adopted as the true one. It preserves the older stock of truths with a minimum of modification, stretching them just enough to make them admit the novelty, but conceiving that in ways as familiar as the case leaves possible. An *outrée* explanation, violating all our preconceptions, would never pass for a true account of a novelty. We should scratch round industriously till we found something less excentric. The most violent revolutions in an individual's beliefs leave most of his old order standing. [...] New truth is always a go-between, a smoother-over of transitions. It marries old opinion to new fact so as ever to show a minimum of jolt, a maximum of continuity."

See also [Quine, 1953, p. 42]: "[T]otal science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Reëvaluation of some statements entails reëvaluation of others, because of their logical interconnections – the logical laws being in turn simply certain further statements of the system, Five years after the publication of Levi's *The Enterprise of Knowledge* (1980), in which formal criteria for rational revisions of epistemic states are defended, Alchourrón, Gärdenfors and Makinson published a very elegant formalization of some of the logical prerequisites for Levi's ideas on knowledge change: the harmonious overture to what has by now built up to a discipline of symphonic proportions within the philosophy of logic, a discipline that is designated by the terms 'belief revision', 'belief change' and 'theory change'.[9]

In the past fifteen years, logical research into *belief change* has boomed, producing a considerable amount and diversity of formal systems. Although there have been elegant and 'deep' results and successful attempts at partial systematization,[10] the reasons and motives for the construction of many systems frequently leave much to be desired: often one tiny little problem gives rise to yet another new system.[11] As Levi is not only an initiator of research into *belief change*, but also its most philosophical advocate – Levi is almost the only one to propound a philosophical embedding for formal theories of belief change[12] –, we will subject Levi's proposals for the modelling of belief change to a critical examination, hoping that the conclusions that will be reached can be extrapolated to competing systems for *belief change*, so that we can get a clearer picture of this branch of philosophical logic.

3.1.2 Epistemic States and Their Representations

Before we actually can write down anything sensible about the mechanisms of belief change, we must first know *what* is supposed to change. Levi distinguishes between *epistemic states* ('states of full belief'), our true states of knowledge on the one hand, and *representations of epistemic states* ('corpora') on the other. Although in his *The Fixation of Belief and Its Undoing* (1991), Levi has gone to many lengths to define and defend his preferred notion of 'epistemic state',[13] we shall adopt Levi's own policies for treating contraction and conditionals, and concentrate on Levi's proposals concerning the *representations* of epistemic states

certain further elements of the field. Having reëvaluated one statement we must reëvaluate some others, which may be statements logically connected with the first or may be the statements of logical connections themselves."

[9] [Alchourrón, Gärdenfors, and Makinson, 1985] For an introduction, see [Gärdenfors, 1988] and [Hansson, 1999c].

[10] See [Rott, 1991], [Rott, 1992], and [Rott and Pagnucco, 1999].

[11] See for instance [Fermé and Hansson, 1999], in which the system, that, unlike other systems of *belief change*, is able to accept a *part* of the new information, is illustrated with the following example: "One day when you return back from work, your son tells you, as soon as you see him: 'A dinosaur has broke grandma's vase in the living-room'. You probably accept one part of the information, namely that the vase has been broken, while rejecting the part of it that refers to a dinosaur." [Fermé and Hansson, 1999, p. 331.]

[12] [Friedman and Halpern, 1999] outlines two alternative 'embeddings'.

[13] See especially [Levi, 1991, §§ 2.1–2.4].

as well as their dynamics.

It goes without saying that a choice for certain types of representations of epistemic states has far-reaching consequences. Although there is some contention about the most suitable kinds of representations,[14] the similarities between the positions defended by logicians overrule the differences. Most protagonists, including Levi, share the presupposition that epistemic states should be represented by *structured sets of descriptive sentences.*

Why has this idea been such a resounding success? The promotion of this tendency to represent our knowledge as a structured set of descriptive sentences was mainly due to the leading role that philosophy of science, which was grafted on neo-positivism, played in the development of epistemology in the twentieth century. Originally, the members of the Vienna Circle subscribed to the traditional ideal of knowledge, although they gradually gave up the requirement to provide a *Begründung* of science. Nevertheless, the Aristotelian ideal of knowledge[15] has in its fall not dragged down the idea that our knowledge ultimately forms a coherent whole of descriptive sentences. The fall of the Aristotelian idea has only fundamentally changed this idea. This coherence consists and consisted of 'inferential' relations between descriptive sentences. During the twentieth century, both the propositional attitude towards sentences included in a system of knowledge and their inferential relations have been reassessed. If it was formerly thought that all *true sentences* could be assigned a specific place in the fabric of our knowledge by working out which axioms and corollaries (or in the case of logical empiricists such as Moritz Schlick, which *Konstatierungen*) were needed to justify them with the help of the canons of reasoning mechanisms, now we give preference to the metaphor of a *web of belief* in which *sentences which are held to be true* are ordered according to their "relative likelihood, in practice, of our choosing one statement rather than another for revision in the event of recalcitrant experience".[16]

Secondly, there is the *holistic* approach of meaning and cognition, propagated by Quine in particular,[17] which has led many epistemologists to be inclined to think that an epistemic state is an idealization of our 'theory of the world', where,

[14]For an already somewhat dated survey, see [Gärdenfors, 1988, p. 21–46].

[15]See [Beth, 1959, p. 31–32] for a detailed discussion.

[16][Quine, 1953, p. 43.]. In recent investigations into *belief change* we come across a similar idea under the name 'epistemic entrenchment'. See among others [Gärdenfors and Makinson, 1988], [Rott, 1991], and [Levi, 1991, § 4.7].

[17]In his famous 'Two Dogma's of Empiricism', Quine discredited the assumption that each meaningful sentence is equivalent to a logical-mathematical construct of observational terms, an assumption that was shared by Peirce and most logical empiricists. Instead, Quine proposes to consider *theories* instead of sentences as the primary carriers of meaning, and then to try to establish the empirical meaning of theories on the basis of the relations between theories and empirical data. Levi also stresses that "in the first instance it is not sentences or other linguistic entities that carry truth value and informational value but potential states of full belief." [Levi, 1996, p. 53.]

obviously, a theory consists of descriptive sentences. The logical empiricist ideal of an *Einheitswissenschaft* has only reinforced that inclination. So, Quine speaks of '[t]he totality of our so-called knowledge or beliefs', 'our own particular world-theory', and 'total science'.[18]

In the third place, there is an influential methodological reason for representing epistemic states by structured sets of descriptive sentences. First, in epistemological inquiry, it is possible to limit oneself, for instance on the basis of Quinean behaviourism, to the part of our knowledge that can be put into words, without committing oneself to the point of view that *all* our knowledge can be thus articulated. Likewise, it seems advisable in epistemology to start by investigating only those beliefs which can be put into a *descriptive* sentence, so that recently booked results in logic can be applied in epistemological analyses, without committing oneself to the point of view that *all* knowledge that can be put into words can also be straitjacketed into a descriptive sentence. This *reculer pour mieux sauter* – to be able to tackle epistemological problems with contemporary logic we will limit ourselves to the part of knowledge which can be expressed in descriptive sentences – is a popular strategy among formally oriented philosophers of science, logicians and researchers in the field of artificial intelligence.

So it is only natural that Levi proposes to represent the knowledge of a certain agent X that can be expressed in some (formal) language L as a deductively closed set K of sentences in L.[19] Levi calls such a deductively closed set of sentences a *corpus*.[20] Within a corpus two classes of sentences can be distinguished.

In the first place, a corpus contains sentences that X will not give up under any circumstance. For instance, those sentences that articulate the metaphysical, ontological and (classical) logical presuppositions to which X is committed. Levi labels the set of sentences which, at least for X, do not qualify for revision, as X's *urcorpus*. It forms the kernel of each corpus of X and includes at least the criteria on the basis of which proposals for changing X's corpus are judged. So the *urcorpus* consists at least of "those assumptions which any corpus should have if an account of the revision of knowledge [...] is to stand a chance of working." [Levi, 1980, p. 7.] In addition to (classical) logic, mathematics and set theory, the *urcorpus* contains a 'conception of error' [Levi, 1980, p. 8] and, I assume, a 'conception of informational value'.[21]

In the second place, X's corpus contains sentences of which X does not rule

[18][Quine, 1953, p. 42]; [Quine, 1960, p. 24]; and [Quine, 1953, p. 42].

[19]Levi assumes that L is rich enough to express arithmetic and set theory and that it complies with (classical) first-order logic. L should not be understood as the language "that the agent uses or would use to communicate his convictions or other attitudes." [Levi, 1991, p. 33.]

[20][Levi, 1980, p. 4.] The problem of logical omniscience, a consequence of the choice for *deductively closed* sets of sentences, is discussed in [Levi, 1976, p. 22–23] and in [Levi, 1980, p. 9–12]. See also [Levi, 1991, § 2.1], in which Levi, by putting forward a distinction between 'doxastic commitment' and 'doxastic performance', distances himself from the naturalist program in epistemology defended notably by Quine.

[21]In the formalizations of Levi's theory of belief change, the *urcorpus* is represented by $Cn(\emptyset)$.

out that they might one day qualify for revision, though they are, at present, infallible.[22] These sentences may be (negations of) singular statements, but also laws, theories, and statistical claims. The fact that X considers a statement ϕ in her corpus at time t to be susceptible to revision, does not alter the degree to which ϕ is considered probable by X:

> From X's point of view at t, every theoretical assumption, statistical claim, universal generalization and observation report in his corpus at t is as certainly and necessarily true as any truth of logic – at least as far as the conduct of practical deliberations and scientific inquiry are concerned. [Levi, 1976, p. 24]

If we now follow Levi's suggestion and represent the 'credal probability' that X attributes to sentences by a function Q complying with the standard axioms of probability theory, such that Q is defined for all sentences in the language L, the above can be summed up as follows: for all sentences ϕ in X's corpus K it holds that $Q(\phi) = 1$.

Our current corpus is, according to Levi, the only standard for what we, at least for the time being, hold possible. *Logical* possibilities form much too large a class for a workable concept of possibility:

> It seems clear that in daily life and scientific inquiry, we discount utterly all sorts of logical possibilities. We do not assign them small probabilities of being true. [Levi, 1976, p. 12.]

Furthermore, it is for Levi a "prima facie obvious fact" that, in scientific inquiry and practical deliberation, X must consider *all* the elements from its present corpus to be certain and infallible. The tenets that our current corpus determines what we hold possible and that we consider all the elements from our current corpus to be certain and infallible are defended on the basis of Levi's definition of the concept 'serious possibility': a sentence ϕ is a serious possibility with respect to a corpus K if and only if ϕ is consistent with K [Levi, 1980, p. 5].[23] A corollary of this interpretation of possibility is that each element ϕ in K is necessary and therefore infallible, since $\neg\phi$ is not a serious possibility with respect to K. Levi sums up both claims with his thesis of *epistemological infallibilism*.[24] In short,

> X is committed to treating all items in the corpus of knowledge he adopts at t as infallibly true in the sense that the logical possibility that one of the items is false is not, as far as he is concerned, a serious one. [Levi, 1976, p. 7]

[22]See the quote of Peirce's (*CP* 5.416) on page 32 of this chapter.

[23]In 1980, Levi defines a corpus K in terms of the set of serious possibilities. In 1996, Levi defines the set of serious possibilities in terms of a corpus K. Logically speaking, these are two sides of the same cookie – see [Levi, 1996, p. 45].

[24]See [Levi, 1980, p. 13].

Epistemological infallibility, however, does not imply that our current corpus K, our one and only standard for serious possibility is impervious to deliberate change: "Certainty does not imply incorrigibility." [Levi, 1991, p. 3.] With good reason, our current corpus can be changed and improved in order to arrive at another corpus K' which then will become our one and only standard for serious possibility. In short, knowledge is corrigible, even though we consider it to be infallible when we have no reason to change it.[25]

In his 'Knowledge and Belief' from 1952, in which he provides a reinterpretation of the traditional distinction between 'knowledge' and 'belief', Norman Malcolm uses a concept of knowledge closely related to Levi's. The omission of the latter to breathe life into his rather formal conception of knowledge with a number of convincing illustrations is compensated by the five real-life examples which Malcolm puts forward to make clear that the answer to the question "Can I discover *in myself* whether I know something or merely believe it?" [Malcolm, 1952, p. 69] must be in the negative.[26] Let us take a closer look at two of his examples:

> Suppose, for example, that several of us intend to go for a walk and that you propose that we walk in Cascadilla Gorge. I protest that I should like to walk beside a flowing stream and that at this season the gorge is probably dry. Consider the following cases: [...]
>
> (4) You say 'I know it won't be dry' and give a stronger reason, e.g., 'I saw a lot of water flowing in the gorge when I passed it this morning'. If we went and found water, there would be no hesitation at all in saying that you knew. [...]
>
> (5) Everything happens as in (4), except that upon going to the gorge we find it to be dry. We should not say that you knew, but that you *believed* that there would be water. And this is true even though you declared you knew, and even though your evidence was the same as it was in case (4) in which you did know. [Malcolm, 1952, p. 69–70]

According to Malcolm, these examples show that "*although you knew you could have been mistaken*" [Malcolm, 1952, p. 71]. Malcolm thinks that it is surely possible for a statement that we consider to be an 'absolute certainty' at present, for example "There is a heart in my body", to turn out to be false on closer examination and, hence, to be eligible for correction.[27] So, absolute certain knowledge and corrigibility are not mutually exclusive. That's all very well, but when can a change of our absolute certain knowledge be called an improvement?

[25]See [Levi, 1980, p. 18].

[26]Levi writes: "In my opinion, there is no relevant difference, from X's point of view at t, between what he knows and what he fully believes" [Levi, 1976, p. 5]. On the relation between 'knowledge' and 'full belief', see also [Levi, 1991, p. 45].

[27]See [Malcolm, 1952, p. 76].

3.1.3 Revision of Corpora

Now that we have represented an epistemic state by a corpus – a deductively closed set of sentences from a language L – we can start thinking of belief changes or revisions of corpora in terms of "shifts from one deductively closed set to another." [Levi, 1976, p. 23.] Levi distinguishes two fundamental types of revision, namely *expansion* and *contraction*. He claims that these types are fundamental because all other kinds of revision of corpora can be understood as a series of expansions and contractions [Levi, 1980, p. 65].[28] Subsequently, Levi concentrates on articulating the conditions under which these two basic types of revision are justified, starting from the following consideration:

> The kind of cognitive aim that, in my opinion, does best in rationalizing scientific practice is one that seeks, on the one hand, to avoid error and, on the other, to obtain valuable information. [Levi, 1996, p. 51]

Consequently, the starting-point of Levi's ideas on belief change is formed by the twin concepts of 'informational value' and 'credal probability'.

Expansion

In expansion, a sentence ϕ is added to a corpus K. For the sake of convenience we will denote the result of such an operation by '$K + \phi$'. Logically speaking, an expansion doesn't amount to much: just take the union of the sets K and $\{\phi\}$ and close that union under deduction. In short,

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

It is plain that this definition does not answer the question under which conditions an expansion is an *improvement* of our corpus. It only indicates how we should change our current corpus once we have decided to expand it with the sentence ϕ. It tells us nothing about the reasonableness of such a decision. Unlike the great majority of researchers in the field of *belief change*, Levi formulates up a standard on the basis of which the legitimacy of the decision to implement an expansion can be judged.[29] Most of Levi's ideas on expansion stem from the monograph

[28] Consequently, in belief change literature, *revision* is usually defined in terms of a *contraction* and an *expansion*: $K \times \phi = (K - \neg\phi) + \phi$. This definition of revision is called the *Levi identity*. See [Gärdenfors, 1988, p. 69].

[29] Friedman and Halpern rightly complain about the fact that in the bulk of belief change literature no-one takes the trouble to investigate into the conditions under which the addition of a sentence to an epistemic state is legitimate, though "deciding when a formula has come to be accepted is nontrivial. [...] Acceptance has a complex interaction with what is already believed." [Friedman and Halpern, 1999, p. 404.] Levi notes the same shortcoming: "The absence of an account of the conditions under which expansion is justified is a serious lacuna in a theory of rational belief change." [Levi, 1996, p. 6.] See also Levi's remarks on Gärdenfors's work on expansion [Levi, 1991, p. 44 and § 3.6].

Gambling with Truth (1967).

Levi distinguishes two types of expansion, namely *deliberate expansion* and *routine expansion*. Both types are necessary to acquire new information. In routine expansion, an external stimulus is converted into a sentence via a previously adopted 'program'. The resulting sentence is then indiscriminately added to the corpus of the agent – what Levi has in mind here is making observations or consulting a witness or an expert. Although we only accept a program for routine expansion if we consider it to be reliable, a hundred percent reliability is an unreasonable demand. Therefore, an accepted program can inject information into our current corpus which is inconsistent with our current corpus, whereby the corpus resulting from the expansion becomes inconsistent and, hence, trivial, since Levi closes corpora under classical logic. In short, a routine expansion implemented according to the rules can unintentionally lead to the inconsistent corpus.[30] Further on, we will see that a correctly implemented deliberate expansion does not suffer from this deficiency.[31]

In deliberate expansion, an agent chooses one sentence from a series of alternatives and then adds it to his corpus. Let us now take a look at what the technical ins and outs of this type of expansion are. The need for a deliberate expansion of our present corpus does not simply come out of the blue – we add new information to our corpus only for a certain purpose. What is that purpose and how do we serve it best? To clarify the issue, Levi notes down the following considerations concerning 'deliberate decision making':

> In deliberate decision making, the agent identifies the options available to him, his goals, and the available relevant evidence concerning the admissibility of the options for the purpose of realizing these goals and values. The option chosen is determined relative to these beliefs and values according to principles of rational choice. [Levi, 1980, p. 36]

Several aspects of our question about the conditions under which an expansion is legitimate can now be specified:

> The options are potential expansion strategies which qualify as potential answers to the question under investigation, and the aim is to gratify the demand for information occasioned by the question while at the same time avoiding error. [Levi, 1980, p. 38–39]

So the aim of an expansion is answering a question with 'new error-free information'. However, in case ϕ as well as $\neg\phi$ are serious possibilities with respect to

[30] See for a more detailed discussion of routine expansion [Levi, 1991, § 3.4].

[31] "[I]f one is living up to one's commitments, one cannot legitimately expand into inconsistency via deliberate expansion. On the other hand, routine expansion can and sometimes does lead to inconsistency even when all commitments are fully met." [Levi, 1991, p. 76.]

K, there is always the risk that if we expand our corpus K with ϕ, we allow a false sentence into our corpus. According to Levi, such an expansion is justified if and only if the information value of ϕ outweighs the risk that ϕ is false:

> On the basis of inquiries [...] we sometimes reach a point where we conclude that the trade offs between risk of error and informational benefits are such as to warrant adding some hypothesis to the corpus and so to convert its status from mere hypothesis to settled, established and infallible truth (where being settled, and established is only for the time being and not necessarily forever). [Levi, 1976, p. 15]

In order to fulfil the aim of getting relevant 'new error-free information' as best as possible, we should ideally proceed as follows: if, given our current corpus K, we have to deal with a problem, we first identify, in a phase which Levi calls 'abduction',[32] all seriously possible problem-solving options; then we trade off the *informational value* and the *credal probability* of all the available options; and finally, we implement an expansion of our current corpus with those subset of options which have come up as the best during this weighing procedure. How does it all fit together formally?

Levi calls the set U of all available options that solve a given problem an *ultimate partition.* Let $U = \{\phi_1, \ldots, \phi_n\}$ be a finite set of available options.[33] U is exhaustive and exclusive. All alternative options ϕ_i in U are serious possibilities with respect to the current corpus K. No single option ϕ_i is an element of K. Moreover, K implies the truth of exactly one element in U, though we do not know which.[34] Levi defines a potential answer as the rejection of a subset R of U. A potential answer can be formulated with a sentence ρ, where ρ stands for the disjunction of all alternative options in R. After choosing a potential answer, agent X should expand his corpus K with $\neg\rho$, that is, with the statement that the correct answer in U is *not* in the subset R of rejected elements from U.[35]

As noticed, we need to balance the risk we take of admitting a false sentence into our corpus when choosing a potential answer ρ against what the choice of

[32]Levi stipulates: "Abductive logic [...] is a system of norms prescribing necessary conditions which a system of potential answers to any legitimate question should satisfy." [Levi, 1976, p. 33] The assessment of the informational values of the potential answers is also the result of abduction. See [Levi, 1980, p. 49] and [Levi, 1998, p. 4].

[33]Levi also discusses, although summarily, infinite sets of options. See [Levi, 1976, p. 41–42] and [Levi, 1980, p. 49]. The technical problems raised by infinite sets of options are irrelevant for the purpose of my argument.

[34]See [Levi, 1967a].

[35]There are two limiting cases. On the one hand, X can decide not to reject any option in U. In that case, the expansion of K with the statement that the correct answer is *not* to be found in the (now empty!) set of rejected alternatives in U leaves the corpus K as it was. On the other hand, X may decide to reject all alternatives in U. Expansion of K with the statement that the correct answer in U is not to be found in the set of rejected alternatives now produces the inconsistent corpus.

ρ brings us, namely new information: after expansion with $\neg\rho$, we indeed *know* that $\neg\rho$. In order to represent this trade off between informational value and credal probability numerically, Levi attaches to both informational value and credal probability a *separate* probability measure. (Levi maintains that these two measures should not be reducible to each other.) So, the *informational value* is fixed with an 'information-determining probability measure' M, and the *credal probability* with a second probability measure, the 'expectation-determining probability measure' Q. (We have already met the latter at our discussion of Levi's notion of a corpus.)

The probability measure M assigns to each sentence ϕ_i in U a probability $M(\phi_i)$, such that $0 \leq M(\phi_i) \leq 1$ and $M(\phi_1) + \ldots + M(\phi_n) = 1$. This probability measure means to represent X's (context dependent) evaluation of the *informational value* of the available options, but says nothing about X's assessment of the *credal probability* of these options: $M(\phi_i)$ is the informational value of *rejecting* ϕ_i [Levi, 1980, p. 48]. Hence, "$1 - M(\phi_i)$ is the informational utility or value of adding ϕ_i to X's corpus [...] when considerations of truth value are ignored." [Levi, 1976, p. 38 – adapted notation.]

In turn, the measure Q assigns a probability $Q(\psi)$ to each sentence ψ in the language L, such that $0 \leq Q(\psi) \leq 1$. This probability measure is meant to represent X's assessment of the *credal probability* of the available options, but tells us nothing about X's evaluation of the *informational value* of these options. It fixes X's 'credal state', a supplement to X's corpus of knowledge:

> [R]elative to his corpus of knowledge X has a 'credal state' represented by a probability function assigning to all sentences in L a numerical probability consistent with the requirement that all items in his corpus bear probability 1. [Levi, 1976, p. 37]

Lastly, the utilities $M(\phi_i)$ and $Q(\phi_i)$ of each option ϕ_i have to be traded off, weighed by a 'degree of boldness' q. This degree of boldness, which, though it always holds that $0 < q \leq 1$, is context-dependent (as we shall see later on), represents the degree to which X is prepared to risk errors in order to acquire new information. Levi's assumptions and argumentations, based on an approach via a maximization of expected epistemic value, finally lead to the following criterion for the choice of an expansion strategy,[36] a criterion that we will designate from now on with 'Rule A':

> Given a corpus $K_{X,t}$, finite partition U, information-determining probability function M defined over the Boolean algebra of elements of U, an expectation-determining probability function Q defined over the

[36]For technical details, see especially [Levi, 1967b].

> same algebra, and an index of caution q, X should reject all and only those elements of U satisfying $Q(\phi_i) < qM(\phi_i)$. [Levi, 1980, p. 53][37]

If the potential answer $R = \{\phi_{i1}, \ldots, \phi_{im}\}$ is the set of options which is rejected on the strength of the abovementioned criterion, then, if we subscribe to Levi's proposals, the expansion of X's current corpus K with $\neg(\phi_{i1} \vee \ldots \vee \phi_{im})$ is legitimate. The result of this expansion, X's new corpus, is then given by $K + \neg(\phi_{i1} \vee \ldots \vee \phi_{im})$, which will serve as X's new standard for serious possibility:

> To be sure, prior to expansion, there is a risk, from X's point of view, that the information to be added to his standard for serious possibility is false. Yet, sometimes X is justified in taking the risk. Once X has implemented the expansion strategy and taken the risk, he evaluates serious possibility according to a new standard relative to which the new information added is no longer possibly false. [Levi, 1980, p. 57]

Contraction

In contraction, a sentence ϕ is deleted from a corpus K, such that ϕ is not a logical consequence of the remaining sentences in the corpus resulting from the contraction of K with ϕ. For the sake of convenience, we will use '$K - \phi$' to denote the resulting corpus.

Other criteria apply for contractions than for expansions. As opposed to what is the case in expansions, avoiding error cannot be a reason for deleting a sentence ϕ from a corpus K, as all the sentences in a corpus K, which after all acts as X's standard for serious possibility, cannot possibly be false: "In contraction, the concern to avoid error is vacuous."[38] On the contrary, in a contraction X gives up a sentence which is definitely true: "For X to contract his corpus is for him to surrender error-free information." [Levi, 1980, p. 58.] Hence, the *credal probability* of the sentences in K can play no part in the formulation of a criterion for legitimate contractions. Levi intends to formulate a theory of contraction that "seeks to show how a consistent account of justified ceasing to believe is feasible even when K is taken to be a standard for serious possibility and all members of K are true in the sense in which avoidance of error is taken to be a desideratum of efforts to improve K by revising it." [Levi, 1991, p. 61.]

Levi's epistemology only allows for two reasons for a contraction. First, as we have indicated briefly above, it is possible to accidentally end up in the inconsistent corpus via a legitimate routine expansion of a consistent corpus K with a sentence ϕ. Because the inconsistent corpus "fails as a standard for serious possibility to be used in inquiry and deliberation" and therefore is of no value

[37]Strictly speaking, the Q-function is not only defined for all boolean combinations of elements in U, but for all elements of the language L.

[38][Levi, 1991, p. 79.] See [Levi, 1998, p. 50].

whatsoever,[39] an agent is obliged to once again arrive at a consistent corpus by means of a coerced contraction:[40]

> When routine expansion injects inconsistency into the inquirer's doctrine, contraction from the inconsistent state is required. An inconsistent state of full belief or corpus fails as a standard for serious possibility for the purpose of subsequent inquiry and for practical deliberation. [Levi, 1991, p. 76–77]

That we need to implement a contraction if we have landed in the inconsistent corpus is beyond dispute. In a coerced contraction, we can restrict ourselves to determining the strategies to extricate ourselves from the inconsistent corpus. The inconsistent corpus was reached by expanding a consistent corpus K already containing the sentence $\neg\phi$ with a sentence ϕ which was obtained via a program held to be reliable. Hence, according to Levi, we can do either of three things:[41] (1) we may call into question the reliability of the program which resulted in the sentence ϕ which was inconsistent with our old corpus K. In this case, we go back to the old corpus K, from which we delete with contraction the claim that the program in question is reliable; (2) we may doubt the background information present in the old corpus which is inconsistent with the sentence ϕ obtained by means of the program. In this case, we remove background sentence $\neg\phi$ with contraction from K and expand the result with ϕ; (3) we may refuse to believe both the program and the relevant background information.[42] In the last case, we take the intersection of the corpora obtained by way of the first two strategies.[43]

[39]Elsewhere Levi writes: "To allow X to consider a contradictory corpus to be feasible does not imply that if he should detect inconsistency in his corpus he should rest content. When X's corpus is inconsistent, it breaks down as a standard of serious possibility. It furnishes a truth definition which is unsuitable for characterizing the aim of avoiding error. It is useless as a resource for inquiry and deliberation." [Levi, 1980, p. 27–28.] Hence, "it is always urgent to contract from an inconsistent state of full belief." [Levi, 1991, p. 68.]

[40]Situations in which one arrives at the inconsistent corpus via a routine expansion "furnish one of the occasions that justify contraction through coming to doubt the information obtained via routine expansion or some other item in the initial corpus or, as I think is normally sensible, both." [Levi, 1991, p. 76.]

[41]Since Levi bases his system on an underlying classical logic which he considers immune to revision, he cannot account for a fourth possibility: an adjustment of the underlying logic. Von Neumann and others argued that the reconciliation of the particle theory and the wave theory of light via Bohr's *principle of complementarity* did not imply a weakening of one of the fundamental principles of the rival theories, but actually a *weakening* of the underlying logic. See [Beth, 1964, p. 8–10].

[42]Although an inconsistent corpus contains all the sentences of L, it apparently does not eat away at our memory and our powers of judgment. After all, one seems not to forget from which corpus the inconsistency is reached, while the corresponding informational values which will turn out to be necessary to implement the said contraction are left undisturbed. How inconsistent is an inconsistent corpus?

[43]A more comprehensive account of coerced contractions can be found in [Levi, 1991, § 4.8].

In the second place, an *uncoerced* contraction comes into consideration when we decide to give 'a hearing' to a hypothesis T_2, which is contradicted by an element T_1 from the current corpus K. Because it initially holds that T_2 is not a serious possibility with respect to K, elements from K have to be deleted, so as "to shift to a position where judgment is suspended between these rival hypotheses so that investigations can be undertaken to decide whether T_1 should be reinstated via inferential expansion or T_2 should take T_1's place." [Levi, 1980, p. 60.] Not every hypothesis qualifies for such a procedure. "There must be some inducement to incur the loss of information" [Levi, 1991, p. 118]:

> To be justified in ceasing to believe what is initially settled, the inquirer must regard the benefits of giving the new proposal a non-question-begging hearing to be great enough to outweigh the costs. [Levi, 1991, p. 4]

For example, a hypothesis T_2, which, though it be incompatible with our current views, gives an explanation of anomalies – phenomena that cannot (yet) be explained by our current theories, whereas they should[44] – is worth considering.[45] The actual corpus K, which contains $\neg T_2$, prohibits an unprejudiced evaluation of T_2. If we still wish to make a fair evaluation possible between T_2 and $\neg T_2$, it is necessary to adapt our actual corpus in such a way that both T_2 and $\neg T_2$ are serious possibilities with respect to the adapted corpus. According to Levi, $K - \neg T_2$ is the best corpus for the intended evaluation, because it differs only minimally from our current corpus. Now, on the basis of this adapted corpus $K - \neg T_2$, using the criteria for expansion discussed above, we can check without prejudice whether T_2 or its negation should be added to the adapted corpus.[46]

After this brief outline of the circumstances that justify the deletion of certain sentences from our current corpus, we shall conclude our discussion of contraction with the contraction method propagated by Levi, in which the central question is: supposing that we wish to remove a sentence ϕ from a given corpus K, *how* should we implement this contraction? Levi answers:

> We need to identify the available options or strategies for contraction by removing ϕ and then examine the goals and values that ought to be promoted in order to decide among them. [Levi, 1991, p. 121]

We will not be able to avoid a modest logical apparatus in order to grasp the technical details of Levi's ideas on contraction.[47] Given the aim of a contraction

[44] See [Levi, 1991, p. 152–153].

[45] Levi writes: "[A] good reason for implementing an uncoerced contraction would be that it allows a promising theory incompatible with current doctrine to be examined without prejudice." [Levi, 1991, p. 153.]

[46] For more information, see [Levi, 1991, § 4.9].

[47] [Hansson and Olsson, 1995] is an excellent study of Levi's contraction operators.

of a corpus K with a sentence ϕ, we can immediately impose three constraints on $K - \phi$, the result of this contraction: (1) $K - \phi$ is a corpus, a deductively closed set of sentences; (2) $K - \phi$ is a subset of K, since a sentence is removed from K; and (3), the sentence ϕ which needs to be removed is *not* an element of $K - \phi$. We shall use $C(K, \phi)$ to denote the set of corpora that meet these three requirements:

$$C(K, \phi) \quad = \quad \{K' \subseteq K : K' = Cn(K') \textit{ and } \phi \notin K'\}.$$

Since each element of $C(K, \phi)$ is the intersection of one or more elements of the set $S(K, \phi)$ of 'saturatable' contractions which skip a sentence ϕ from a corpus K,[48] we can stick without loss of generality to the set $S(K, \phi)$:

$$S(K, \phi) \\ = \\ \{K' \subseteq K : K' = Cn(K') \textit{ and } Cn(K' \cup \{\neg\phi\}) \textit{ is maximally consistent in } L\}.$$

The set of saturatable contractions $S(K, \phi)$ is used by Levi as a starting-point for his definition of the contraction of a corpus K with a sentence ϕ: after all, each possible contraction of K with ϕ is the intersection of the elements of a subset of $S(K, \phi)$. Hence, the problem of finding the right contraction can be reduced to the question which elements of $S(K, \phi)$ we should choose for this subset. Since each subset of $S(K, \phi)$ meets the logical contraints imposed on a contraction, logical considerations alone will fall short if, like Levi, we require that "when all relevant factors in a given context are taken into account, one change at most should be legitimate or justified. Hence, given the initial corpus K and *all other relevant factors* (whatever these may be), and given that adding or deleting ϕ is legitimate or justified, the new belief state to which one shifts legitimately or with justification is uniquely determined."[49]

Since, at the present state of logical research, no purely logical definition of a sensible contraction operator can be given within classical logic,[50] Levi resorts to an information-determining measure M which assigns an M-value $M(K')$ to all possible contractions K' in $C(K, \phi)$ and so solves the problem of making a well-founded choice from the elements of $S(K, \phi)$ in order to define the required unique subset of $S(K, \phi)$. In this way, the uniqueness of the contraction of K with ϕ is warranted,[51] as Levi defines this contraction as the intersection of all

[48] See [Levi, 1991, p. 122] and [Levi, 1996, p. 20]. The elements of $S(K, \phi)$ are *saturatable*, because for each element K' in $S(K, \phi)$ it holds that $Cn(K' \cup \{\neg\phi\})$ is maximally consistent in L.

[49] [Levi, 1991, p. 67 – adapted notation.]

[50] These results have been summarized in Observation 2.1 of [Alchourrón and Makinson, 1982] and in Proposition 5.3 of [Hansson and Olsson, 1995].

[51] Levi starts from "the normative assumption that there should not be more than one admissible change of belief state in a given context" [Levi, 1991, p. 180].

the elements of the said subset of $S(K, \phi)$. Levi advises the agent who is planning to contract a sentence ϕ from her corpus K to use this information-determining measure M as follows:

> [T]he inquiring agent should evaluate the various contraction strategies available to her with respect to the informational value incurred and should choose a contraction strategy that minimizes the loss of informational value if a minimizing strategy exists. [Levi, 1991, p. 122]

In short, choose those saturatable contractions from $S(K, \phi)$ which have the highest informational value. If we now assume that an information-determining measure M assigns to all deductively closed subsets of K' an informational value $\mathit{Cont}(K') = 1 - M(K')$, only requiring that $A \subset B$ implies $\mathit{Cont}(A) \leq \mathit{Cont}(B)$,[52] then the needed selection function γ, which chooses the elements from $S(K, \phi)$ with the highest informational value, can be defined as follows: if $S(K, \phi) = \emptyset$, then $\gamma(S(K, \phi)) = \{K\}$; if $S(K, \phi) \neq \emptyset$, then

$$\gamma(S(K, \phi)) \\ = \\ \{K' \in S(K, \phi) : \textit{ for all } K'' \in S(K, \phi) \textit{ it holds that } \mathit{Cont}(K'') \leq \mathit{Cont}(K')\}.$$

The formal apparatus developed in this subsection allows Levi to determine the admissible contraction $K - \phi$, given a corpus K, a sentence ϕ which has to be deleted and an informativity measure M over $C(K, \phi)$, by means of the set of saturatable contractions $S(K, \phi)$ and the selection function γ:[53]

(i) If $\phi \in K$, then $K - \phi = \cap\gamma(S(K, \phi))$,
(ii) If $\phi \notin K$, then $K - \phi = K$.

Levi's contraction operator does not have all the properties of Alchourrón, Gärdenfors and Makinson's contraction operator. The latter operator is characterized by the following six postulates:[54]

[52] On the basis of the informational values of K and of the elements of $S(K, \phi)$, Levi also defines the concept 'damped informational value' to rule out the possibility that for two elements K' and K'' of $S(K, \phi)$ with $\mathit{Cont}(K') = \mathit{Cont}(K'')$, it can hold that $\mathit{Cont}(K' \cap K'') < \mathit{Cont}(K')$. We leave this extra complication aside, because it does not affect the logical characteristics of Levi's contraction operator. See [Levi, 1991, § 4.4]. [Levi, 1998] also presents a variant of this adapted informativity concept, which leads to a somewhat more strict contraction operator 'mild contraction', characterized in [Rott and Pagnucco, 1999]. A discussion of Levi's recent refinements of his ideas on contraction would lead to far afield in a study of an introductory nature.

[53] [Levi, 1991, p. 130.] Hansson and Olsson showed that clause (ii), which is lacking in Levi's original definition, is indispensable [Hansson and Olsson, 1995, p. 108]. In [Levi, 1996, p. 23], this minor flaw is corrected.

[54] See [Alchourrón, Gärdenfors, and Makinson, 1985, p. 513].

$(K-1)$	$K-\phi$ is deductively closed,	*Closure*
$(K-2)$	$K-\phi \subseteq K$,	*Inclusion*
$(K-3)$	If $\phi \notin K$, then $K-\phi = K$,	*Vacuity*
$(K-4)$	If $\phi \notin Cn(\emptyset)$, then $\phi \notin K-\phi$,	*Success*
$(K-5)$	$K \subseteq (K-\phi)+\phi$,	*Recovery*
$(K-6)$	If $Cn(\phi) = Cn(\psi)$, then $K-\phi = K-\psi$.	*Extensionality*

Levi's contraction operator meets all these postulates except $(K-5)$, which is known in the literature as *Recovery*.[55] This postulate has come under critical fire, not in the least from Levi himself:

> Consider, for example, a situation where it is believed that Jones was HIV positive, received a drug treatment and subsequently showed HIV negative. Contract the corpus by giving up "Jones received the drug treatment." The conviction that Jones initially showed HIV positive would be retained. But the judgment that Jones showed HIV negative later on would be abandoned. Moreover, restoring the judgment that Jones received the drug treatment would not resurrect the conviction that Jones subsequently showed HIV negative unless the inquirer had the well entrenched conviction initially that the drug treatment always eliminates the HIV virus. If this belief were not well entrenched or if all that is believed is that the drug treatment is followed by cure in some percentage of cases less than 100%, the Recovery Condition would be violated. [Levi, 1998, p. 9][56]

Levi's contraction operator, on the other hand, is characterized by the following five postulates:[57]

$(K-1)$	$K-\phi$ is deductively closed,	*Closure*
$(K-2)$	$K-\phi \subseteq K$,	*Inclusion*
$(K-3a)$	If $\phi \notin K$ or $\phi \in Cn(\emptyset)$, then $K-\phi = K$,	*Vacuity*
$(K-4)$	If $\phi \notin Cn(\emptyset)$, then $\phi \notin K-\phi$,	*Success*
$(K-5)$	If $Cn(\phi) = Cn(\psi)$, then $K-\phi = K-\psi$.	*Extensionality*

3.1.4 Two Applications

A short outline of Levi's treatment of modal and conditional statements concludes this brief exposé of Levi's epistemological and logical positions. As we have seen,

[55] Let L be the language that consists of all truth functional combinations of p and q, and let $K = Cn(\{p,q\})$. Then $Cn(\{q \rightarrow p\}) \in S(K,p)$. Choose γ such that $\gamma(S(K,p)) = \{Cn(\{q \rightarrow p\})\}$. Then $K-p = Cn(\{q \rightarrow p\})$ and $(K-p)+p = Cn(\{p\})$. In this case, K is not a subset of $(K-p)+p$, so *Recovery* does not hold. See [Hansson and Olsson, 1995, p. 112].

[56] See also [Levi, 1991, p. 134–135] and [Hansson, 1999a]. [Levi, 1998, p. 37] presents a second counterexample to *Recovery*.

[57] See [Hansson and Olsson, 1995, p. 109] and [Rott and Pagnucco, 1999, p. 512].

Levi imposes two requirements on a plausible theory for revising epistemic states: (1) such a theory should articulate inference rules which, unlike classical inference, allow us to draw those conclusions from an epistemic state which we would also draw in practical deliberation and scientific inquiry; and (2), it should *formally* justify transitions – expansions as well as contractions – from one epistemic state to another epistemic state. Levi's recent *For the Sake of the Argument* (1996) develops one general, quasi-formal framework with which formal theories that aim at meeting these two requirements can be compared on the basis of abstract characteristics.

All types of sentences that have been in the spotlight of philosophical logic for decades, can be handled within the system developed by Levi. Tautologies, mathematical truths, indicative statements, laws, dispositional statements, modal statements, conditionals, counterfactuals, inductive statements, defaults – they all have their place in Levi's framework. In this subsection we will briefly discuss the way in which Levi accounts for modal and conditional statements within his system. In Levi's view, these types of statements have *no* truth values, unlike the elements of a corpus, though they can be accepted or rejected on the basis of a corpus.

Modal Statements

Levi is of the opinion that modal claims express a serious possibility. A sentence is only held to be seriously possible with respect to an epistemic state represented by a corpus K. However, it does not follow from this that if ϕ is consistent with K, the sentence "ϕ is a serious possibility" is an *element* of K. Levi shows that such a 'realistic' interpretation of modal statements is contrary to his ideas on expansion and contraction.[58] Modal statements cannot be part of a corpus. As an alternative Levi proposes to interpret them as claims *about* a corpus, as claims expressing a certain *property* of such a corpus. Levi sees it as an important advantage of his interpretation of modal statements that it can do without the Kripke semantics which abounds in philosophical logic:[59]

$$\text{“}\phi\text{ is seriously possible” } \textit{is acceptable with respect to } K$$
$$\Longleftrightarrow$$
$$K \cup \{\phi\} \textit{ is consistent.}$$

[58] See [Levi, 1991, § 3.7] for a concise presentation of these arguments.

[59] Levi opposes modal extensions of the language in which our opinions are articulated. In Levi's view, modal logic and its applications to epistemological and metaphysical problems are a "retrograde step in philosophy" [Levi, 1980, p. xvi]. Possible worlds semantics is, strictly speaking, superfluous: "[M]any advocates of the usefulness of possible worlds semantics for the purpose of explicating judgments of possibility and conditionals appeal to examples that may be given a straightforward epistemic or, in the case of conditionals, belief-change treatment." [Levi, 1991, p. 114.] Indeed, "[c]onditionals understood in terms of imaging become mere artifacts of the metaphysician's fevered imagination." [Levi, 1996, p. 76.]

Conditionals

According to Levi, conditionals are a second class of statements which never are an element of a corpus. They do, however, express properties of a corpus, properties which tell us how a corpus will behave under certain revisions. Levi interprets conditionals on the basis of *Ramsey's Test*, which takes its name from a cursory remark by Frank Ramsey on the interpretation of conditionals:

> If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, $\neg q$' are contradictories. [Ramsey, 1929, p. 143n]

Levi now proposes to interpret conditionals on the basis of Ramsey's Test using the techniques for *belief change* outlined above, a choice which obliges him, under penalty of triviality, to refuse to admit conditionals as elements of corpora.[60] Therefore, like modal statements, conditionals do not have truth, but only acceptability conditions. Then Levi proposes a theory figuring three different variations of Ramsey's Test.[61] According to Levi, this theory provides an adequate solution to the well-known problems concerning conditional sentences: "[A]ll 'if' sentences customarily classified by contemporary philosophers as indicatives and as subjunctives are explicated by various versions of the Ramsey test" [Levi, 1996, p. 13].

In order to find out whether a conditional "If ϕ, then ψ" is acceptable with respect to a consistent corpus K, we add the antecedent ϕ to our current 'stock of knowledge', that is, to the actual corpus K. If this addition produces an inconsistent set of sentences, we make minimal changes in the inconsistent set so as to make it consistent while retaining the sentence ϕ. This procedure, carried out along Levi's criteria for revisions of corpora, produces a new consistent corpus K' which contains ϕ and differs only minimally from the old corpus K. Then we check whether the consequent ψ is an element of this new corpus K'. If that indeed is the case, then the conditional "If ϕ, then ψ" is acceptable with respect to the original corpus K.[62]

All three variations of Ramsey's Test discussed by Levi can be formulated in terms of contraction and expansion. Levi argues that a conditional "If ϕ, then ψ" should ideally be interpreted on the basis of a corpus that contains neither ϕ nor $\neg\phi$. Hence, the corpus K must be processed such that both ϕ and $\neg\phi$ are serious possibilities with respect to the adapted corpus. In some cases, this may

[60] See [Rott, 1989] for an elegant proof of Gärdenfors's Triviality Theorem from 1986.

[61] See [Levi, 1996, p. 18–50] and [Gärdenfors, 1988, p. 147–148].

[62] Similar epistemic interpretations of conditionals have been formalized previously in [Rescher, 1964] and [Veltman, 1976]. From the beginning, the idea that counterfactuals need an epistemic interpretation has been criticized. See [Kratzer, 1981], [Lewis 1973], and [Stalnaker, 1968]. For a more recent discussion, see [Rott, 1999] and [Stalnaker, 1992].

mean that ϕ has to be eliminated by contraction from K, in others, (think of counterfactuals), $\neg\phi$ has to be eliminated. This variant – Levi's favourite – can be defined as follows [Levi, 1996, p. 31]:

$$\text{“If } \phi\text{, then } \psi\text{” } \textit{is acceptable with respect to } K$$
$$\iff$$
$$\psi \in ((K - \neg\phi) - \phi) + \phi.$$

3.2 A Critique

3.2.1 Corpora?

A corpus represents an epistemic state. As the epistemic states of an individual agent are highly continuous, even when the agent is placed in a multitude of situations, we cannot assume that epistemic states depend only on context. Let us assume that we wish to articulate our actual epistemic state by a corpus in a certain formal language L. Levi assumes corpora to be consistent and deductively closed. Hence, if we wish to identify the corpus that represents our actual epistemic state, we only need to determine which sentences in L are elements of the corpus and which are not. Is there a criterion for this? Are the contextual certainties we apply in the current context C_1 – epistemic states are always located within some context – perhaps precisely the elements sought after? (The contextual certainties in C_1 are exactly all sentences ϕ with $Q_1(\phi) = 1$, where Q_1 is our expectation-determining probability measure in context C_1.) The certainties we use in context C_1 may, however, very well be incompatible with our certainties in a context C_2.[63] Since the corpus we are looking for must be consistent, we cannot immediately admit contextual certainties into the corpus. This is due to the circumstance that although an element of a corpus is in fact necessarily a contextual certainty, the converse does not apply: imagine a situation where an agent X with corpus K in context C_1 accepts that $Q_1(\phi) = 1$, while the same agent with the same corpus in context C_2 accepts that $Q_2(\neg\phi) = 1$. According to Levi, it cannot be concluded from the fact that $Q_1(\phi) = 1$ that ϕ is an element of K.[64] Of course this holds for $\neg\phi$ as well. Consequently, X is free to apply mutually inconsistent certainties within different contexts. Contextual certainties, therefore, give us little to go on in determining our current corpus K.

But how do we determine whether a certain sentence is an element of the corpus K which represents our current epistemic state? If we can imagine a context C in which, given the current corpus K, a sentence ϕ is relevant, while $Q(\phi) \neq 1$, then ϕ is *not* part of our corpus, because in this case $\neg\phi$ is a serious possibility with respect to K. Might that be a feasible criterion?

[63] Cf. [Batens, 1992, p. 202].

[64] Levi writes: “According to X's credal state, all items in his corpus receive probability 1 (although the converse need not hold).” [Levi, 1976, p. 10.]

We have already seen that the certainties in our current context C_1 do not give a definitive answer, but at best a clue as to whether those certainties are an element of our corpus. The abovementioned criterion instructs us to verify whether $Q_i(\phi) = 1$ holds in *all* imaginable contexts C_i when our corpus is the corpus K. A tricky question presents itself, even supposing that it is feasible to scour all imaginable contexts: how do we know if our epistemic state, represented by corpus K, stays the same with a context shift? It is impossible to aim, on the one hand, at finding the elements of K, and to make sure, on the other, that the corpus applied in context C_i is identical to K, because after all we have to know what K's elements are before we can make that comparison. How do we know for sure whether in such a context shift we have not inadvertently expanded or contracted K?

Although it is a direct consequence of Levi's epistemology, the criterion formulated in this paragraph for the determination of the elements of a corpus K will have to do without any application. For want of a better criterion with which to decide whether a certain sentence is an element of a corpus we can only come to the provisional conclusion that Levi's notion of a corpus violates Quine's maxim: "No entity without identity" [Quine, 1969, p. 23].

3.2.2 Contextual Parameters?

The preceding discussion of Levi's theories of expansion and contraction show that Levi's epistemology can only be applied if we have at our disposal (estimations of) numerical values for the 'system of contextual parameters' he uses, consisting of, among others, the 'ultimate partition' U, the information-determining measure M, the expectation-determining measure Q and the degree of boldness q.[65] This system of contextual parameters is part of X's epistemic state at time t. In his inquiry into the mechanisms of belief change, Levi just presupposes that we have found sufficiently specific values for these parameters and argues for his criteria concerning legitimate expansions and contractions on the basis of this presupposition. Levi defends his crucial presupposition with an *ad consequentiam* argumentation:

> Of course, investigation may reveal that no system of contextual parameters can be identified such that, given specific values for these parameters, the legitimacy of X's modification of bodies of knowledge would be subjective or context dependent in a sense which put them beyond critical control. However, we would be obstructing the course of inquiry to assume that this is so at the outset. [Levi, 1976, p. 2]

[65] After a brief discussion of deliberate expansion, Levi writes: "This sketch presupposes that there is no indeterminacy in evaluations of informational value and that a definite value for the index q has been fixed." [Levi, 1991, p. 93.]

Following this argument, the first question we could ask ourselves is *which* course of inquiry is obstructed if we do not assume that sufficiently specific values can be found for Levi's system of contextual parameters? It would seem a bit far-fetched that Levi is alluding here to research in the field of theoretical physics or in the area of comparative literature. These (and other) disciplines do not need Levi's methodological considerations to get along anyway. More likely, Levi is referring to epistemological investigations which aim at grasping 'the logic of inquiry', investigations that also include Levi's own epistemology.

Second, we can check, on the basis of clues from Levi's own works, if and to what extent the presupposition that Levi's contextual parameters have been specified accurately enough is plausible. We will limit ourselves to the assessments of the probability measure M and the degree of boldness q.

Informational Value

Assessing an information-determining M-function is a context dependent matter. Among others, it depends on the cognitive aims pursued by an agent X in a given context. As a consequence, there are hardly any gains to be expected from the search for a universal, context independent M-function.[66] Since Levi's criteria for both expansion and contraction can only be applied once we have, among others, assessed the M-function, this function should be assessed before the intended evaluation of the proposed expansion or contraction can take place. Such assessments

> are part of the abductive task. To some extent, these assessments may be regulated by criteria which are applicable to a large class of problems. It may, perhaps, be possible to identify certain desiderata which determine explanatory power and simplicity relevant to the assessment of informational value in inquiries where the aim is to obtain explanations of some kind. It is doubtful, however, that such desiderata can be converted into criteria for the evaluation of informational value which render it irrelevant to consider the peculiarities of the particular demands for information motivating specific inquiries. Indeed, such restrictions on the assessment of informational value are likely to be very weak. Such assessment is, in my opinion, heavily context dependent. [Levi, 1980, p. 47]

In the end, in the adoption of a certain (class of) M-function(s), an 'abductive logic' is the deciding factor:

[66]Levi writes: "The considerations that enter into an evaluation of informational value are diverse, often competing, and heavily context dependent. Different kinds of inquiries impose different demands for new information, so that it is not to be expected that evaluations of informational value will meet the same requirements in all contexts. And inquiries addressing the same issues may be committed to different research programs generating different demands for information." [Levi, 1991, p. 83.]

> Arguments concerning the adoption of one M-function rather than another are to be evaluated (insofar as there is a right and a wrong to the matter) according to principles of abductive logic [Levi, 1976, p. 40–41].

We can, however, hardly see this reference to an 'abductive logic' as anything else but a shortcut, as long as Levi keeps us guessing about the peculiarities of such an abductive logic. Hence, the appeal to an 'abductive logic' does not contribute anything to the assessment of an M-function. Luckily, Levi gives us some rather more tangible clues for the determination of an M-function, even though the assumption of an M-function is considered by him to be "excessively unrealistic":[67] different kinds of values, for example precision, simplicity and explanatory power, "constitute different dimensions that contribute to the assessment of what I call *informational value*."[68] Moreover, Levi writes:

> [T]he demands for informational value that animate the inquirer's deliberations [...] may reflect commitments to research programs and ideals of explanatory adequacy, simplicity, systematicity, precision, and the like, including commitments to certain types of theoretical frameworks [Levi, 1991, p. 150–151].

This does not help us make any headway either: the original problem of finding a numerical specification for one parameter is now 'reduced' to a messy multitude of problems. Is Levi's list complete? And how do we assess the different values on the list? How can the relative importance of these values be assessed? Many questions, but no answers.

Are we perhaps just splitting hairs by demanding excessive preciseness in the estimation of Levi's M-values? Isn't it enough if, off the top of our head, we make an ordering of the informational values of those sentences or corpora that require an M-value in a given context, so that the correct expansion or contraction can then be assessed?[69] Still, in order to apply Levi's contraction operator, a

[67][Levi, 1976, p. 37.]. Nevertheless, Levi thinks that "some light can be shed on the more realistic situations of daily life and scientific inquiry by considering these idealisations." [Levi, 1976, p. 37.]

[68][Levi, 1991, p. 145] See also [Levi, 1998, p. 16].

[69]Levi writes: "[A]ny two evaluations that yield the same weak ordering of the potential contraction strategies will yield the same recommendations. One might argue that quantitative differences in the assessment of informational value [...] are irrelevant to the assessment of contraction strategies." [Levi, 1998, p. 50.] Therefore, "in contraction quantitative dimensions of informational value do not seem relevant" [Levi, 1998, p. 51], since "only ordinal considerations are needed in contraction" [Levi, 1998, p. 52]. In expansion, however, such quantitative considerations do play a prominent role, because after all, an evaluation of a potential expansion strategy amounts to the trading off of the informational value and credal probability of the available options weighed by a degree of boldness: "The task of aggregating informational value and risk of error in a single assessment [...] calls for the quantitative characterization of informational value." [Levi, 1998, p. 21.]

certain precision is needed: if we want to apply Levi's selection function γ without difficulties, we should be able to establish whether there is a negligible or in fact a considerable difference between the informativities $Cont(K')$ and $Cont(K'')$ of two saturatable contractions K' and K'' in $S(K, \phi)$. Numerical identity is, after all, too much to ask, since Levi's clues for finding the right numerical values for his M-function cannot warrant the preciseness required in the case of numerical identity. It would be in Levi's line of argument to sidestep the problem discussed here by conjuring up yet another parameter, ϵ, which represents the doubtlessly context dependent degree of precision of an agent X, so that by referring to ϵ, it is possible to assess whether the informativities $Cont(K')$ and $Cont(K'')$ are negligibly different from each other or not:

$$Cont(K') \sim Cont(K'') \quad \Longleftrightarrow \quad |Cont(K') - Cont(K'')| \leq \epsilon.$$

It is up to the reader to determine whether this strategy gives a satisfactory solution for our problem.

The assessment of an information-determining M-function takes more doing than Levi wishes us to believe by simply assuming that it has already been assessed. In spite of all Levi's clues it is altogether implausible that the M-function could ever be determined with sufficient preciseness in a given context. Since Levi, given a corpus K and a sentence ϕ, eventually reduces the problem of finding the admissible contraction $K - \phi$ to the problem of determining an M-function over the saturatable contractions in $S(K, \phi)$, we can at least come to the conclusion that something is lacking in Levi's epistemology, considering the difficulties in assessing the M-function. Unfortunately it is not entirely clear to me how this lacuna in Levi's epistemology could be filled appropriately.

A second problem with the assessment of informational values arises when we wish to test the acceptability of *counterfactuals* via Levi's criteria for conditionals. At first glance, Levi's approach seems to be preferable to David Lewis's and Robert Stalnaker's treatment of counterfactuals with possible worlds semantics, in which a relation of similarity between worlds plays a crucial role. After all, it is possible, within Levi's system, to sidestep counterexamples to David Lewis's and Robert Stalnaker's analysis. In 1976, Pavel Tichý described a situation in which Lewis's and Stalnaker's analysis of counterfactuals produces a result that is completely at odds with our pretheoretical intuitions:

> [C]onsider a man – call him Jones – who is possessed of the following dispositions as regards wearing his hat. Bad weather invariably induces him to wear his hat. Fine weather, on the other hand, affects him neither way: on fine days he puts his hat on or leaves it on the peg, completely at random. Suppose, moreover, that actually the weather is bad, so Jones *is* wearing his hat. [Tichý, 1976, p. 271]

The statement “If the weather were fine, Jones would be wearing his hat”, which, on the strength of our pretheoretical intuitions, is unacceptable, would be acceptable according to Lewis’s and Stalnaker’s analysis.

If we write K for the, obviously consistent, corpus that contains at least the abovementioned information, ϕ for “The weather is fine” and ψ for “Jones is wearing his hat”, then it is clear that both $\neg\phi \in K$ and $\psi \in K$. To check with Levi’s method whether the conditional in question is acceptable with respect to K, we first have to contract the corpus K by $\neg\phi$. (A contraction of the resulting corpus with ϕ will not be necessary, as $\phi \notin K$.) At this point it is important to arrive at a corpus that leaves open whether ψ is the case or not, because if ψ were to remain in $K - \neg\phi$, the statement “If ϕ, then ψ” would be acceptable with respect to K.

At first sight, it seems that the corpus $K - \neg\phi$ must contain ψ. As ψ is obviously relevant to the problem at hand, it seems likely that each corpus K^* such that $K^* \subseteq K$ and $\psi \notin K^*$ must have a lower informational value than $K - \neg\phi$. It seems an unavoidable conclusion that $\psi \in K - \neg\phi$, because Levi demands of an admissible contraction $K - \neg\phi$ minimal loss of informational value. There is, however, an emergency exit: according to Levi’s *Weak Monotonicity* postulate, it holds that $K - \neg\phi$ has at least as much informational value as all its subsets, but, on the other hand, real subsets K^* of $K - \neg\phi$ can have the same informational value as $K - \neg\phi$ itself. Hence, I propose to leave these sceptical considerations aside and to try to find, in line with Levi’s proposals, an admissible corpus $K - \neg\phi$ such that both ψ and $\neg\psi$ are serious possibilities with respect to that corpus.

This we can do by finding out which element of $C(K, \neg\phi)$ is most suited to serve as a starting-point for a further expansion with ϕ. This element must be a deductively closed subset of K which contains neither $\neg\phi$ nor ψ nor $\neg\psi$ nor $\phi \rightarrow \psi$ nor $\phi \rightarrow \neg\psi$. Let K^* be this subset. Then there is a subset S^* of $S(K, \neg\phi)$, such that $K^* = \cap S^*$. If we now choose the informational values of the elements of $S(K, \neg\phi)$ in such a way that exactly all elements in S^* have the highest informational value, then it holds that $\gamma(S(K, \neg\phi) = S^*$ and, hence, that $K - \neg\phi = K^*$. It is now plain that, in accordance with our pretheoretical intuitions, the following sentences are both unacceptable with respect to K:

(1) “If the weather were fine, Jones would be wearing his hat.”
(2) “If the weather were fine, Jones wouldn’t be wearing his hat.”

However, something still does not feel right. It turns out that Tichý’s example is, indeed, not a counterexample to Levi’s account of conditionals as long as we are absolutely free to choose the informational values of the elements of $S(K, \neg\phi)$ in such a way that we reach the result that was prescribed by our pretheoretical intuitions. Therefore, the desired result holds sway over the assessment of the informational values of the elements of $S(K, \neg\phi)$. It is even possible to generalize this observation to a theorem: Let $\neg\phi \in K$, such that $\neg\phi \notin Cn(\emptyset)$ and K is consistent. Then, for each ψ in L such that $\neg\phi \notin Cn(\neg\phi \vee \psi)$ there is a choice of informational values over $S(K, \neg\phi)$ such that "If ϕ, then ψ" is acceptable with respect to K.[70] In short, we can always construct the required contractions in such a way that each and every counterfactual turns out to be acceptable. What would be the explanatory power of a theory of conditionals which can validate every counterfactual? 100%? (If Oswald had not killed Kennedy, Jack Ruby would have been the first man on the moon.)

Degree of Boldness

In Rule A, which gives the conditions under which a deliberate expansion is legitimate, the degree of boldness q weighs the informational value $M(\phi_i)$ of rejecting an option ϕ_i against ϕ_i's credal probability $Q(\phi_i)$. Hence, the higher q, the more options in U escape rejection, the bolder X is in accepting new information. With the introduction of the degree of boldness, which Levi also calls the 'index of caution', it was stipulated that $0 < q \leq 1$. Obviously, we cannot apply Rule A unless we have determined the parameter q, even if we have already found the appropriate numerical values for the probability measures M and Q. Levi pays a relatively large amount of attention to a method with which q can be determined with greater accuracy. This method for finding the 'true value' [Levi, 1976, p. 46] of q (or an appropriate interval for values of q) is based on Rule A and on a process which Levi calls 'bookkeeping'. He describes this bookkeeping process as follows:

> Suppose that X has a corpus K_1 and a question under consideration. Rule A (or one of its variations) recommends a given conclusion at an appropriately high level of caution and X's problem situation does not contain any other questions for which rule A recommends conflicting conclusions. X is in a position to accept as evidence the conclusion which K_1 and rule A warrant his merely accepting. In that case, X shifts to the corpus K_2 which is the result of such expansion. He can then proceed to consider whether relative to K_2 rule A licenses adding still more information pertinent to the question X is considering. X can ask the same question over and over again until he has exhausted the resources of using rule A in this reiterated way. [Levi, 1976, p. 42–43]

[70]Suppose that $\neg\phi \in K$ and that K is consistent. Since K is deductively closed, it holds that $\neg\phi \vee \psi \in K$ for all ψ in L. Suppose that $\neg\phi \notin Cn(\neg\phi \vee \psi)$. Then there is a deductively closed subset K^* of K such that $\neg\phi \notin K^*$ and $\neg\phi \vee \psi \in K^*$. There is a subset S^* of $S(K, \neg\phi)$, such that $K^* = \cap S^*$. Choose the informational values of the elements of $S(K, \neg\phi)$ in such a way that exacly all elements in S^* have the highest informational value. Then, by definition, $K - \neg\phi = K^*$. As K is consistent, it holds that $\phi \notin K$. By *Inclusion*, it holds that $\phi \notin K - \neg\phi$. By *Vacuity*, it holds that $(K - \neg\phi) - \phi = K - \neg\phi = K^*$. Hence, $\neg\phi \vee \psi \in (K - \neg\phi) - \phi$. Hence, $\psi \in ((K - \neg\phi) - \phi) + \phi$. Therefore, "If ϕ, then ψ" is acceptable with respect to K.

Levi throws light on his positive evaluation of "the usefulness of the bookkeeping device in imposing constraints on degrees of caution" [Levi, 1976, p. 43] by means of the following considerations. Let U be a finite ultimate partition, and let $q = 1$.[71] If we write $Z(\phi)$ for the quotient $Q(\phi)/M(\phi)$, it can be proven that within a finite number of reiterations of the bookkeeping process all elements from U that do not have maximum Z-value are rejected. As a rule, a lesser degree of boldness brings it about that a number of elements from U that do not have maximum Z-value escape rejection [Levi, 1976, p. 44]. One example then shows that in some cases it *must* be that $q < 1$: given that $q = 1$, if we apply Rule A and the bookkeeping process to determine which options from U have to be rejected, where all options from U are understood to be equally informative, then precisely all the options with the greatest plausibility escape rejection. In this case, a situation can be imagined which clearly contradicts our pretheoretical intuitions:

> If I toss a coin known to be unbiased ten times, I am not prepared to predict that it will land heads exactly 5 times. If I toss it 100 times, I am not prepared to predict that it will land heads exactly 50 times. My reluctance is not based on any doubts as to the truth of the claim that the chance of heads is 0.5. I would not make the predictions even if I was certain that 0.5 is, indeed the chance of heads. The point is that I would refuse to rule out hypotheses asserting that the relative frequency of heads will differ from heads by some small amount. [Levi, 1976, p. 44]

On the other hand, in other cases the desired result can only be reached if $q = 1$. Here are two of Levi's examples:

> In estimating the value p of the chance of obtaining heads on a toss of a coin, imagine that X observes a large number of tosses and they all land heads. [...] It will become clear that under suitably specified conditions, X will be able to conclude that p falls in an interval from $1 - \epsilon$ to 1 when the value of q is less than 1 but will not be able to conclude that p equals 1. Yet, on some occasions, it may be appropriate to conclude that the coin will always land heads on tosses and $p = 1$.
>
> [...]

[71] In the formulation of the conditions for the theorem discussed here, the degree of boldness q is fixed for the entire bookkeeping process. This, however, is not mandatory. According to Levi, the degree of boldness may also be changed during the process: "[O]ccasions can arise where an investigator has reached a stage of inquiry where his demands for information may very well induce him to modify his ultimate partition and change his degree of caution." [Levi, 1976, p. 48.] During the bookkeeping process the probability measures M and Q are adjusted via conditionalization. See [Levi, 1980, p. 54n].

> Suppose that a population of organisms which are all hybrid with respect to some pair of alleles are randomly mated and data as to the percentage of offspring which are purebred dominant, purebred recessive and hybrid collected. On the basis of the data, one might conclude that the chance of obtaining a purebred dominant offspring from hybrid parents is approx[i]mately $\frac{1}{4}$, of obtaining purebred recessive offspring is also approximately $\frac{1}{4}$ and of obtaining hybrid offspring is approximately $\frac{1}{2}$. However, there is considerable pressure to conclude that the values for these chances are exactly $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{2}$ respectively. [Levi, 1976, p. 45–46]

Finally, after discussing the views of critics and rivals on the assessment of a numerical value for q, Levi comes to a modest conclusion:

> I suggest that the q-value ought to be substantially less than 1 unless there is some excusing circumstance along the lines described above. I have no firm conviction as to what an appropriate numerical value for q ought to be which will satisfy "all but the virtually sceptical" but I suspect that many sceptics will remain dissatisfied unless it is substantially less than 0.5. [Levi, 1976, p. 51–52][72]

Although Levi does not state it in so many words, he cannot do without reference to a "presystematic judgment"[73] on which expansions are justified in which cases, if his suggestion to make a more accurate approximation of the value of q by way of the bookkeeping process is to wash. After all, as long as we do not assume a presystematic judgment on which expansions are legitimate at what time, we have, applying Levi's bookkeeping technique for an approximation of q, no reason whatsoever to prefer one specific value (or interval of values) for q over another. In short, the desired result plays a decisive role in the choice of a specific value for the degree of boldness q.

3.3 Conclusion

We noted that even the assessment of the elements of a corpus K results in problems which are as yet unsolved. Moreover, even if we know which elements are part of a corpus K, the values of Levi's M-function, Q-function and parameter q can always be chosen in such a way that each expansion or contraction can be legitimized within Levi's epistemological framework with hindsight, if that expansion or contraction meets at least a number of minimal constraints. No

[72]Four years later, Levi writes: "[N]ormally q should be less than 1 (and indeed, less than .5)" [Levi, 1980, p. 55].

[73]See also Levi's refutation of a proposal by Keith Lehrer [Levi, 1976, p. 44].

credence can be attached to the *normative* status of his theory, to which Levi sticks through thick and thin,[74] as long as Levi fails to formulate convincing criteria – criteria that are *independent* of the desired result – with which (1) the identity of a corpus K can be found out, and (2) the system of contextual parameters (the parameter q and the probability measures M and Q) can be assessed adequately.[75] It is a misleading strategy to nonchalantly *assume* that "appropriate inputs are present (such as the demand for information, ultimate partition, degree of caution, credal state, and so on)" [Levi, 1991, p. 107] and then to formulate criteria for legitimate expansions and contractions on the basis of these values, since not much is gained with a 'solution' in which a given problem is reduced to the values of a number of parameters of which it is absolutely unclear how they should be assessed. Would we really be explaining much less if we limited ourselves to one parameter, l, which stipulates the legitimacy of belief changes: a 'measure of legitimacy'?[76]

May Levi's dynamic epistemology have a second life, this time in the form of a descriptive theory? Descriptivity implies the obligation of testability, which Levi's theory, at least until now, fails to meet: as long as the numerical values of the parameters of our epistemic state before and after a controlled change cannot be *measured* with sufficient accuracy, a test (as opposed to the usual reference in belief change literature to the 'intuitiveness' of basic assumptions and postulates[77]) of Levi's theory, supposing it is testable, remains a futuristic idea. Though Levi antagonizes Lewis and Stalnaker by proclaiming that their treatment of conditionals and conditional logic, based on possible worlds semantics, are

[74] "I am concerned with conditions under which changes in doxastic commitments are legitimate. The concern is prescriptive, not explanatory" [Levi, 1991, p. 107] and "[m]y preoccupation is in the final analysis with identifying standards of rational health in reasoning. Logicians, mathematicians, and computer scientists make an important contribution to identifying what those standards might be and to the design of technologies that can contribute to enhancing our limited abilities to realize these standards. I am interested in defending a view not of what these standards might be but what they should be." [Levi, 1996, p. xiii.]

[75] Levi tries to justify his refusal to have anything to do with inquiries into criteria for finding adequate values for his parameters with an argument by analogy: "Thermodynamics and some branches of economic theory illustrate comparative statical theories which investigate changes in equilibrium states of systems suitably specified without scrutinizing the details of the paths such systems follow in moving from one equilibrium state to another. The normative analogue of such theories of the sort I am aiming to construct here prescribes shifts from one state of cognitive equilibrium to another without prescribing details of the psychological or social changes which are made in implementing the revision." [Levi, 1980, p. 11.] This does not wash, because the assessing of the values of the required parameters in 'thermodynamics and some branches of economic theory' is done relatively unproblematically, while it is a major obstacle for Levi's logico-epistemological system. Accordingly, thermodynamic theories are usually testable, while theories in the field of *belief change* are not.

[76] Note that this mock proposal does not differ in principle, but only *in degree* from Levi's approach.

[77] Levi writes: "[T]he appeal to postulates that seem intuitively compelling at first blush is a risky business." [Levi, 1998, p. 9.]

"formalisms in search for an as yet undiscovered application" [Levi, 1996, p. 82], they should actually be seeing eye to eye with one another

While Levi's proposals for expansion, once it has been decided to expand a corpus K with a sentence ϕ, are unproblematic from a logical point of view, the forementioned difficulties of determining an appropriate information-determining M-function prevent us from applying Levi's contraction operator, even if the decision has already been taken to contract a corpus K with a sentence ϕ. Levi's proposal to concretize the ubiquitous reference to a selection function γ in belief change literature with an information-determining probability measure M does not yield much, since Levi's contraction operator cannot be applied without the required informational values. This problem cannot be avoided by simply assigning the same informational value to all the elements from a not-empty $S(K, \phi)$, in which case $\gamma(S(K, \phi)) = S(K, \phi)$, because then, as Hansson and Olsson demonstrated, it holds that $K - \phi = Cn(\emptyset)$. The standard approach of [Alchourrón, Gärdenfors, and Makinson, 1985], in which the selection function γ is not explained in any further detail, suffers from the same deficiency.

Therefore, it is important to strive for an approach of *belief change* in which the usual reference to an extra-logical element, such as a selection function or an ordering of sentences in a corpus on the basis of their corrigibility, is avoided, so that a contraction, as well as an expansion, can be implemented directly, that is, *without* appealing to an extra-logical element. This is even more important in cases where we have no indication of the relative corrigibility of the elements of the corpus to be contracted. Only when we have construced such theories, can we verify on the basis of empirical tests whether the proposed contraction operator holds good. Within a classical logical framework this wish (until now) can only be fulfilled under penalty of totally unacceptable results. An underlying logic which is weaker than classical logic might open up new perspectives.

Chapter 4

Finite State Belief Dynamics

Suppose that we get the information that “Cervantes lost his left hand in battle” (p), “Cervantes did not lose both hands in battle” ($\neg(p \wedge q)$), and “Góngora was born in 1561” (r). Suppose, furthermore, that all information is on equal footing, that is, the sources from which the information stems are equally reliable. Which beliefs can we plausibly extract from these sentences? Obviously, the information offered is consistent, so, if we have to rely solely on the given sentences, there is no objection to believing all of them (and their logical consequences) at the same time.

But suppose, furthermore, that we get the additional information that “Cervantes lost his right hand in battle” (q). First, now all sentences cannot be true at the same time, although it is possible, of course, to *get* these pieces of information. Assuming our beliefs to be based on the information we have, it makes sense to say that in the information base consisting of our four sentences it is plausible to believe that Góngora was born in 1561 and that Cervantes lost *a* hand in battle, although we do not have plausible beliefs as to whether it was his right or his left one (or even both!). New information can, therefore, put a strain on old beliefs: this course of reasoning is nonmonotonic.

Suppose now, that, for some reason or other, the information that “Cervantes lost his left hand in battle” turns out to be unreliable. We eliminate the information that “Cervantes lost his left hand in battle” from the information we have, which is thus *contracted*. This means that, *ceteris paribus*, we are entitled to believe that “Cervantes lost his right hand in battle”, although our previous set of sentences did not enable us to do so.

If we apply classical logic to the inconsistent information base just described, we fail to save these appearances, because of classical logic’s property of *ex falso quodlibet*. Obtaining all formulas from a contradiction is inappropriate for the information base described above, for then we would lose all structure and information we had: if we have everything, we have nothing at all. Therefore, if we want to construct a formal theory that describes the phenomena sketched above, classical logic will not do. We do need a logic which can handle inconsistencies, so some relevance logic or paraconsistent logic seems convenient. This logic has to have another important characteristic as well: it must be able to cope with incomplete information, as our information base will never be complete. Here, we shall deal with the matter using a *partial*, *paraconsistent* logic.

4.1 Paraconsistent Belief Revision?

The literature on paraconsistent logics is focused on taming the inconsistent. A range of systems has been built in which a contradiction does not necessarily lead to triviality. Dynamics usually does not play a role in these systems. On the other hand, customarily, the consistency of theories is assumed in the belief revision tradition initiated by [Alchourrón, Gärdenfors, and Makinson, 1985].[1] Little work has been done to integrate paraconsistent logics with qualitative logics for theory change. To the best of my knowledge, only [Restall and Slaney, 1995] presents belief revision using a paraconsistent underlying logic. Let us take a closer look at Restall and Slaney's reasons for such an approach.

Restall and Slaney use arguments taken from the field of relevant logic to motivate their paraconsistent approach to belief revision. In almost all systems for belief revision, Restall and Slaney note, the underlying logic is *superclassical.*

> This is a theoretical simplification. No-one believes that belief is closed under that sort of consequence. If it were, we would believe all tautologies, and furthermore, we would only have inconsistent beliefs when believing everything. [Restall and Slaney, 1995, p. 1]

Though Restall and Slaney reject the view that two contradictory statements can both be true, and hence grant that 'knowledge brooks no contradiction', they accept inconsistent beliefs, but retain the idealization that beliefs be closed under logical consequence. Thus, they make way for a straightforward substitution for the underlying logic in systems for belief revision: instead of using (supra)classical logic they opt for the paraconsistent logic first degree entailment as the underlying logic and show that all standard representation theorems (epistemic entrenchment, transitively relational partial meet contraction, and spheres) are preserved under the proposed substitution. Restall and Slaney admit inconsistent beliefs, but also provide a method to extract consistent subsets from an inconsistent theory.

We share Restall and Slaney's misgivings about using (at least) classical logic as underlying logic in systems for belief revision, and we favor their proposal to use first degree entailment instead. Nevertheless, we refuse to allow for inconsistent beliefs in a logical system, yet we cannot but agree with an aphorism of Charles

[1] There are some attempts to deal with inconsistent belief sets within this AGM tradition. See, for instance, [Fuhrmann, 1991], [Hansson, 1993], and [Wasserman, 2000].

Caleb Colton's: 'Man is an embodied paradox, a bundle of contradictions.' Just as a theorist of correct reasoning does not aim to validate fallacies, regardless of how often people commit them, we favor a theory of belief revision that shuns inconsistent beliefs, even in cases where the information on which these beliefs are grounded is inconsistent.

This asymmetry mirrors the fact that it is in our power to aspire to consistent beliefs (and perhaps even to attain them), though the inconsistency of information offered is beyond our control. The system to be expounded in the present chapter is a *normative* system, since it exemplifies the norm that beliefs ought to be consistent. A similar maxim cannot be maintained for information: it is pointless to precribe that information be consistent. Of course, Restall and Slaney subscribe to the norm that beliefs ought to be consistent, but their paper 'Realistic Belief Revision' has a different object: it aims at a *description* of actual reasoning practices.[2] Hence, they allow for inconsistent beliefs, not only in practice, but in logical theory as well.

4.1.1 A Novel Approach

Paraconsistent logic can, indeed, handle inconsistent sets of formulas satisfactorily. Nevertheless, its commitment to inconsistent beliefs or even contradictory beliefs, notwithstanding philosophical defences of *dialethism*,[3] gives it an exotic flavour, which is a serious hindrance for its introduction in epistemic contexts. Secondly, our introductory considerations showed that we need a nonmonotonic logic to appreciate the fact that additional information that contradicts previous information may lead to giving up some of our beliefs. Paraconsistent logic is far too liberal as regards the admission of beliefs. In my view, the admission of inconsistent *beliefs* is a major flaw in all the more or less paraconsistent treatments of inconsistent *information*.

Here, we argue for a distinction between *information* and *belief*. On the one hand, we shall set forth interrelated techniques for representing, expanding, contracting and revising *information*. Information may, of course, be inconsistent. Henceforth, the devices representing our information can contain contradictory and even inconsistent sentences. On this level, nonmonotonicity does not play a role. Hence, we may use the paraconsistent monotonic logic first degree entailment as the logic governing expansion, contraction, and revision of information. On the other hand, operations are offered to extract *beliefs* from the represented information. These beliefs will always be consistent and are closed under logical consequence. Here, nonmonotonicity seems imperative.

[2]Note that the phenomena discussed in the introduction to this chapter cannot be described within Restall and Slaney's framework, because their underlying logic (first degree entailment) is monotonic. In the present chapter, the underlying logic is a nonmonotonic logic *based* on first degree entailment.

[3]Dialethism is the view that there are true contradictions.

Hansson, in his editorial in the thematic issue on belief revision of the *Journal of Logic, Language, and Information*, makes a case for such a two-fold approach:

> [T]he dynamics of belief states involves two major types of operations. One is *operations of change*, transformations from one belief state to another. [...] The other major type can be called *operations of retrieval*. The task of such an operation is to find, for a given belief state, the set of sentences to which the agent has a certain epistemic attitude. [...]
>
> In the presence of conflicting information, selections are necessary. We have a choice between making these selections as part of the operations of change when new information is received and making them as part of the operations of retrieval when information is recovered from the system [...]. [Hansson, 1998, p. 125]

The aim of the present study is to develop a suitable system for belief dynamics with the following characteristics. First, the system must be capable of representing inconsistent information. Second, the set of beliefs licensed by possibly inconsistent information must *always* be consistent.

Section 4.2 sets forth the method for representing static information. The dynamics of information change are studied in Section 4.3. To model information and information change, a simple inconsistency-tolerant four-valued logic, known as first degree entailment, is used. This logic is a generalization of classical propositional logic in the sense that it admits, next to classical propositional logic's total and consistent valuations, partial and inconsistent valuations as well.

In Section 4.4, operations to extract consistent beliefs from a representation of static information are explored. Several methods for processing sets of four-valued valuations to obtain a set of consistent three-valued valuations are investigated. The resulting three-valued valuations are valuations in the sense of Kleene's strong three-valued logic and, hence, guarantee that the set of formulas they all validate is consistent. The logics governing these extraction operators are nonmonotonic three-valued logics based on the four-valued logic first degree entailment.

4.2 Preliminaries

The basic logic, underlying all concepts and systems in the present chapter, will be first degree entailment (henceforth fde), which pertains to implications of the form $\phi \rightarrow \psi$, where ϕ and ψ are truth functional, not containing any implications themselves. In [Anderson and Belnap, 1962], this logic has been defined proof-theoretically. Later, systems of formal semantics for this system were provided in [Routley and Routley, 1972], which propounded a two-valued semantics for fde, and in [Dunn, 1976], which propounded a four-valued semantics for fde. Here, Dunn's semantics will be used. For a system of natural deduction for fde, the reader may have recourse to the Appendix of this thesis.

Throughout the chapter, we use a language for propositional logic, built from an infinite set of propositional variables $PV = \{p, q, r, \ldots\}$.

4.2.1. Definition. [Language] The set of all formulas of fde, denoted by $\mathcal{F}$, is the least set satisfying the following conditions:

(i) $PV \subseteq \mathcal{F}$
(ii) If $\phi \in \mathcal{F}$ and $\psi \in \mathcal{F}$, then $(\phi \wedge \psi) \in \mathcal{F}$ and $(\phi \vee \psi) \in \mathcal{F}$
(iii) If $\phi \in \mathcal{F}$, then $\neg\phi \in \mathcal{F}$.

Note that $\top$ and $\bot$ do not occur among the formulas of the object language.

4.2.2. Definition. [Valuations] Let w be a set of propositional literals. Then w is a *valuation.*

In the following, we shall use v and w to denote valuations. A valuation w is inconsistent if and only if for some propositional variable p both $p \in w$ and $\neg p \in w$. In Section 4.4, to define an extractor, we need the concept of the *co-valuation* of a valuation, where such contradictory propositional variables are eliminated from the valuation. So, if $w = \{p, \neg p, q\}$, then its co-valuation should be $\{q\}$:

4.2.3. Definition. [Co-Valuations] Let w be a valuation. Then its *co-valuation* $\overline{w}$ is defined to be:

$$\overline{w} \quad = \quad \{p \in w : \neg p \notin w\} \cup \{\neg p \in w : p \notin w\}.$$

Obviously, a valuation is consistent if and only if it is identical with its co-valuation.

Our basic semantics is just a rewriting of Dunn's semantics for fde. In Dunn's sense, a valuation is a map $\nu : PV \mapsto \wp(\{\text{TRUE}, \text{FALSE}\})$ from the set of propositional variables to subsets of the set of truth-values TRUE and FALSE. Hence, a propositional variable p can have both truth-values, only one truth-value, and no truth-value.

Here, a valuation w defines such a map as follows: $p \in w$ if and only if $\text{TRUE} \in \nu(p)$, and $\neg p \in w$ if and only if $\text{FALSE} \in \nu(p)$. As a consequence, given the valuation $\{p, \neg p, q\}$, it holds that $\nu(p) = \{\text{TRUE}, \text{FALSE}\}$, $\nu(q) = \{\text{TRUE}\}$, and $\nu(x) = \emptyset$ for all other propositional variables x.

The assignment of truth-values to literals is extended to all formulas of the language as follows:[4]

[4] Definite restrictions on the set of valuations give rise to other, familiar logics: if we consider only consistent valuations, the resulting logic is *Kleene's strong three-valued logic* [Kleene, 1952]; if we consider only total valuations, we have Priest's *logic of paradox* LP [Priest, 1979]; if we consider only consistent and total valuations, we end up with classical propositional logic.

4.2.4. Definition. [Semantical Rules] Let w be a valuation. Then

(i)	$w \models p$	iff	$p \in w,$	if $p \in PV$
(ii)	$w \dashv p$	iff	$\neg p \in w,$	if $p \in PV$
(iii)	$w \models \phi \wedge \psi$	iff	$w \models \phi$ and $w \models \psi$	
(iv)	$w \dashv \phi \wedge \psi$	iff	$w \dashv \phi$ or $w \dashv \psi$	
(v)	$w \models \phi \vee \psi$	iff	$w \models \phi$ or $w \models \psi$	
(vi)	$w \dashv \phi \vee \psi$	iff	$w \dashv \phi$ and $w \dashv \psi$	
(vii)	$w \models \neg\phi$	iff	$w \dashv \phi$	
(viii)	$w \dashv \neg\phi$	iff	$w \models \phi.$	

For instance, $\{p, \neg p, q\} \models p \wedge \neg p$, but $\{p, \neg p, q\} \not\models q \wedge \neg q$. Hence, fde semantics keeps inconsistencies *local*. An inconsistency does not lead to triviality, in the sense that from an inconsistency everything follows.

Validity is defined in the obvious way:

4.2.5. Definition. [Validity] Let ϕ and ψ be formulas. Then ϕ *implies* ψ, denoted by $\phi \models \psi$, is defined as follows:

$$\phi \models \psi \quad \text{iff} \quad \forall w (w \models \phi \rightarrow w \models \psi).$$

In the following, we shall frequently use a partial ordering on the set of valuations. The ordering is given by set-theoretical inclusion. This simple ordering enables us to define least valuations satisfying a formula.

4.2.6. Definition. [Minimality] Let ϕ be a formula. A valuation w is ϕ-*minimal* if both

(i) $w \models \phi$
(ii) $\forall v (v \subset w \rightarrow v \not\models \phi)$.

Of course, several valuations can be ϕ-minimal (the valuations $\{p\}$ and $\{\neg q\}$ both are $p \vee \neg q$-minimal), but every formula defines a non-empty unique set of smallest finite valuations satisfying that formula. This set contains exactly all different fde-valuations that are minimally sufficient for validating the formula under consideration. A subset – empty, if the formula is inconsistent – of this set is the set of consistent ϕ-minimal valuations. These notions are defined as follows:

4.2.7. Definition. [Minimal Valuations] Let ϕ be a formula. Then

$$[\![\phi]\!] \quad = \quad \{w : w \text{ is } \phi\text{-minimal}\}.$$

We shall use $[\![\phi]\!]^{\top}$ to denote the set of consistent valuations in $[\![\phi]\!]$, and likewise, $[\![\phi]\!]^{\perp}$ to denote the the set of inconsistent valuations in $[\![\phi]\!]$.

Of course, $[\![\phi]\!] = [\![\phi]\!]^{\top} \cup [\![\phi]\!]^{\perp}$ and $[\![\phi]\!]^{\top} \cap [\![\phi]\!]^{\perp} = \emptyset$.

So, $[\![(p \vee q) \wedge \neg p]\!]^{\top} = \{\{q, \neg p\}\}$ and $[\![(p \vee q) \wedge \neg p]\!]^{\perp} = \{\{p, \neg p\}\}$.

These definitions entail that if the set of ϕ-minimal valuations is identical with the set of ψ-minimal valuations, then ϕ and ψ are equivalent under fde. The converse holds as well:

4.2.8. LEMMA (EXTENSIONALITY). *Let ϕ and ψ be formulas. Then*

$$[\![\phi]\!] = [\![\psi]\!] \quad \textit{iff} \quad \phi \models \psi \textit{ and } \psi \models \phi.$$

Proof. Assume that $[\![\phi]\!] = [\![\psi]\!]$. Suppose that $w \models \phi$. Sift this w to obtain v such that $v \subseteq w$ and $v \in [\![\phi]\!]$. By our assumption, it must be that $v \in [\![\psi]\!]$. Then $v \models \psi$. Hence, as $v \subseteq w$, we have $w \models \psi$. Therefore, $\phi \models \psi$. As the other case is proved similarly, it must be that $\phi \models \psi$ and $\psi \models \phi$.

Assume that $\phi \models \psi$ and $\psi \models \phi$. Suppose that $w \in [\![\phi]\!]$. Then $w \models \phi$ and, hence, $w \models \psi$. Suppose that $w \notin [\![\psi]\!]$. Then there must be a v such that $v \subset w$ and $v \models \psi$. Then, as we assumed that $\psi \models \phi$, we have $v \models \phi$, contradicting the ϕ-minimality of w. Hence $w \in [\![\psi]\!]$. The other inclusion is proved similarly. Therefore, $[\![\phi]\!] = [\![\psi]\!]$. □

4.2.1 An Algorithm for Finding ϕ's Minimal Valuations

Finding ϕ's minimal valuations can be cumbersome. How does one know that no single minimal valuation has been overlooked? An algorithm that lists them all gives the necessary assurance. Before formulating such an algorithm, we need some operations on valuations and sets of valuations.

4.2.9. DEFINITION. [Set Minimalization] Let $\mathcal{W}$ be a set of valuations. Then the *minimalization of* $\mathcal{W}$, denoted by $\mathsf{min}(\mathcal{W})$, is defined to be

$$\mathsf{min}(\mathcal{W}) \quad = \quad \{w \in \mathcal{W} : \forall v (v \in \mathcal{W} \rightarrow v \not\subset w)\}.$$

4.2.10. DEFINITION. [Minimal Union and Minimal Product] Let ϕ and ψ be formulas. Then the *minimal union of* $[\![\phi]\!]$ *and* $[\![\psi]\!]$, denoted by $[\![\phi]\!] \otimes [\![\psi]\!]$, and the *minimal product of* $[\![\phi]\!]$ *and* $[\![\psi]\!]$, denoted by $[\![\phi]\!] \oplus [\![\psi]\!]$, are defined to be

$$\begin{array}{lrcl} \text{(i)} & [\![\phi]\!] \otimes [\![\psi]\!] & = & \mathsf{min}([\![\phi]\!] \cup [\![\psi]\!]) \\ \text{(ii)} & [\![\phi]\!] \oplus [\![\psi]\!] & = & \mathsf{min}(\{v \cup w : v \in [\![\phi]\!] \text{ and } w \in [\![\psi]\!]\}). \end{array}$$

The algorithm is based on the following Deconstruction Rules. We prove the correctness of these rules immediately. We avoid a direct definition of $[\![\neg\phi]\!]$ in terms of some set-theoretical operation on $[\![\phi]\!]$ by splitting cases according to the main connective of the negated formula.

4.2.11. Lemma (Deconstruction Rules). *Let ϕ and ψ be formulas. Then*

(i)	$[\![p]\!]$	$=$	$\{\{p\}\},$	*if* $p \in PV$
(ii)	$[\![\neg p]\!]$	$=$	$\{\{\neg p\}\},$	*if* $p \in PV$
(iii)	$[\![\phi \wedge \psi]\!]$	$=$	$[\![\phi]\!] \oplus [\![\psi]\!]$	
(iv)	$[\![\phi \vee \psi]\!]$	$=$	$[\![\phi]\!] \otimes [\![\psi]\!]$	
(v)	$[\![\neg(\phi \wedge \psi)]\!]$	$=$	$[\![\neg\phi]\!] \otimes [\![\neg\psi]\!]$	
(vi)	$[\![\neg(\phi \vee \psi)]\!]$	$=$	$[\![\neg\phi]\!] \oplus [\![\neg\psi]\!]$	
(vii)	$[\![\neg\neg\phi]\!]$	$=$	$[\![\phi]\!].$	

Proof. (i) and (ii) are obvious.

(iii) Suppose that $w \in [\![\phi \wedge \psi]\!]$. Then $w \models \phi$ and $w \models \psi$. Sift w to obtain a $w_1 \in [\![\phi]\!]$ and a $w_2 \in [\![\psi]\!]$. Obviously, $w_1 \cup w_2 \subseteq w$. Moreover, the strong inclusion does not hold: suppose that $w_1 \cup w_2 \subset w$. Then, because of the $\phi \wedge \psi$-minimality of w, it must be that $w_1 \cup w_2 \not\models \phi \wedge \psi$. Hence, $w_1 \not\models \phi$ or $w_2 \not\models \psi$, contradicting the fact that $w_1 \in [\![\phi]\!]$ and $w_2 \in [\![\psi]\!]$. Therefore, there are $w_1 \in [\![\phi]\!]$ and $w_2 \in [\![\psi]\!]$, such that $w = w_1 \cup w_2$.

It remains to be shown that w is an element of the *minimal* product of $[\![\phi]\!]$ and $[\![\psi]\!]$. Suppose it is not. Then there are $v_1 \in [\![\phi]\!]$ and $v_2 \in [\![\psi]\!]$, such that $v_1 \cup v_2 \subset w$. But then, because of the $\phi \wedge \psi$-minimality of w, it must be that $v_1 \cup v_2 \not\models \phi \wedge \psi$. Hence, $v_1 \not\models \phi$ or $v_2 \not\models \psi$, contradicting the fact that $v_1 \in [\![\phi]\!]$ and $v_2 \in [\![\psi]\!]$. Therefore, w is an element of the minimal product of $[\![\phi]\!]$ and $[\![\psi]\!]$ – that is, $w \in [\![\phi]\!] \oplus [\![\psi]\!]$.

Suppose that $w \in [\![\phi]\!] \oplus [\![\psi]\!]$. Then there are $w_1 \in [\![\phi]\!]$ and $w_2 \in [\![\psi]\!]$, such that $w = w_1 \cup w_2$. Obviously, $w \models \phi \wedge \psi$. Suppose that w is not $\phi \wedge \psi$-minimal. Then there is a v, such that $v \subset w$ and $v \models \phi \wedge \psi$. Hence $v \models \phi$ and $v \models \psi$. Sift this v to obtain $v_1 \in [\![\phi]\!]$ and $v_2 \in [\![\psi]\!]$. Of course, $v_1 \cup v_2 \subseteq v \subset w$. Hence, $v_1 \cup v_2 \subset w$. But then $w \notin [\![\phi]\!] \oplus [\![\psi]\!]$ and we have a contradiction. Therefore, w is $\phi \wedge \psi$-minimal – that is, $w \in [\![\phi \wedge \psi]\!]$.

(iv) Suppose that $w \in [\![\phi \vee \psi]\!]$. Then $w \models \phi \vee \psi$ and $\forall v(v \subset w \rightarrow v \not\models \phi \vee \psi)$, that is, $\forall v(v \subset w \rightarrow v \not\models \phi)$ and $\forall v(v \subset w \rightarrow v \not\models \psi)$. Hence $w \in [\![\phi]\!] \cup [\![\psi]\!]$. It remains to be shown that w is in the minimalization of $[\![\phi]\!] \cup [\![\psi]\!]$. Suppose it is not. Then there is a $v \in [\![\phi]\!] \cup [\![\psi]\!]$ with $v \subset w$. Then, as $v \in [\![\phi]\!] \cup [\![\psi]\!]$, it must be that $v \models \phi$ or $v \models \psi$. But, as $v \subset w$, by the $\phi \vee \psi$-minimality of w, it must be that $v \not\models \phi$ and $v \not\models \psi$: a contradiction. Therefore, $w \in \mathsf{min}([\![\phi]\!] \cup [\![\psi]\!])$, that is, $w \in [\![\phi]\!] \otimes [\![\psi]\!]$.

Suppose that $w \in [\![\phi]\!] \otimes [\![\psi]\!]$. Then $w \in \mathsf{min}([\![\phi]\!] \cup [\![\psi]\!])$. Then $w \in [\![\phi]\!] \cup [\![\psi]\!]$ and $\forall v(v \in [\![\phi]\!] \cup [\![\psi]\!] \rightarrow v \not\subset w)$. Hence, $w \in [\![\phi]\!]$ or $w \in [\![\psi]\!]$.

Case 1: Suppose that $w \in [\![\phi]\!]$. Then $w \models \phi$ and $\forall v(v \subset w \rightarrow v \not\models \phi)$. Hence, $w \models \phi \vee \psi$. Suppose that there is a v with $v \subset w$ and $v \models \phi \vee \psi$. As w is ϕ-minimal, it must be that $v \not\models \phi$. Hence, $v \models \psi$. Sift v to obtain a v_1 with $v_1 \subseteq v$ and $v_1 \in [\![\psi]\!]$. Then $v_1 \subset w$ and $v_1 \in [\![\phi]\!] \cup [\![\psi]\!]$: a contradiction. Therefore, $w \in [\![\phi \vee \psi]\!]$.

Case 2: Suppose that $w \in [\![\psi]\!]$. Then $w \models \psi$ and $\forall v(v \subset w \rightarrow v \not\models \psi)$. Hence, $w \models \phi \vee \psi$. Suppose that there is a v with $v \subset w$ and $v \models \phi \vee \psi$. As w is ψ-minimal, it must be that $v \not\models \psi$. Hence, $v \models \phi$. Sift v to obtain a v_1 with $v_1 \subseteq v$ and $v_1 \in [\![\phi]\!]$. Then $v_1 \subset w$ and $v_1 \in [\![\phi]\!] \cup [\![\psi]\!]$: a contradiction. Therefore, $w \in [\![\phi \vee \psi]\!]$.

Therefore, $w \in [\![\phi \vee \psi]\!]$.

(v), (vi) and (vii) follow from the fact that the De Morgan rules and the Law of Double Negation hold for fde, from Lemma 4.2.8 and from (iii) and (iv). ☐

4.2.12. Definition. [Minimal Valuations Algorithm] Let ϕ be a formula. Then, the *Minimal Valuations Algorithm applied to* ϕ is defined as follows:

1. Put ϕ between double brackets. Then apply Deconstruction Rules (iii), (iv), (v), (vi) and (vii), until no further application of one of these Deconstruction Rules is possible. Use brackets, in order to avoid confusion.

2. Apply Deconstruction Rules (i) and (ii), and solve, bottom up, the operations $\oplus$ and $\otimes$ according to their definitions, until all occurences of $\oplus$ and $\otimes$ have been treated.

4.2.13. Theorem. *Let ϕ be a formula. Then the Minimal Valuations Algorithm applied to ϕ generates exactly all ϕ-minimal valuations.*

Proof. By structural induction on ϕ. Use Lemma 4.2.11. □

4.2.2 Finite States and Their Determiners

In possible worlds semantics, the meaning of a formula ϕ is identified with the set $[\phi]$ of all possible worlds that validate ϕ, where a possible world is a *total* and *consistent* valuation. Partial information is represented by a multitude of possible worlds. In the present context, the meaning of a formula ϕ is identified with the set $[\![\phi]\!]$ of *minimal* valuations required to validate ϕ. Restricting our discussion to the consistent case, these consistent minimal valuations, which then are partial valuations of Kleene's strong three-valued logic, state the minimal requirements possible worlds must fulfill in order to be included in $[\phi]$. Hence, a possible world w is an element of $[\phi]$ if and only if there is a consistent v in $[\![\phi]\!]$ such that $v \subseteq w$. Thus, it seems that the same information is conveyed by $[\phi]$ and by $[\![\phi]\!]$. However, the accounts differ in their notion of relevance, which has an important repercussion for the concept of 'information span', which is discussed below.

Relevance is definitely *not* a characteristic of the usual possible worlds approach. Each possible world that validates "Góngora was portrayed by Velázquez" is opinioned on every other formula of the language, regardless of its relevance to the subject matter of the represented information. This huge amount of additional 'information' is the price paid for using a set of *total* valuations to represent *partial* information. In this chapter, we avoid this situation by concentrating on the set of minimal valuations required to validate the information to be represented. Thus, a notion of relevance (albeit a crude one) is central to our way of information representation.

In the inconsistent case, possible worlds semantics could be construed using, instead of classical logic, Priest's LP as underlying logic. LP employs total, but possibly inconsistent valuations. Thus, we would stay as close as possible to the standard possible worlds semantics. (See for instance [Mares, 1998] on 'counterpossible conditionals'). Here, we pursue the same strategy as in the consistent case sketched above. We shall use minimal fde-valuations to describe the minimal requirements an 'inconsistent world' must fulfill in order to be included in $[\phi]$.

In a possible worlds framework, an *epistemic state* σ is identified with a set of possible worlds. A formula ϕ is *supported* by σ if and only if every world w in σ validates ϕ. Here, following our intentions to discuss epistemic changes using minimal valuations, we shall define analogues of these notions. A *finite state* $\mathcal{K}$ is a set of valuations that satisfies four conditions:

4.2.14. Definition. [Finite State] Let $\mathcal{K}$ be a set of valuations. Then $\mathcal{K}$ is a *finite state*, if

(i) $\mathcal{K} \neq \emptyset$,
(ii) $\mathcal{K}$ is finite,
(iii) every w in $\mathcal{K}$ is finite,
(iv) $\forall v \forall w((v \in \mathcal{K} \text{ and } w \in \mathcal{K}) \rightarrow v \not\subset w)$.

If $\mathcal{K} = \{\emptyset\}$, then $\mathcal{K}$ is *trivial*.

Since every formula of the language is satisfiable in fde, a finite state, which is a representation of all the information we recieved thus far, can not be empty. In case we do not have any information at all, the finite state is *trivial* and should not impose any constraint on the choice of minimal valuations validating the incoming information. This situation is represented adequately by the finite state $\{\emptyset\}$. As $\emptyset$ is a subset of *all* total valuations w, the possible worlds analogue of $\{\emptyset\}$ would be the set of all possible worlds, which indeed represents the epistemic state of total ignorance.

Every formula ϕ has a finite number of propositional variables, so we need only consider finitely many valuations under which that formula is true. Moreover, each valuation itself will be finite as well, as the number of its truth-assigments to propositional variables is bounded by ϕ's number of propositional variables. Therefore,

4.2.15. Lemma. *Let ϕ be a formula. Then $[\![\phi]\!]$ is a finite state.*

Each non-trivial finite state $\mathcal{K}$ has a formula $\partial(\mathcal{K})$, that characterizes it. In [Gärdenfors, 1988, p. 26], this formula is called the *determiner* of $\mathcal{K}$. Contrary to Gärdenfors's notion of a determiner, we do not need to extend the language to incorporate infinite conjunctions. The determiner of a non-empty, finite set of valuations $\mathcal{W}$ such that every $w_i \in \mathcal{W}$ is finite is just a formula of the language $\mathcal{F}$:

4.2.16. Definition. [Determiner] Let $\mathcal{W} = \{w_1, \ldots, w_n\}$ a non-empty, finite set of valuations such that each $w_i \in \mathcal{W}$ is finite. Then the *determiner of* $\mathcal{W}$, denoted by $\partial(\mathcal{W})$, is defined to be

$$\partial(\mathcal{W}) \quad = \quad \bigvee_{i=1}^{n} (\bigwedge w_i).$$

As all the w_i's in $\mathcal{W}$ are finite, $\partial(\mathcal{W})$ is a formula.

For example, let $\mathcal{W}$ be $\{\{p, \neg q\}, \{p, \neg p\}, \{q\}\}$. Then $\mathcal{W}$ is a non-empty, finite set and each $w_i \in \mathcal{W}$ is finite. Hence, $\partial(\mathcal{W}) = (p \wedge \neg q) \vee (p \wedge \neg p) \vee q$.

The next lemma shows that the determiner $\partial(\mathcal{K})$ characterizes a non-trivial finite state $\mathcal{K}$. Note that the *trivial* finite state does not have a determiner.[5]

4.2.17. Lemma (Characterization). *Let $\mathcal{K} = \{w_1, \ldots, w_n\}$ be a non-trivial finite state. Then*

$$[\![\partial(\mathcal{K})]\!] \quad = \quad \mathcal{K}.$$

Proof. Suppose that $w \in [\![\partial(\mathcal{K})]\!]$. Then, (i) $w \models \bigvee_{i=1}^{n}(\bigwedge w_i)$, and (ii) $\forall v(v \subset w \rightarrow v \not\models \partial(\mathcal{K}))$. From (i) it follows by the Semantical Rules of fde that there is an i with $1 \leq i \leq n$ such that $w \models \bigwedge w_i$. Hence, there is an i with $1 \leq i \leq n$ such that $w_i \subseteq w$. Suppose that $w_i \subset w$. Obviously, $w_i \models \bigwedge w_i$ and, therefore, $w_i \models \partial(\mathcal{K})$. By (ii), however, it must be that $w_i \not\models \partial(\mathcal{K})$, contradicting the previous statement. Hence, $w_i \not\subset w$ and thus $w_i = w$. Therefore, $w \in \mathcal{K}$.

Suppose that $w \in \mathcal{K}$, that is, there is an i with $1 \leq i \leq n$ such that $w = w_i$. Obviously, $w_i \models \bigwedge w_i$ and, hence, $w_i \models \partial(\mathcal{K})$. Therefore, $w \models \partial(\mathcal{K})$. Suppose $\exists v(v \subset w \textit{ and } v \models \partial(\mathcal{K}))$. Sift this v to obtain a v' such that $v' \in [\![\partial(\mathcal{K})]\!]$ and $v' \subseteq v$. Then $v' \subset w$, and, by the reasoning in the first part of this proof, from $v' \in [\![\partial(\mathcal{K})]\!]$ we obtain $v' \in \mathcal{K}$. Hence, there are $w \in \mathcal{K}$ and $v' \in \mathcal{K}$ such that $v' \subset w$, contradicting the fact that $\mathcal{K}$ is a finite state. Hence, $\forall v(v \subset w \rightarrow v \not\models \partial(\mathcal{K}))$. Therefore, $w \in [\![\partial(\mathcal{K})]\!]$. □

4.2.18. Lemma. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be finite states. Suppose that both*

(i) $\forall v(v \in \mathcal{K}_1 \rightarrow \exists w(w \in \mathcal{K}_2 \textit{ and } v \subseteq w))$
(ii) $\forall w(w \in \mathcal{K}_2 \rightarrow \exists v(v \in \mathcal{K}_1 \textit{ and } w \subseteq v))$.

Then $\mathcal{K}_1 = \mathcal{K}_2$.

Proof. Suppose v in $\mathcal{K}_1$. Then there is a w in $\mathcal{K}_2$ such that $v \subseteq w$. Suppose that $v \neq w$. Then $v \subset w$ and there is a $v' \in \mathcal{K}_1$ such that $w \subseteq v'$. Hence, there are v and v' in $\mathcal{K}_1$ such that $v \subset v'$, contradicting the fact that $\mathcal{K}_1$ is a finite state. Therefore, $v = w$. Hence, it must be that $\forall v(v \in \mathcal{K}_1 \rightarrow \exists w(w \in \mathcal{K}_2 \textit{ and } v = w))$. By the same reasoning, starting with a w in $\mathcal{K}_2$, it must be that $\forall w(w \in \mathcal{K}_2 \rightarrow \exists v(v \in \mathcal{K}_1 \textit{ and } w = v))$. Hence, $\mathcal{K}_1 \subseteq \mathcal{K}_2$ and $\mathcal{K}_2 \subseteq \mathcal{K}_1$. Therefore, $\mathcal{K}_1 = \mathcal{K}_2$. □

A finite state $\mathcal{K}$ *supports* a formula ϕ, if ϕ is validated by *all* valuations in $\mathcal{K}$. This notion of support must be distinguished from the notion of 'information span' which will be discussed more thoroughly in the next subsection. A formula

[5] One might consider to introduce a formula $\top$ in the language, which is true under all valuations, and stipulate that $\partial(\{\emptyset\}) = \top$. This addition proves to be unnecessary for the results in this chapter and poses additional technical complications, which are the main reasons for not following this policy.

ϕ is within the *information span* of a finite state $\mathcal{K}$, if ϕ is validated by *at least one* valuation in $\mathcal{K}$. It should be noted that the latter concept does not function in a standard possible worlds setting: *every* formula that is consistent with an epistemic state σ would be, regardless of the state's content, within that state's information span, for each world in σ is fully opinioned.

To conclude this subsection, we note the fact that $\mathcal{K}$ supports ϕ if and only if $\mathcal{K}$'s determiner implies ϕ:

4.2.19. Lemma. *Let $\mathcal{K}$ be a non-trivial finite state and let ϕ be a formula. Then*

$$\partial(\mathcal{K}) \models \phi \quad \textit{iff} \quad \forall w(w \in \mathcal{K} \rightarrow w \models \phi).$$

Proof. Let $\mathcal{K}$ be $\{w_1, \ldots, w_n\}$. Assume that $\partial(\mathcal{K}) \models \phi$. Suppose that $w \in \mathcal{K}$. Then there is an i with $1 \leq i \leq n$ such that $w = w_i$. Obviously, $w \models \bigwedge w_i$. Hence, $w \models \bigvee_{i=1}^{n}(\bigwedge w_i)$. Then $w \models \partial(\mathcal{K})$. Therefore, by assumption, $w \models \phi$.

Assume that $\forall w(w \in \mathcal{K} \rightarrow w \models \phi)$. Suppose that $v \models \partial(\mathcal{K})$. Then $v \models \bigvee_{i=1}^{n}(\bigwedge w_i)$. Hence, there is an i with $1 \leq i \leq n$ such that $v \models \bigwedge w_i$. Obviously, $w_i \subseteq v$. As $w_i \in \mathcal{K}$, by assumption, $w_i \models \phi$. Therefore, $v \models \phi$. □

4.2.3 Finite States: Information Span

Before we are in a position to study the properties of the information dynamics of finite states, we must be able to say something about the information that is within the information span of a finite state. In the present section, this notion of 'information span' shall be explored. The definition below enables us to define a relation on pairs of finite states ('The information span of $\mathcal{K}_1$ is contained in the information span of $\mathcal{K}_2$'). Information spans play a crucial role in our representation theorem for contraction later on.

4.2.20. Definition. [Information Span] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then ϕ *is in the information span of* $\mathcal{K}$, notation: $\phi \sqsubseteq \mathcal{K}$, if the following holds:

$$\phi \sqsubseteq \mathcal{K} \quad \text{iff} \quad \exists w(w \in \mathcal{K} \textit{ and } w \models \phi).$$

'The information span of $\mathcal{K}_1$ is contained in the information span of $\mathcal{K}_2$' can now be expressed as follows:

4.2.21. Definition. Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be finite states. Then

$$\mathcal{K}_1 \trianglelefteq \mathcal{K}_2 \quad \text{iff} \quad \forall \phi(\phi \sqsubseteq \mathcal{K}_1 \rightarrow \phi \sqsubseteq \mathcal{K}_2).$$

From these definitions it follows that if two states have the same information span, they are identical:

4.2.22. LEMMA. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be finite states. Then*

$$\mathcal{K}_1 \trianglelefteq \mathcal{K}_2 \text{ and } \mathcal{K}_2 \trianglelefteq \mathcal{K}_1 \quad \text{iff} \quad \mathcal{K}_1 = \mathcal{K}_2.$$

Proof. Suppose that $\mathcal{K}_1 \trianglelefteq \mathcal{K}_2$ and $\mathcal{K}_2 \trianglelefteq \mathcal{K}_1$. As $\mathcal{K}_1$ is a finite state, it has a finite number of elements. Let $\mathcal{K}_1 = \{v_1, \ldots, v_n\}$. For all v in $\mathcal{K}_1$ it holds that $v \models \bigwedge v$. Hence, for all v in $\mathcal{K}_1$ it holds that $\bigwedge v \sqsubseteq \mathcal{K}_1$. Hence, by supposition, for all v in $\mathcal{K}_1$ it holds that $\bigwedge v \sqsubseteq \mathcal{K}_2$. Hence, by definition, for all v in $\mathcal{K}_1$ there is a w in $\mathcal{K}_2$, such that $w \models \bigwedge v$. Hence, by the Semantical Rules of fde, for all v in $\mathcal{K}_1$ there is a w in $\mathcal{K}_2$ such that $v \subseteq w$, that is, $\forall v(v \in \mathcal{K}_1 \rightarrow \exists w(w \in \mathcal{K}_2 \textit{ and } v \subseteq w))$. By the same reasoning, it must be that $\forall w(w \in \mathcal{K}_2 \rightarrow \exists v(v \in \mathcal{K}_1 \textit{ and } w \subseteq v))$. By Lemma 4.2.18, $\mathcal{K}_1 = \mathcal{K}_2$.

The converse is obvious. □

Hence the relation $\trianglelefteq$ has some properties analogous to set-theoretical inclusion. Let us introduce union and intersection for finite states as well:

4.2.23. DEFINITION. [Set Maximalization] Let $\mathcal{W}$ be a set of valuations. Then the *maximalization of* $\mathcal{W}$, denoted by $\mathsf{max}(\mathcal{W})$, is defined to be

$$\mathsf{max}(\mathcal{W}) \quad = \quad \{w \in \mathcal{W} : \forall v(v \in \mathcal{W} \rightarrow w \not\subset v)\}.$$

4.2.24. DEFINITION. [Maximal Union and Intersection] Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be finite states. Then the *maximal union of* $\mathcal{K}_1$ *and* $\mathcal{K}_2$, denoted by $\mathcal{K}_1 \sqcup \mathcal{K}_2$, and the *maximal intersection of* $\mathcal{K}_1$ *and* $\mathcal{K}_2$, denoted by $\mathcal{K}_1 \sqcap \mathcal{K}_2$, are defined to be

(i) $\mathcal{K}_1 \sqcup \mathcal{K}_2 \;=\; \mathsf{max}(\mathcal{K}_1 \cup \mathcal{K}_2)$
(ii) $\mathcal{K}_1 \sqcap \mathcal{K}_2 \;=\; \mathsf{max}(\{v \cap w : v \in \mathcal{K}_1 \text{ and } w \in \mathcal{K}_2\})$.

Union and intersection for finite states behave as they should:

4.2.25. LEMMA. *Let $\mathcal{K}$ be a finite state. Then*

(i) $\mathcal{K} \sqcup \mathcal{K} \;=\; \mathcal{K}$
(ii) $\mathcal{K} \sqcap \mathcal{K} \;=\; \mathcal{K}$.

The following two lemmas extend the analogy between the relation $\trianglelefteq$ and set-theoretical inclusion. Since it is easy to check that the first lemma holds, only the second is proved.

4.2.26. LEMMA. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be finite states. Then*

(i) $\mathcal{K}_1 \sqcap \mathcal{K}_2 \trianglelefteq \mathcal{K}_1$
(ii) $\mathcal{K}_1 \sqcap \mathcal{K}_2 \trianglelefteq \mathcal{K}_2$
(iii) $\mathcal{K}_1 \trianglelefteq \mathcal{K}_1 \sqcup \mathcal{K}_2$
(iv) $\mathcal{K}_2 \trianglelefteq \mathcal{K}_1 \sqcup \mathcal{K}_2$.

4.2.27. Lemma. *Let $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_3$, and $\mathcal{K}_4$ be finite states. Suppose that $\mathcal{K}_1 \trianglelefteq \mathcal{K}_3$ and $\mathcal{K}_2 \trianglelefteq \mathcal{K}_4$. Then*

(i) $\mathcal{K}_1 \sqcup \mathcal{K}_2 \trianglelefteq \mathcal{K}_3 \sqcup \mathcal{K}_4$
(ii) $\mathcal{K}_1 \sqcap \mathcal{K}_2 \trianglelefteq \mathcal{K}_3 \sqcap \mathcal{K}_4$.

Proof. (i) Suppose that $\phi \sqsubseteq \mathcal{K}_1 \sqcup \mathcal{K}_2$. Then $\phi \sqsubseteq \mathsf{max}(\mathcal{K}_1 \cup \mathcal{K}_2)$. Hence, there is a v in $\mathsf{max}(\mathcal{K}_1 \cup \mathcal{K}_2)$ such that $v \models \phi$. It holds that $v \in \mathcal{K}_1$ or $v \in \mathcal{K}_2$. Hence, $\phi \sqsubseteq \mathcal{K}_1$ or $\phi \sqsubseteq \mathcal{K}_2$. Hence, by the supposition that $\mathcal{K}_1 \trianglelefteq \mathcal{K}_3$ and $\mathcal{K}_2 \trianglelefteq \mathcal{K}_4$, it must be that $\phi \sqsubseteq \mathcal{K}_3$ or $\phi \sqsubseteq \mathcal{K}_4$. Therefore, by Lemma 4.2.26(iii) and (iv), $\phi \sqsubseteq \mathcal{K}_3 \sqcup \mathcal{K}_4$.

(ii) Suppose that $\phi \sqsubseteq \mathcal{K}_1 \sqcap \mathcal{K}_2$. Then $\phi \sqsubseteq \mathsf{max}(\{v \cap w : v \in \mathcal{K}_1 \textit{ and } w \in \mathcal{K}_2\})$. Hence, there are v_i in $\mathcal{K}_1$ and v_j in $\mathcal{K}_2$ such that $v_i \cap v_j \models \phi$. Consider v_i. Obviously, it holds that $v_i \models \bigwedge v_i$. Hence, it holds that $\bigwedge v_i \sqsubseteq \mathcal{K}_1$. Hence, by the assumption that $\mathcal{K}_1 \trianglelefteq \mathcal{K}_3$, it must be that $\bigwedge v_i \sqsubseteq \mathcal{K}_3$. Hence, there is a $w_k \in \mathcal{K}_3$ such that $v_i \subseteq w_k$. By the same reasoning, there is a $w_l \in \mathcal{K}_4$ such that $v_j \subseteq w_l$. Obviously, $v_i \cap v_j \subseteq w_k \cap w_l$. Hence, $w_k \cap w_l \models \phi$. As the maximalization of $\{v \cap w : v \in \mathcal{K}_3 \textit{ and } w \in \mathcal{K}_4\}$ only skips $w_k \cap w_l$ in case there is a w' in $\{v \cap w : v \in \mathcal{K}_3 \textit{ and } w \in \mathcal{K}_4\}$ such that $w_k \cap w_l \subset w'$, it must be that $\phi \sqsubseteq \mathsf{max}(\{v \cap w : v \in \mathcal{K}_3 \textit{ and } w \in \mathcal{K}_4\})$. Therefore, $\phi \sqsubseteq \mathcal{K}_3 \sqcap \mathcal{K}_4$. □

The quasi-set-theoretical notions developed in this section serve as tools for analysing the properties of contraction. The notion of a determiner of a finite state will be instrumental in probing expansion.

4.3 Operations of Information Change

4.3.1 Expansion

In classical systems of belief revision, expansion of a belief set $\mathcal{K}$ with a formula ϕ is the least problematic operation. So it is here. There is, nevertheless, a difference in interpretation of the operation. In the system to be propounded here, an expansion of the finite state $\mathcal{K}$ with a formula ϕ amounts to changing our finite state $\mathcal{K}$, which is a systematic representation of all the information we received thus far, to incorporate the information that ϕ. The result of the adaptation of our finite state is a new finite state $\mathcal{K} + \phi$.

4.3.1. Definition. [Expansion] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then the *expansion of $\mathcal{K}$ with ϕ*, denoted by $\mathcal{K} + \phi$, is defined as follows:

$$\mathcal{K} + \phi \quad = \quad \mathsf{min}(\{v \supseteq w : w \in \mathcal{K} \text{ and } v \models \phi\}).$$

In order to facilitate the proofs to come, we need the following lemma:

4.3.2. Lemma. *Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then*

$$\mathcal{K} + \phi \quad = \quad \mathcal{K} \oplus [\![\phi]\!].$$

Proof. Suppose that $v \in \mathsf{min}(\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\})$. Then $v \models \phi$ and there is a $w_i \in \mathcal{K}$ such that $w_i \subseteq v$. Take v_j in $[\![\phi]\!]$ with $v_j \subseteq v$. It holds that $w_i \in \mathcal{K}$ and $v_j \models \phi$, and $v_j \cup w_i \supseteq w_i$. Hence, $w_i \cup v_j \in \{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. Obviously, $w_i \cup v_j \subseteq v$. Hence, it must be that $w_i \cup v_j = v$, since otherwise v would have been skipped by the minimalization of $\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$, which would contradict our initial supposition. Therefore, $v \in \{v \cup w : w \in \mathcal{K} \text{ and } v \in [\![\phi]\!]\}$. Suppose that $v \notin \mathsf{min}(\{v \cup w : w \in \mathcal{K} \textit{ and } v \in [\![\phi]\!]\})$. Then there must be a v' in $\{v \cup w : w \in \mathcal{K} \textit{ and } v \in [\![\phi]\!]\}$ such that $v' \subset v$. Then there must be a w'_k in $\mathcal{K}$ and a v'_l in $[\![\phi]\!]$ such that $v' = w'_k \cup v'_l$. It holds that $w'_k \in \mathcal{K}$ and $v'_l \models \phi$ and $w'_k \cup v'_l \supseteq w'_k$. Hence, $w'_k \cup v'_l \in \{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. Therefore, $v' \in \{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. But then v is skipped by the minimalization of $\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. Hence, $v \notin \{v \supseteq w : w \in \mathcal{K} \text{ and } v \models \phi\}$: a contradiction. Therefore, $v \in \mathsf{min}(\{v \cup w : w \in \mathcal{K} \textit{ and } v \in [\![\phi]\!]\})$, that is, $v \in \mathcal{K} \oplus [\![\phi]\!]$.

To prove the other inclusion, suppose that $v \in \mathsf{min}(\{v \cup w : w \in \mathcal{K} \text{ and } v \in [\![\phi]\!]\})$. Then there is a w_i in $\mathcal{K}$ and a v_j in $[\![\phi]\!]$ such that $v = w_i \cup v_j$. It holds that $w_i \in \mathcal{K}$ and $w_i \cup v_j \models \phi$ and $w_i \cup v_j \supseteq w_i$. Hence, $w_i \cup v_j \in \{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. Therefore, $v \in \{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$. Suppose that $v \notin \mathsf{min}(\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\})$. Then there must be a v' in $\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\}$ such that $v' \subset v$. Then $v' \models \phi$ and there is a $w'_k \in \mathcal{K}$ such that $w'_k \subseteq v'$. Take v'_l in $[\![\phi]\!]$ with $v'_l \subseteq v'$. Consider $w'_k \cup v'_l$. It holds that $w'_k \cup v'_l \subseteq v'$. Hence, $w'_k \cup v'_l \subset v$. Moreover, since $w'_k \in \mathcal{K}$ and $v'_l \in [\![\phi]\!]$, it must be that $w'_k \cup v'_l \in \{v \cup w : w \in \mathcal{K} \text{ and } v \in [\![\phi]\!]\}$. But then v is skipped by the minimalization of $\{v \cup w : w \in \mathcal{K} \text{ and } v \in [\![\phi]\!]\}$. Hence, $v \notin \mathsf{min}(\{v \cup w : w \in \mathcal{K} \text{ and } v \in [\![\phi]\!]\})$: a contradiction. Therefore, $v \in \mathsf{min}(\{v \supseteq w : w \in \mathcal{K} \textit{ and } v \models \phi\})$, that is, $v \in \mathcal{K} + \phi$. □

The next lemma will be put to use in our representation theorem for expansion. Note that if $\mathcal{K}$ is trivial, then, by definition, $\mathcal{K} + \phi = [\![\phi]\!]$.

4.3.3. Lemma. *Let $\mathcal{K}$ be a non-trivial finite state and let ϕ be a formula. Then*

$$\mathcal{K} + \phi \quad = \quad [\![\partial(\mathcal{K}) \wedge \phi]\!].$$

Proof. By Lemma 4.3.2, Lemma 4.2.17 and Lemma 4.2.11(iii). □

Properties of the Expansion Function

The following five postulates are similar to the first five (out of six) postulates in [Gärdenfors, 1988, p. 48–51]:[6]

[6] *Update semantics*, as set forth in [Veltman, 1996], can be mimicked with this expansion function, on the condition that the starting set σ of possible worlds always only contains worlds that differ finitely. σ must be describable by a *finite* formula.

4.3.4. Definition. [Postulates for Expansion] Let $\mathcal{K}$ and $\mathcal{K}'$ be non-trivial finite states and let ϕ be a formula. An *expansion operator*, †, is any operator satisfying the following postulates:

$(\mathcal{K} \dagger 1)$	$\mathcal{K} \dagger \phi$ is a non-trivial finite state
$(\mathcal{K} \dagger 2)$	$\partial(\mathcal{K} \dagger \phi) \models \phi$
$(\mathcal{K} \dagger 3)$	$\partial(\mathcal{K} \dagger \phi) \models \partial(\mathcal{K})$
$(\mathcal{K} \dagger 4)$	If $\partial(\mathcal{K}) \models \phi$, then $\partial(\mathcal{K}) \models \partial(\mathcal{K} \dagger \phi)$
$(\mathcal{K} \dagger 5)$	If $\partial(\mathcal{K}) \models \partial(\mathcal{K}')$, then $\partial(\mathcal{K} \dagger \phi) \models \partial(\mathcal{K}' \dagger \phi)$.

Our next theorem shows that, using finite states rather than belief sets, we do not, unlike Gärdenfors, need an additional postulate to secure uniqueness.

4.3.5. Theorem (Uniqueness). *Let † and ‡ be expansion operators, satisfying $(\mathcal{K} \dagger 1)$ through $(\mathcal{K} \dagger 5)$. Then for all formulas ϕ it holds that $\mathcal{K} \dagger \phi = \mathcal{K} \ddagger \phi$.*

Proof. It will suffice to show that both (a) $\partial(\mathcal{K} \dagger \phi) \models \partial(\mathcal{K} \ddagger \phi)$ and (b) $\partial(\mathcal{K} \ddagger \phi) \models \partial(\mathcal{K} \dagger \phi)$, for, by Lemma 4.2.8 and Lemma 4.2.17, it follows from (a) and (b) that $\mathcal{K} \dagger \phi = \mathcal{K} \ddagger \phi$. First, because of $(\mathcal{K} \dagger 1)$, both $\partial(\mathcal{K} \dagger \phi)$ and $\partial(\mathcal{K} \ddagger \phi)$ are formulas. By $(\mathcal{K} \dagger 2)$, it must be that $\partial(\mathcal{K} \dagger \phi) \models \phi$. Hence, by $(\mathcal{K} \dagger 4)$, we obtain (i) $\partial(\mathcal{K} \dagger \phi) \models \partial((\mathcal{K} \dagger \phi) \ddagger \phi)$. By $(\mathcal{K} \dagger 3)$, it must be that $\partial(\mathcal{K} \dagger \phi) \models \partial(\mathcal{K})$. Hence, by $(\mathcal{K} \dagger 5)$, we obtain (ii) $\partial((\mathcal{K} \dagger \phi) \ddagger \phi) \models \partial(\mathcal{K} \ddagger \phi)$. From (i) and (ii), it follows that (a). Of course, (b) can be proved analogously. Therefore, $\mathcal{K} \dagger \phi = \mathcal{K} \ddagger \phi$. □

4.3.6. Theorem (Representation Theorem for Expansion). *Let $\mathcal{K}$ be a non-trivial finite state and let ϕ be a formula. Then $\mathcal{K} \dagger \phi$ satisfies $(\mathcal{K} \dagger 1)$ through $(\mathcal{K} \dagger 5)$ iff $\mathcal{K} \dagger \phi = \mathcal{K} + \phi$.*

Proof. Because of Theorem 4.3.5, it will suffice to show that $\mathcal{K} + \phi$ satisfies $(\mathcal{K} \dagger 1)$ through $(\mathcal{K} \dagger 5)$. The first three follow from Lemma 4.3.3.

$(\mathcal{K} \dagger 4)$ Suppose that $\partial(\mathcal{K}) \models \phi$. Since fde is reflexive, it holds that $\partial(\mathcal{K}) \models \partial(\mathcal{K})$. Hence, $\partial(\mathcal{K}) \models \partial(\mathcal{K}) \wedge \phi$. From Lemma 4.3.3, we get $\partial(\mathcal{K}) \wedge \phi \models \partial(\mathcal{K} + \phi)$. Therefore, $\partial(\mathcal{K}) \models \partial(\mathcal{K} + \phi)$, since fde is transitive.

$(\mathcal{K} \dagger 5)$ Suppose that $\partial(\mathcal{K}) \models \partial(\mathcal{K}')$. From Lemma 4.3.3 we obtain both $\partial(\mathcal{K}+\phi) \models \phi$ and $\partial(\mathcal{K} + \phi) \models \partial(\mathcal{K})$. Combining the latter with our supposition, it must be that $\partial(\mathcal{K} + \phi) \models \partial(\mathcal{K}')$, since fde is transitive. Hence, $\partial(\mathcal{K} + \phi) \models \partial(\mathcal{K}') \wedge \phi$. From Lemma 4.3.3 we get $\partial(\mathcal{K}') \wedge \phi \models \partial(\mathcal{K}' + \phi)$. Therefore, $\partial(\mathcal{K} + \phi) \models \partial(\mathcal{K}' + \phi)$. □

The following theorem relates iterated expansion to conjunction.

4.3.7. Theorem (Iterated Expansion). *Let $\mathcal{K}$ be a finite state and let ϕ and ψ be formulas. Then*

$$(\mathcal{K} + \phi) + \psi \quad = \quad \mathcal{K} + (\phi \wedge \psi).$$

Proof. If $\mathcal{K} = \{\emptyset\}$, then for all χ it holds that $\mathcal{K} + \chi = [\![\chi]\!]$, by definition. Lemma 4.2.11(iii) does the job. Otherwise, by Lemma 4.3.2, it holds that $(\mathcal{K}+\phi)+\psi = (\mathcal{K} \oplus [\![\phi]\!]) \oplus [\![\psi]\!]$. By Lemma 4.2.17, $\mathcal{K}$ has a determiner such that $\mathcal{K} = [\![\partial(\mathcal{K})]\!]$. The proof can be completed, using Lemma 4.2.8, Lemma 4.2.11(iii), and Lemma 4.3.2. □

Since conjunction is commutative, the order of the formulas with which a finite state is expanded is irrelevant to the result:

4.3.8. Corollary. *Let $\mathcal{K}$ be a finite state and let ϕ and ψ be formulas. Then*

$$(\mathcal{K} + \phi) + \psi \quad = \quad (\mathcal{K} + \psi) + \phi.$$

4.3.2 Contraction

Intuitively, if we contract a finite state $\mathcal{K}$ with a formula ϕ, we skip all sufficient evidence for ϕ from the elements of our finite state, so that ϕ can not be among the formulas which are within the information span of the resulting finite state, as there will be no residual evidence for ϕ left in it. Hence, ϕ is not within the information span of $\mathcal{K} - \phi$. Moreover, all logical consequences which were in the information span of $\mathcal{K}$ 'just because' ϕ was within the information span of $\mathcal{K}$ will be removed as well.

First, we provide two definitions of contraction, one of which is, to the best of my knowledge, the first *recursive* definition of contraction in the literature, and secondly, we shall show that these definitions are equivalent. The recursive definition offers more insight into the handling of complex formulas than the direct definition. Moreover, we can use the results of earlier contractions of $\mathcal{K}$ to compute the contraction of $\mathcal{K}$ with a more complex formula. On the other hand, the direct definition is more convenient in case we study the metalogical properties of the system.

4.3.9. Definition. [Contraction] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then the *contraction of $\mathcal{K}$ with ϕ*, denoted by $\mathcal{K} - \phi$, is defined as follows:

$$\mathcal{K} - \phi \quad = \quad \mathsf{max}(\{v \subseteq w : w \in \mathcal{K} \textit{ and } v \not\models \phi\}).$$

4.3.10. Definition. [Recursive Contraction] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then the *recursive contraction of $\mathcal{K}$ with ϕ*, denoted by $\mathcal{K} \sim \phi$, is defined recursively as follows:

$$\begin{array}{rlcl}
\text{(i)} & \mathcal{K} \sim p & = & \mathsf{max}(\{w \backslash \{p\} : w \in \mathcal{K}\}) \\
\text{(ii)} & \mathcal{K} \sim \neg p & = & \mathsf{max}(\{w \backslash \{\neg p\} : w \in \mathcal{K}\}) \\
\text{(iii)} & \mathcal{K} \sim (\psi_1 \wedge \psi_2) & = & (\mathcal{K} \sim \psi_1) \sqcup (\mathcal{K} \sim \psi_2) \\
\text{(iv)} & \mathcal{K} \sim (\psi_1 \vee \psi_2) & = & (\mathcal{K} \sim \psi_1) \sqcap (\mathcal{K} \sim \psi_2).
\end{array}$$

These rules suffice to define the contraction of any formula ϕ from $\mathcal{K}$, since the De Morgan rules and the Law of Double Negation hold for fde. For instance, $\mathcal{K} \sim \neg(\psi_1 \wedge \psi_2)$ equals $\mathcal{K} \sim (\neg\psi_1 \vee \neg\psi_2)$, which, according to (iv), amounts to $(\mathcal{K} \sim \neg\psi_1) \sqcap (\mathcal{K} \sim \neg\psi_2)$.

By structural induction on ϕ, we can show that Contraction and Recursive Contraction are equivalent:

4.3.11. Theorem. *Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then*

$$\mathcal{K} - \phi \quad = \quad \mathcal{K} \sim \phi.$$

Properties of the Contraction Function

In the present context, Gärdenfors's postulates for contraction cannot be straightforwardly applied, as the standard postulates have been devised for belief sets. We propose to translate them into the language of finite states, using previous definitions. We only need four of Gärdenfors's eight postulates plus an additional postulate for our representation theorem. The reader can easily check, however, that all translations of Gärdenfors's eight postulates hold, except for the hotly debated *Recovery* postulate.[7]

4.3.12. Definition. [Postulates for Contraction] Let $\mathcal{K}$ and $\mathcal{K}'$ be finite states and let ϕ be a formula. A *contraction operator*, $\dot{-}$, is any operator satisfying the following postulates:

$(\mathcal{K}\dot{-}1)$	$\mathcal{K}\dot{-}\phi$ is a finite state	(*Closure*)
$(\mathcal{K}\dot{-}2)$	$\mathcal{K}\dot{-}\phi \trianglelefteq \mathcal{K}$	(*Inclusion*)
$(\mathcal{K}\dot{-}3)$	If $\phi \not\trianglelefteq \mathcal{K}$, then $\mathcal{K}\dot{-}\phi = \mathcal{K}$	(*Vacuity*)
$(\mathcal{K}\dot{-}4)$	$\phi \not\trianglelefteq \mathcal{K}\dot{-}\phi$	(*Success*)
$(\mathcal{K}\dot{-}M)$	If $\mathcal{K} \trianglelefteq \mathcal{K}'$, then $\mathcal{K}\dot{-}\phi \trianglelefteq \mathcal{K}'\dot{-}\phi$	(*Monotonicity*).

The last property of this list is not among the AGM-postulates.[8] Gärdenfors discusses the property under the same name $(\mathcal{K}\dot{-}M)$ and argues against it, because it is, in the context of belief sets, equivalent to its counterpart $(\mathcal{K} \times M)$, which he previously showed to be unsound [Gärdenfors, 1988, p. 59–60]. In the present setting, $(\mathcal{K}\dot{-}M)$ is satisfied, but $(\mathcal{K} \times M)$ not. [Pais and Jackson, 1992] introduces a similar but weaker postulate *Partial Monotonicity*.

[7]In the present setting, *Recovery* would amount to the following: $\mathcal{K} \trianglelefteq (\mathcal{K}\dot{-}\phi) + \phi$. For a counterexample: let $\mathcal{K}$ be $[\![p \wedge q]\!]$, and let ϕ be $p \vee q$. As is well-known, *Recovery* is, in case belief revision is studied for epistemological reasons, the least plausible of Gärdenfors's postulates for contraction. [Hansson, 1999a] presents the central counter-examples which hinge on the justificatory structure of the beliefs involved. As the other translations of Gärdenfors's postulates are satisfied by our contraction function, our contraction function is, following [Makinson, 1987], a finite state based *withdrawal function*.

[8][Alchourrón, Gärdenfors, and Makinson, 1985].

All traditional contraction functions, such as partial meet contraction[9], safe contraction[10], and contraction based on epistemic entrenchment[11], depend on an extra-logical element. The chief argument for adopting an extra-logical element, such as a selection function or an ordering of the formulas, is the fact that the only contraction function defined by logical means alone, that is, *full meet contraction* [Alchourrón, Gärdenfors, and Makinson, 1985, p. 512], gives rise to a trivial operation.[12] As we have seen, in the present context, contraction has been defined with logical means alone. Since our contraction function does not lead to the undesirable results that full meet contraction had within the usual classical setting of belief change, AGM's arguments against defining contraction without an extra-logical element are out of place in our approach.

Standard AGM-contractions allow for different non-equivalent contraction functions satisfying the postulates ($\mathcal{K}\dot{-}1$) through ($\mathcal{K}\dot{-}8$). ($\mathcal{K}\dot{-}M$), however, guarantees uniqueness of the function, if ($\mathcal{K}\dot{-}1$) through ($\mathcal{K}\dot{-}4$) are present:

4.3.13. Theorem (Uniqueness). *Let $\dot{-}$ and $\ddot{-}$ be contraction functions, satisfying ($\mathcal{K}\dot{-}1$) through ($\mathcal{K}\dot{-}4$) and ($\mathcal{K}\dot{-}M$). Then for all formulas ϕ it holds that $\mathcal{K}\dot{-}\phi = \mathcal{K}\ddot{-}\phi$.*

Proof. On the basis of ($\mathcal{K}\dot{-}4$), it holds that $\phi \not\sqsubseteq \mathcal{K}\dot{-}\phi$ and $\phi \not\sqsubseteq \mathcal{K}\ddot{-}\phi$. By ($\mathcal{K}\dot{-}1$), both $\mathcal{K}\dot{-}\phi$ and $\mathcal{K}\ddot{-}\phi$ are finite states. Hence, by ($\mathcal{K}\dot{-}3$), we obtain $(\mathcal{K}\dot{-}\phi)\ddot{-}\phi = \mathcal{K}\dot{-}\phi$ and $(\mathcal{K}\ddot{-}\phi)\dot{-}\phi = \mathcal{K}\ddot{-}\phi$. Moreover, by ($\mathcal{K}\dot{-}2$), we have $\mathcal{K}\dot{-}\phi \trianglelefteq \mathcal{K}$ and $\mathcal{K}\ddot{-}\phi \trianglelefteq \mathcal{K}$. From this, using ($\mathcal{K}\dot{-}M$), it follows that $(\mathcal{K}\dot{-}\phi)\ddot{-}\phi \trianglelefteq \mathcal{K}\ddot{-}\phi$ and $(\mathcal{K}\ddot{-}\phi)\dot{-}\phi \trianglelefteq \mathcal{K}\dot{-}\phi$. Using the equalities established previously, it must be that $\mathcal{K}\dot{-}\phi \trianglelefteq \mathcal{K}\ddot{-}\phi$ and $\mathcal{K}\ddot{-}\phi \trianglelefteq \mathcal{K}\dot{-}\phi$. From Lemma 4.2.22 it follows that $\mathcal{K}\dot{-}\phi = \mathcal{K}\ddot{-}\phi$. □

4.3.14. Theorem (Representation Theorem for Contraction). *Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then $\mathcal{K}\dot{-}\phi$ satisfies ($\mathcal{K}\dot{-}1$) through ($\mathcal{K}\dot{-}4$) and ($\mathcal{K}\dot{-}M$) iff $\mathcal{K}\dot{-}\phi = \mathcal{K} - \phi$.*

Proof. Because of Theorem 4.3.13, it will suffice to show that $\mathcal{K} - \phi$ satisfies ($\mathcal{K}\dot{-}1$) through ($\mathcal{K}\dot{-}4$) and ($\mathcal{K}\dot{-}M$). The first four postulates follow directly from the definition of $\mathcal{K} - \phi$. To prove ($\mathcal{K}\dot{-}M$), assume that $\mathcal{K} \trianglelefteq \mathcal{K}'$. Suppose that $\psi \sqsubseteq \mathcal{K} - \phi$. Then there is a v in $\mathcal{K} - \phi$ with $v \models \psi$. Hence, there is a w in $\mathcal{K}$ with $v \subseteq w$ and $v \models \psi$ and $v \not\models \phi$.

[9] [Alchourrón, Gärdenfors, and Makinson, 1985].

[10] [Alchourrón and Makinson, 1985].

[11] [Gärdenfors and Makinson, 1988].

[12] If K is a theory such that $\phi \in K$, then the *full meet contraction* of ϕ from K, denoted by $K \sim \phi$, is defined to be $\cap(K \perp \phi)$, where $K \perp \phi$ is the set of all maximal subsets K' of K such that $K' \nvdash \phi$. Observation 2.1 of [Alchourrón and Makinson, 1982] shows what goes wrong: if $\phi \in K$, then $K \sim \phi = K \cap Cn(\neg\phi)$, which is far too small, since the contraction skips all sentences from K that are not consequences of $\neg\phi$. For instance, let K be $Cn(p \wedge q)$ and let ϕ be q. Then $p \notin K \sim q$. (Using the machinery of the present chapter, let $\mathcal{K}$ be $[\![p \wedge q]\!]$ and let ϕ be q. Then $\mathcal{K} \oslash [\![q]\!] = [\![p]\!]$.) The situation for *full meet Levi contractions* is even worse. See Theorem 5.3 of [Hansson and Olsson, 1995].

Hence, $\bigwedge w \sqsubseteq \mathcal{K}$. Therefore, by assumption, $\bigwedge w \sqsubseteq \mathcal{K}'$. Hence, there is a w' in $\mathcal{K}'$ with $w \subseteq w'$. Summarizing, there is a w' in $\mathcal{K}'$ with $v \subseteq w'$ and $v \not\models \phi$ and $v \models \psi$. Hence, $v \in \{v \subseteq w' : w' \in \mathcal{K}' \textit{ and } v \not\models \phi\}$. As the maximalization of this set only skips v in favour of a superset of v, there will be an element in $\mathcal{K}' - \phi$ that validates ψ. Hence, $\psi \sqsubseteq \mathcal{K}' - \phi$. Therefore, if $\mathcal{K} \trianglelefteq \mathcal{K}'$, then $\mathcal{K} - \phi \trianglelefteq \mathcal{K}' - \phi$. □

Iterated contraction is related to disjunction:

4.3.15. Theorem (Iterated Contraction). *Let $\mathcal{K}$ be a finite state and let ϕ and ψ be formulas. Then*

$$(\mathcal{K} - \phi) - \psi \quad = \quad \mathcal{K} - (\phi \vee \psi).$$

Proof. Since, by Definition 4.3.10 and Theorem 4.3.11, it holds that $\mathcal{K} - (\phi \vee \psi) = (\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)$, it suffices to prove (a) $(\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi) \trianglelefteq (\mathcal{K} - \phi) - \psi$ and (b) $(\mathcal{K} - \phi) - \psi \trianglelefteq (\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)$. The identity is then established by Lemma 4.2.22.

(a) By Lemma 4.2.26, it must be that both (i) $(\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi) \trianglelefteq \mathcal{K} - \phi$ and (ii) $(\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi) \trianglelefteq \mathcal{K} - \psi$. Because of the Representation Theorem and ($\mathcal{K}\dot{-}4$) it holds that $\psi \not\sqsubseteq \mathcal{K} - \psi$, it follows from (ii) that $\psi \not\sqsubseteq (\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)$. Hence, by ($\mathcal{K}\dot{-}3$), it holds that (iii) $((\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)) - \psi = (\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)$. By (i) and ($\mathcal{K}\dot{-}M$), it must be that $((\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)) - \psi \trianglelefteq (\mathcal{K} - \phi) - \psi$, which, together with (iii), gives us (a).

(b) By the Representation Theorem and ($\mathcal{K}\dot{-}2$) it must be that $(\mathcal{K} - \phi) - \psi \trianglelefteq \mathcal{K} - \phi$ and $\mathcal{K} - \phi \trianglelefteq \mathcal{K}$. Applying ($\mathcal{K} - M$) to the latter gives us $(\mathcal{K} - \phi) - \psi \trianglelefteq \mathcal{K} - \psi$. By Lemma 4.2.27 and Lemma 4.2.25, it must be that $(\mathcal{K} - \phi) - \psi \trianglelefteq (\mathcal{K} - \phi) \sqcap (\mathcal{K} - \psi)$. □

Due to the commutativity of disjunction, it immediately follows that the order in which formulas are contracted from a finite state has no influence on the result:

4.3.16. Corollary. *Let $\mathcal{K}$ be a finite state and let ϕ and ψ be formulas. Then*

$$(\mathcal{K} - \phi) - \psi \quad = \quad (\mathcal{K} - \psi) - \phi.$$

4.3.3 Revision

Finally, revision of a finite state $\mathcal{K}$ with a formula ϕ can be thought of as changing a finite state $\mathcal{K}$ with the claim that there is only evidence *for* (and not against) ϕ. This can be modelled in two ways, the second of which is an adaption of a proposal in [Hansson, 1993] to the present context.

4.3.17. Definition. [Levi Identity] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then the *revision of* $\mathcal{K}$ *with* ϕ, denoted by $\mathcal{K} \times \phi$, is defined to be

$$\mathcal{K} \times \phi \quad = \quad (\mathcal{K} - \neg\phi) + \phi.$$

4.3.18. DEFINITION. [Reversed Levi Identity] Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then the *reversed revision of* $\mathcal{K}$ *with* ϕ, denoted by $\mathcal{K} \times_r \phi$, is defined to be

$$\mathcal{K} \times_r \phi \quad = \quad (\mathcal{K} + \phi) - \neg\phi.$$

Using the Levi Identity always leads to a finite state $\mathcal{K} \times \phi$ that *supports* ϕ. Hence, $\partial(\mathcal{K} \times \phi) \models \phi$. Note that $\neg\phi$ might be within the information span of $\mathcal{K} \times \phi$. (Consider, for instance, $\mathcal{K} \times (p \wedge \neg p)$.) Conversely, applying the Reversed Levi Identity always results in a finite state $\mathcal{K} \times_r \phi$ that does not have $\neg\phi$ within its information span. Hence, it holds that $\neg\phi \not\trianglelefteq \mathcal{K} \times_r \phi$. It is possible, though, that ϕ is not supported by $\mathcal{K} \times_r \phi$ either. (Consider, once again, $\mathcal{K} \times_r (p \wedge \neg p)$.) Special representation theorems for these revision functions have not been thoroughly studied yet.

4.4 Extractors: Operations of Belief Retrieval

The formal apparatus developed so far enables us to attach to each formula, consistent or inconsistent it may be, a finite state. Which set of beliefs can we plausiby extract from this state, under the constraint that this set of extracted beliefs has to be *consistent*? Several alternative consequence relations, constructed along different lines, will be adequate to this task.[13] Hence, our investigations into these consequence relations will be of a tentative sort, since there are no *prima facie* reasons in favor of one of the four alternatives which we shall discuss.

Here, we shall first distinguish two functions on finite states that make inconsistent valuations consistent and define extractors in terms of these functions. The first function, the consistency forcer f, just ignores literals which appear contradictorily in a contradictory valuation. The second one, the consistency forcer g, takes maximal consistent subvaluations of a contradictory valuation, thereby retaining literals which appear contradictorily in it.

Next, we consider a selection function that serves as a criterion on the basis of which we may neglect some valuations. Before applying one of our two consistency-forcing functions, we may, but must not, use this selection function. Hence, four different consequence relations are generated. Given the present results, a principled choice between these consequence relations can not be made. Instead, we shall list some arguments in favor of each of them.

[13] A systematic overview of consequence relations devised for reasoning from inconsistent unprioritized *knowledge bases* is presented in [Benferhat, Dubois, and Prade, 1997]. Semantics for their consequence relations have been provided recently (See [Batens, manuscript]). Benferhat, Dubois and Prade also investigate into the 'syntax-sensitivity' of their consequence relations, distinguishing four types of syntax-sensitivity. Contrary to the consequence relations discussed by them, the consequence relations to be defined in the present chapter are insensitive to *all* four types.

4.4.1 Making $\mathcal{K}^{\perp}$ Consistent

Informally, a simple *preference relation* on partial valuations motivates the definition of these consequence relations. The minimal and consistent partial valuations satisfying a certain formula are preferred over the minimal and inconsistent valuations satisfying the same formula, if any. If there are such consistent valuations in the finite state under discussion, we do not want any surprising logic. (Here, the logic ruling the consistent case amounts to Kleene's strong three-valued logic.) Therefore, if the finite state which serves as the starting-point has a consistent valuation among its elements, all four extractors coincide.

On the other hand, if there are no minimal and consistent valuations satisfying the formula in question, then our reasoning has to be carried out on the basis of the minimal and inconsistent valuations satisfying that formula. These inconsistent valuations must be processed in some way or other, since all of them satisfy some contradiction. We shall describe two related ways of processing the inconsistent valuations guaranteeing that the beliefs that are based on them are consistent.

[Restall and Slaney, 1995] uses a *vocabulary* to extract consistent subsets from a set of formulas K closed under fde. A vocabulary V is a set of propositional variables, and the restriction of K with V equals the set of formulas in K built from the propositional variables in V. If one chooses only propositional variables p in V which are consistent in K (that is, $p \wedge \neg p \notin K$), then the restriction of K with V is consistent.

In the present setting, Restall and Slaney's proposal amounts to skipping all contradictory literals from all valuations in $\mathcal{K}^{\perp}$.[14] Restricting all valuations w in $\mathcal{K}^{\perp}$ to its consistent literals is just taking the co-valuations $\overline{w}$ of all w in $\mathcal{K}$. So, if $\mathcal{K}^{\top} = \emptyset$, in order to guarantee that the reliable conclusions based on the given inconsistent information are classically consistent, their proposal tells us to work with the *co-valuations* $\overline{w}$ of valuations w in $\mathcal{K}^{\perp}$. In this way, the joint consistency of the 'beliefs' licensed by a finite state is secured, also in case when that state only contains inconsistent valuations.

A similar (as a matter of fact slightly stronger) extractor is defined thus: in case $\mathcal{K}^{\top} = \emptyset$, we might consider the maximal consistent subvaluations of the inconsistent valuations in $\mathcal{K}^{\perp}$. For instance, consider $p \wedge \neg p \wedge q$. Then its minimal valuations are given by the finite state $\{\{p, \neg p, q\}\}$, whose single valuation is inconsistent. Taking its maximal consistent subvaluations would give us $\{\{p, q\}, \{\neg p, q\}\}$. Then $\partial(\{\{p, q\}, \{\neg p, q\}\}) \models p \vee \neg p$, although the first consequence relation does not give us $p \vee \neg p$ as a conclusion, since $\partial(\{\{q\}\}) \not\models p \vee \neg p$.

Let us make these ideas precise with a couple of definitions:

[14] For $\mathcal{K}^{\top}$ and $\mathcal{K}^{\perp}$, see Definition 4.2.7.

4.4.1. Definition. [Consistency Forcers] Let $\mathcal{K}$ be a finite state. Then the *co-valuations of* $\mathcal{K}$, denoted by $f(\mathcal{K})$, and the *maximally consistent subvaluations of* $\mathcal{K}$, denoted by $g(\mathcal{K})$, are defined to be

$$\begin{array}{rcl} f(\mathcal{K}) & = & \{\overline{w} : w \in \mathcal{K}\} \\ g(\mathcal{K}) & = & \mathsf{max}(\{v : v \subseteq w \textit{ and } w \in \mathcal{K} \textit{ and } v \textit{ is consistent }\}). \end{array}$$

4.4.2. Definition. Let $\mathcal{K}$ be a finite state and let $x \in \{f, g\}$. Then

$$\mathcal{K}_x \quad = \quad \begin{cases} x(\mathcal{K}^{\perp}), & \text{if } \mathcal{K}^{\top} = \emptyset \\ \mathcal{K}^{\top}, & \text{otherwise.} \end{cases}$$

4.4.3. Definition. [Belief Extractors] Let $\mathcal{K}$ be a finite state, let ϕ be a formula, and let $x \in \{f, g\}$. If $\mathcal{K}_x = \{\emptyset\}$, then $\mathcal{K} \not\approx_x \phi$ for all $\phi \in \mathcal{F}$. Otherwise,

$$\mathcal{K} \mid\!\approx_x \phi \quad \text{iff} \quad \partial(\mathcal{K}_x) \models \phi.$$

As an example, let $\mathcal{K}$ be $[\![p \wedge \neg(p \wedge q) \wedge r]\!]$, the first information base discussed in the introductory Cervantes-example. Obviously, $\mathcal{K} \mid\!\approx_f p \wedge \neg(p \wedge q) \wedge r$. Expanding $\mathcal{K}$ with q, we obtain $\mathcal{K} + q \mid\!\approx_f (p \vee q) \wedge r$, but $\mathcal{K} + q \not\approx_f p$ and $\mathcal{K} + q \not\approx_f q$. If we now contract this last state with p, then $(\mathcal{K} + q) - p \mid\!\approx_f q \wedge \neg(p \wedge q) \wedge r$. The reader can easily check that all these statements hold for g-based extraction as well. Unlike f-based extraction, it holds that $\mathcal{K} + q \mid\!\approx_g p \vee \neg p$ and $\mathcal{K} + q \mid\!\approx_g q \vee \neg q$.

There are two arguments in favor of f-based extraction. First, it is the weakest extractor of the four alternative extractors we shall discuss (Theorem 4.4.10). Secondly, it conforms to six out of seven conditions for Preferential Reasoning as described by Kraus, Lehmann and Magidor (Theorem 4.4.4). It is far from obvious, however, that these conditions have normative force within the present context.

Likewise, there are two arguments in favor of g-based extraction. First, unlike the f-based extraction, g-based contraction gives literals appearing inconsistently in the relevant valuations and literals that do not appear in the relevant valuations a different treatment. Secondly, the consistency forcer g is a special case of contraction (Theorem 4.4.6).

Some Properties of f-Based and g-Based Extraction

Let us check some properties of our extractors, using the categorization of nonmonotonous systems in [Kraus, Lehmann, and Magidor, 1990]. As noted, it is not clear whether Kraus, Lehmann and Magidor's conditions have normative force within the present context. Theorem 4.4.4 does enable a comparison of some properties of f-based extraction with familiar conditions on nonmonotonic reasoning. Representation theorems for f-based extraction or g-based extraction have not been found yet.

We shall show that the f-based extractor satisfies four out of five of Kraus, Lehmann and Magidor's conditions for Cumulative Reasoning,[15] whereas the condition that does not hold, that is, *Reflexivity*, can be modified into *Conditional Reflexivity*. Since our extractors do not satisfy *Reflexivity*, *And* loses its status as a derived rule and becomes an independent one. Additionally, the rule *Or* holds for f-based extraction.[16] Hence, the properties of f-based extraction are, except for the change of *Reflexivity* into *Conditional Reflexivity*, similar to Kraus, Lehmann and Magidor's Preferential Reasoning.

4.4.4. Theorem. *Let ϕ, ψ, and χ be formulas and let $x \in \{f, g\}$. Then*

(i) $$\frac{[\![\phi]\!] \approx_x \psi}{[\![\psi]\!] \approx_x \psi}$$ *(Conditional Reflexivity)*

(ii) $$\frac{\phi \models \psi \qquad \psi \models \phi \qquad [\![\phi]\!] \approx_x \chi}{[\![\psi]\!] \approx_x \chi}$$ *(Left Logical Equivalence)*

(iii) $$\frac{[\![\phi]\!] \approx_x \psi \qquad \psi \models \chi}{[\![\phi]\!] \approx_x \chi}$$ *(Right Weakening)*

(iv) $$\frac{[\![\phi \wedge \psi]\!] \approx_f \chi \qquad [\![\phi]\!] \approx_f \psi}{[\![\phi]\!] \approx_f \chi}$$ *(Cut)*

(v) $$\frac{[\![\phi]\!] \approx_f \psi \qquad [\![\phi]\!] \approx_f \chi}{[\![\phi \wedge \psi]\!] \approx_f \chi}$$ *(Cautious Monotonicity)*

(vi) $$\frac{[\![\phi]\!] \approx_x \psi \qquad [\![\phi]\!] \approx_x \chi}{[\![\phi]\!] \approx_x \psi \wedge \chi}$$ *(And)*

(vii) $$\frac{[\![\phi]\!] \approx_f \chi \qquad [\![\psi]\!] \approx_f \chi}{[\![\phi \vee \psi]\!] \approx_f \chi}$$ *(Or)*

[15] Cumulative Reasoning consists of the inference rules *Reflexivity*, *Left Logical Equivalence*, *Right Weakening*, *Cut*, and *Cautious Monotonicity*. Preferential Reasoning consists of the inference rules for Cumulative Reasoning and the inference rule *Or*.

[16] Within g-based extraction, *Cut*, *Cautious Monotonicity* and *Or* are not valid. (Since the three counterexamples hinge on formulas which are tautologies in classical logic, it might be that these nonmonotonic inference relations do hold for the classical version of g-based extraction.)

To refute *Cut*, take $\phi = (p \wedge \neg p) \vee (q \wedge \neg q) \vee (p \wedge r \wedge \neg r) \vee (\neg p \wedge r \wedge \neg r) \vee (q \wedge s \wedge \neg s) \vee (\neg q \wedge s \wedge \neg s)$, $\psi = ((p \vee \neg p) \wedge (r \vee \neg r)) \vee ((q \vee \neg q) \wedge (s \vee \neg s))$, and $\chi = (p \vee \neg p) \wedge (q \vee \neg q)$. Then $[\![\phi \wedge \psi]\!] \approx_g \chi$ and $[\![\phi]\!] \approx_g \psi$ and $[\![\phi]\!] \not\approx_g \chi$.

To refute *Cautious Monotonicity*, take $\phi = (p \wedge \neg p) \vee (p \wedge q \wedge \neg q) \vee (\neg p \wedge q \wedge \neg q)$, $\psi = q \vee \neg q \vee r$, and $\chi = q \vee \neg q$. Then $[\![\phi]\!] \approx_g \psi$ and $[\![\phi]\!] \approx_g \chi$ and $[\![\phi \wedge \psi]\!] \not\approx_g \chi$.

To refute *Or*, take $\phi = (p \wedge \neg p) \vee (p \wedge q \wedge \neg q)$, $\psi = q \wedge \neg q$, and $\chi = \neg p \vee q \vee \neg q$. Then $[\![\phi]\!] \approx_g \chi$ and $[\![\psi]\!] \approx_g \chi$ and $[\![\phi \vee \psi]\!] \not\approx_g \chi$.

Proof. (i) Suppose that $[\![\phi]\!] \approx_x \psi$. Then ψ is consistent. Hence, $[\![\psi]\!]^\top \neq \emptyset$. Then, $[\![\psi]\!]_x = [\![\psi]\!]^\top$. Hence, $\partial([\![\psi]\!]_x) = \partial([\![\psi]\!]^\top)$. Since $w \models \psi$ for all w in $[\![\psi]\!]^\top$, it follows, by Lemma 4.2.19, that $\partial([\![\psi]\!]_x) \models \psi$. Therefore, $[\![\psi]\!] \approx_x \psi$.

(ii) Immediately from Lemma 4.2.8.

(iii) Suppose that $[\![\phi]\!] \approx_x \psi$ and $\psi \models \chi$. Then, by definition, $\partial([\![\phi]\!]_x) \models \psi$. Since '$\models$' is transitive, it follows that $\partial([\![\phi]\!]_x) \models \chi$. Therefore, $[\![\phi]\!] \approx_x \chi$.

(iv) Suppose that $[\![\phi \wedge \psi]\!] \approx_f \chi$ and $[\![\phi]\!] \approx_f \psi$.

Suppose that $[\![\phi \wedge \psi]\!]^\top \neq \emptyset$. Then $[\![\phi]\!]^\top \neq \emptyset$. Take a w in $[\![\phi]\!]^\top$. Then $w \models \phi$ and, by the global supposition, $w \models \psi$, as $\partial([\![\phi]\!]^\top) \models \psi$. Hence, $w \models \phi \wedge \psi$. Then, there is a v in $[\![\phi \wedge \psi]\!]^\top$ with $v \subseteq w$. Since by the global assumption it holds that $v \models \chi$, it must be that $w \models \chi$. Hence, $\partial([\![\phi]\!]^\top) \models \chi$. Therefore, $[\![\phi]\!] \approx_f \chi$.

Suppose that $[\![\phi \wedge \psi]\!]^\top = \emptyset$. Suppose that $[\![\phi]\!]^\top \neq \emptyset$. Take a w in $[\![\phi]\!]^\top$. Then w is consistent, $w \models \phi$, and, by the same argument as in the previous case, $w \models \psi$. Hence, $\phi \wedge \psi$ is consistent. Therefore, $[\![\phi \wedge \psi]\!]^\top \neq \emptyset$: a contradiction. Hence, $[\![\phi]\!]^\top = \emptyset$. It is proved that under the given circumstances $[\![\phi]\!] = [\![\phi \wedge \psi]\!]$: Suppose that $w \in [\![\phi]\!]$. Then $w \models \phi$ and, by the global assumption, $\overline{w} \models \psi$. Hence, $w \models \phi \wedge \psi$. Suppose that $v \subset w$. Then $v \not\models \phi$. Hence, $v \not\models \phi \wedge \psi$. Therefore, $w \in [\![\phi \wedge \psi]\!]$. Suppose, to prove the other inclusion, $w \in [\![\phi \wedge \psi]\!]$. Then $w \models \phi$. Suppose that $v \subset w$. Then $v \not\models \phi \wedge \psi$. Suppose that $v \models \phi$. Then it must be that $v \not\models \psi$. From $v \models \phi$, it follows that there is a v' with $v' \subseteq v$ and $v' \in [\![\phi]\!]$. By the global assumption, it must be that $\overline{v'} \models \psi$. Hence, $v' \models \psi$ and $v \models \psi$: a contradiction. Therefore, $v \not\models \phi$. Therefore, $w \in [\![\phi]\!]$. Finally, $[\![\phi]\!] \approx_f \chi$.

(v) Suppose that $[\![\phi]\!] \approx_f \psi$ and $[\![\phi]\!] \approx_f \chi$.

Suppose that $[\![\phi \wedge \psi]\!]^\top \neq \emptyset$. Then $[\![\phi]\!]^\top \neq \emptyset$. Take a w in $[\![\phi \wedge \psi]\!]^\top$. Then $w \models \phi$. Then there is a v in $[\![\phi]\!]^\top$ with $v \subseteq w$. By the global assumption, it must be that $v \models \chi$. Then, $w \models \chi$. Hence, $\partial([\![\phi \wedge \psi]\!]^\top) \models \chi$. Therefore, $[\![\phi \wedge \psi]\!] \approx_f \chi$.

Suppose that $[\![\phi \wedge \psi]\!]^\top = \emptyset$. By the same train of arguments as in the proof of (iv), it must be that $[\![\phi \wedge \psi]\!] = [\![\phi]\!]$. Therefore, $[\![\phi \wedge \psi]\!] \approx_f \chi$.

(vi) Suppose that $[\![\phi]\!] \approx_x \psi$ and $[\![\phi]\!] \approx_x \chi$. Then, by definition, $\partial([\![\phi]\!]_x) \models \psi$ and $\partial([\![\phi]\!]_x) \models \chi$. Hence, $\partial([\![\phi]\!]_x) \models \psi \wedge \chi$. Therefore, $[\![\phi]\!] \approx_x \psi \wedge \chi$.

(vii) Suppose that $[\![\phi]\!] \approx_f \chi$ and $[\![\psi]\!] \approx_f \chi$.

Suppose that $[\![\phi \vee \psi]\!]^\top \neq \emptyset$. Take a w in $[\![\phi \vee \psi]\!]^\top$. Then, by Lemma 4.2.11(iv), it must be that $w \in [\![\phi]\!]^\top$ or $w \in [\![\psi]\!]^\top$. In both cases, by the global supposition, it must be that $w \models \chi$. Hence, $\partial([\![\phi \vee \psi]\!]^\top) \models \chi$. Therefore, $[\![\phi \vee \psi]\!] \approx_f \chi$.

Suppose that $[\![\phi \vee \psi]\!]^\top = \emptyset$. Then $[\![\phi]\!]^\top = \emptyset$ and $[\![\psi]\!]^\top = \emptyset$. Take a w in $[\![\phi \vee \psi]\!]^\perp$. Then, by Lemma 4.2.11(iv), it must be that $w \in [\![\phi]\!]^\perp$ or $w \in [\![\psi]\!]^\perp$. In both cases, by the global supposition, it must be that $\overline{w} \models \chi$. Hence, $\partial(f([\![\phi \vee \psi]\!]^\perp)) \models \chi$. Therefore, $[\![\phi \vee \psi]\!] \approx_f \chi$. □

To establish a connection between the consistency forcer g and the contraction function of the previous section, we introduce the notion of a finite state's *inconsistency determiner*, a formula indicating which propositional variables occur inconsistently in one of the valuations of the finite state. Note that different finite states can have identical inconsistency determiners.

4.4.5. Definition. [Inconsistency Determiner] Let $\mathcal{K}$ be a finite state, such that $\mathcal{K}^{\perp} \neq \emptyset$. Then the *inconsistency determiner of* $\mathcal{K}$, denoted by $\partial_{\perp}(\mathcal{K})$, is defined to be

$$\partial_{\perp}(\mathcal{K}) \quad = \quad \bigvee\{p \wedge \neg p : p \in \mathrm{PV} \text{ and } p \wedge \neg p \sqsubseteq \mathcal{K}\}.$$

The inconsistency determiner of a finite state $\mathcal{K}$ can be used to define a special contraction operation that makes $\mathcal{K}$ consistent. This operation resembles Hansson's operator of *consolidation* [Hansson, 1999b, p. 420].

Contracting a finite state $\mathcal{K}$ that contains at least one contradictory valuation with its inconsistency determiner $\partial_{\perp}(\mathcal{K})$ amounts to taking its maximal consistent valuations $g(\mathcal{K})$. In this way, a strong connection has been established between information contraction and g-based extraction.

4.4.6. Theorem. *Let $\mathcal{K}$ be a finite state, such that $\mathcal{K}^{\perp} \neq \emptyset$. Then*

$$\mathcal{K} - \partial_{\perp}(\mathcal{K}) \quad = \quad g(\mathcal{K}).$$

Proof. Let $\{p_1, \ldots, p_n\}$ be the set of all atoms p for which there is a valuation w in $\mathcal{K}$ with $\{p, \neg p\} \subseteq w$. This set is non-empty by hypothesis, and it is finite because $\mathcal{K}$ is a finite state. Hence, $\mathcal{K} - \partial_{\perp}(\mathcal{K})$

$$\begin{aligned}
&= \mathcal{K} - ((p_1 \wedge \neg p_1) \vee \cdots \vee (p_n \wedge \neg p_n)) \\
&= \mathsf{max}(\{v \subseteq w : w \in \mathcal{K} \textit{ and } v \not\models (p_1 \wedge \neg p_1) \vee \cdots \vee (p_n \wedge \neg p_n)\}) \\
&= \mathsf{max}(\{v \subseteq w : w \in \mathcal{K} \textit{ and } v \not\models p_1 \wedge \neg p_1 \textit{ and } \ldots \textit{ and } v \not\models p_n \wedge \neg p_n\}) \\
&= \mathsf{max}(\{v \subseteq w : w \in \mathcal{K} \textit{ and } v \textit{ is consistent}\}) \\
&= g(\mathcal{K}).
\end{aligned}$$

□

4.4.2 Selecting $\mathcal{K}^{\perp}$'s Most Consistent Valuations

Consider the set of formulas $\{p, \neg p, \neg q, p \vee r, q \vee s\}$.[17] The finite state $\mathcal{K} = \{\{p, \neg p, q, \neg q\}, \{p, \neg p, \neg q, s\}\}$ contains the minimal valuations of the conjunction of all formulas in this set. It is clear from the start that p must behave inconsistently. With regard to q, on the other hand, there are two possibilities: q can behave inconsistently as well as consistently. The extractors f and g treat these possibilities on equal terms, that is, no preference is given to 'only p is inconsistent' above 'both p and q are inconsistent'. Hence, $\mathcal{K} \not\approx_f s$ and $\mathcal{K} \not\approx_g s$.

[17] The example has been taken from [Batens, forthcoming]. The choice function to be defined in the present section leads to an *inconsistency-adaptive logic* in Batens's sense. These are logics between a monotonic upper limit logic and a paraconsistent monotonic lower limit logic interpreting formulas as consistently as possible. In the terminology of the present chapter, we may choose as an upper limit logic Kleene's strong three-valued logic and as a lower limit logic fde. An introduction into inconsistency-adaptive logics can be found in the series of papers [Batens, 1980], [Batens, 1989], and [Batens, 1986].

In defining f-based and g-based extraction, a choice function, albeit a crude one, was already put in use: if $\mathcal{K}$ contains consistent valuations, stick to $\mathcal{K}^{\top}$, otherwise use $\mathcal{K}^{\perp}$. There are no pressing reasons to demand that *all* of $\mathcal{K}^{\perp}$'s inconsistent valuations must be given equal consideration. We could ignore those valuations which require more literals to behave inconsistently than other inconsistent valuations, thereby choosing the *most consistent* valuations in $\mathcal{K}$. In the example above, this amounts to restricting $\mathcal{K}$ to $\{\{p, \neg p, \neg q, s\}\}$, since all of its inconsistent literals are contained in the valuation $\{p, \neg p, q, \neg q\}$, which contains an extra inconsistently behaving literal as well. Denoting the intended restriction of $\mathcal{K}$ by $\mathcal{K}^\star$, it now holds that $\mathcal{K}^\star \models\!\approx_f s$ and $\mathcal{K}^\star \models\!\approx_g s$.

To make this idea precise, we need to define a choice function selecting the most consistent valuations from a set of valuations. This can be done as follows:

4.4.7. Definition. Let $\mathcal{K}$ be a finite state. The set of *$\mathcal{K}$'s most consistent valuations*, denoted by $\mathcal{K}^\star$, is defined to be:

$$\mathcal{K}^\star \quad = \quad \{w \in \mathcal{K} : w - \overline{w} \in \mathsf{min}(\{w - \overline{w} : w \in \mathcal{K}\})\}.$$

Let us first show that this definition is well behaved:

4.4.8. Lemma. *Let $\mathcal{K}$ be a finite state. Then*

(i) *If $\mathcal{K}^{\top} = \emptyset$, then $\mathcal{K}^\star \subseteq \mathcal{K}^{\perp}$.*
(ii) *If $\mathcal{K}^{\top} \neq \emptyset$, then $\mathcal{K}^\star = \mathcal{K}^{\top}$.*

Proof. (i) Suppose that $\mathcal{K}^{\top} = \emptyset$. Then $\mathcal{K} = \mathcal{K}^{\perp}$, since $\mathcal{K}$ contains only inconsistent valuations. Hence, $\mathcal{K}^\star \subseteq \mathcal{K}^{\perp}$.

(ii) Suppose that $\mathcal{K}^{\top} \neq \emptyset$. Take a u in $\mathcal{K}^{\top}$. Then $u \in \mathcal{K}$ and $u - \overline{u} = \emptyset$. Hence, $\emptyset \in \{w - \overline{w} : w \in \mathcal{K}\}$. Suppose that $v \in \mathcal{K}^\star$. Then it must be that $v - \overline{v} = \emptyset$, since otherwise $\emptyset \subset v - \overline{v}$ and v would be skipped because of the presence of u. Therefore, $v \in \mathcal{K}^{\top}$. To prove the other inclusion, suppose that $v \in \mathcal{K}^{\top}$. Then $v \in \mathcal{K}$ and $v - \overline{v} = \emptyset$. Hence, since there can be no w in $\mathcal{K}$ with $w - \overline{w} \subset v - \overline{v}$, it must be that $v \in \mathcal{K}^\star$. □

Now, we can combine the f-based and g-based extractors with our function selecting the most consistent valuations.

4.4.9. Definition. [Selective Belief Extractors] Let $\mathcal{K}$ be a finite state, let ϕ be a formula, and let $x \in \{f, g\}$. Then

$$\mathcal{K} \models\!\approx^\star_x \phi \quad \text{iff} \quad \partial((\mathcal{K}^\star)_x) \models \phi.$$

Which relations hold between our four extractors? Let us start with noting that $\mathcal{K} \models\!\approx_g \phi$ does not imply $\mathcal{K} \models\!\approx^\star_g \phi$.[18] The following relations hold:

[18]Consider the finite state $\mathcal{K} = \{\{p, \neg p, q\}, \{q, \neg q\}, \{p, \neg p, \neg q, r, \neg r, s\}\}$. Then $\partial(g(\mathcal{K}^\star)) \not\models \partial(g(\mathcal{K}))$, even if '$\models$' in Definitions 4.4.3 and 4.4.9 is taken to denote *classical* validity. It is easy to see, though, that Theorem 4.4.10 still holds under this modification. As an immediate corollary, it is established that $\mathcal{K} \models\!\approx_f \phi$ and $\mathcal{K} \models\!\approx_g \phi$ are non-equivalent even under classical propositional logic.

4.4.10. Theorem (Relative Strength). *Let $\mathcal{K}$ be a finite state and let ϕ be a formula. Then*

(i) *If $\mathcal{K} \approx_f \phi$, then $\mathcal{K} \approx_g \phi$*
(ii) *If $\mathcal{K} \approx_f \phi$, then $\mathcal{K} \approx^\star_f \phi$*
(iii) *If $\mathcal{K} \approx^\star_f \phi$, then $\mathcal{K} \approx^\star_g \phi$.*

Proof. (i) Assume that $\mathcal{K} \approx_f \phi$. If $\mathcal{K}^\top \neq \emptyset$, the implication follows trivially. Hence, suppose that $\mathcal{K}^\top = \emptyset$. Take a valuation u such that $u \models \partial(\mathcal{K}_g)$. Then $u \models \partial(g(\mathcal{K}^\perp))$. Then there is a v in $g(\mathcal{K}^\perp)$, such that $u \models \bigwedge v$. Hence, $v \subseteq u$ and $v \in \mathsf{max}(\{v \subseteq w : w \in \mathcal{K}$ *and* v *is consistent*$\})$. Take a w in $\mathcal{K}^\perp$ with $v \subseteq w$. First, we show that $\overline{w} \subseteq v$. Suppose $l \in \overline{w}$, where l denotes a literal and $(\neg)l$ its literal negation. Then $l \in w$ and $(\neg)l \notin w$. Hence, for all consistent subvaluations x of w it holds that $x \cup \{l\}$ is consistent. Hence $l \in v$, since v is a maximally consistent subvaluation of w. Since $\overline{w} \subseteq v$ and $v \subseteq u$, it must be that $\overline{w} \subseteq u$. Hence, $u \models \bigwedge \overline{w}$. Hence, $u \models \partial(f(\mathcal{K}^\perp))$. Hence, $u \models \partial(\mathcal{K}_f)$. Hence, by our initial assumption, $u \models \phi$. Then it must be that $\partial(\mathcal{K}_g) \models \phi$. Therefore, $\mathcal{K} \approx_g \phi$.

(ii) Assume that $\mathcal{K} \approx_f \phi$. Then it holds that $\partial(\mathcal{K}_f) \models \phi$. Split cases as to whether $\mathcal{K}^\top$ is empty or non-empty. Suppose, in the first case, that $\mathcal{K}^\top \neq \emptyset$. Then $\mathcal{K}^\star = \mathcal{K}^\top$, by Lemma 4.4.8. Hence, $(\mathcal{K}^\star)^\top \neq \emptyset$. Therefore, $\mathcal{K}_f = \mathcal{K}^\top = \mathcal{K}^\star = (\mathcal{K}^\star)^\top = (\mathcal{K}^\star)_f$. Therefore, by the intitial assumption, $\partial((\mathcal{K}^\star)_f) \models \phi$. Therefore, $\mathcal{K} \approx^\star_f \phi$. Suppose now, in the second case, that $\mathcal{K}^\top = \emptyset$. Then, by Lemma 4.4.8, $\mathcal{K}^\star \subseteq \mathcal{K}^\perp$. Hence, $(\mathcal{K}^\star)^\top = \emptyset$ and $(\mathcal{K}^\star)^\perp = \mathcal{K}^\star$. Suppose that $u \models \partial((\mathcal{K}^\star)_f)$. Then $u \models \partial(f((\mathcal{K}^\star)^\perp))$. Hence, there is a v in $f((\mathcal{K}^\star)^\perp)$ with $u \models \bigwedge v$. Then there is a w in $(\mathcal{K}^\star)^\perp$ with $v = \overline{w}$. Then there is a w in $\mathcal{K}^\perp$ with $v = \overline{w}$. Then there is a v in $f(\mathcal{K}^\perp)$ with $u \models \bigwedge v$. Hence, $u \models \partial(f(\mathcal{K}^\perp))$. Therefore, $u \models \partial(\mathcal{K}_f)$. By the assumption, it follows that $u \models \phi$. Therefore, $\mathcal{K} \approx^\star_f \phi$.

(iii) The proof is similar to that in (i). □

4.5 Future Work

4.5.1 Finite State Belief Revision

The four extractors discussed in the preceding section can be amended to obtain their classical counterparts. We take $\mathcal{K} \Vert\!\approx^\circ_x \phi$ as a notation for the general form of the *classical extractors*, which are obtained by interpreting the '$\models$' in Definitions 4.4.3 and 4.4.9 to denote *classical* validity. Following this amendment, it is feasible to describe four systems of belief revision, as follows:

4.5.1. Definition. Let $\mathcal{K}$ be a finite state, let $x \in \{f, g\}$, and let $\circ$ denote either selected or unselected extraction. The *classical belief set based on* $\mathcal{K}$, denoted by K, is defined to be

$$K \quad = \quad \{\phi : \mathcal{K} \Vert\!\approx^\circ_x \phi\}.$$

Expansion and contraction for belief sets can now be defined in terms of expansion and contraction for finite states. Revision and reversed revision can, of course, be defined similarly.

4.5.2. DEFINITION. Let K be a classical belief set based on a finite state $\mathcal{K}$ and let ϕ be a formula. Then *the expansion of* K *with* ϕ, denoted by K^+_ϕ, and *the contraction of* K *with* ϕ, denoted by K^-_ϕ, are defined to be

$$\begin{array}{lcl} K^+_\phi & = & \{\psi : \mathcal{K} + \phi \mathrel{|\!\!\approx^\circ_x} \psi\} \\ K^-_\phi & = & \{\psi : \mathcal{K} \dot{-} \phi \mathrel{|\!\!\approx^\circ_x} \psi\}. \end{array}$$

Obviously, this definition gives rise to four different systems of belief revision, depending on the classical extractor that has been used to obtain the classical belief set. (A detailed study of the merits and drawbacks of our extractors may reduce this number.) The four systems share some remarkable general properties: (i) expansion and contraction are functions from classically closed belief sets to classically closed belief sets, as in almost all existing systems of belief revision; (ii) belief sets *always* are consistent, hence the systems are all non-prioritized systems of belief revision,[19] since the *Succes Postulate for Revision* ($\phi \in K^\times_\phi$) does not hold in case the belief set is revised by a contradiction; and (iii) iterated revision presents no problems, as the belief set resulting from the application of an operation of change is uniquely determined. No ordering of (sets of) propositions, that needs to be reassessed after each contraction or revision, is used, contrary to most systems presented in the literature.

A study into the sets of postulates characterizing these systems of belief revision would facilitate a comparison between them and the systems of non-prioritized belief revision studied in the literature.

4.5.2 Relative Trustworthiness of the Sources

In a sense, our study is complementary to [Cantwell, 1998]. Cantwell studies methods governing the acceptance of information stemming from sources which are at least partially ordered according to their relative trustworthiness, treating all the information provided by a source *en bloc*. At first sight, it might seem that the present study considers just a special case of Cantwell's theory, since we assumed all information to be equally trustworthy and equality is a partial ordering. The next quote from Cantwell's paper suffices to dispel this idea:

> [T]here are two quite different ways of treating information from a source. Each source provides a *scenario* for what has happened or what the world is like. Either one can treat the scenario as a whole, dismissing it in its entirety if it turns out wrong, or one can treat each piece of information separately, allowing a source to be wrong

[19] For an overview of non-prioritized belief revision, see [Hansson, 1999b].

> in certain respects and right in others. [...] I shall treat the scenario presented by a source as a whole and not sentence by sentence. [Cantwell, 1998, p. 193]

Of course, this does not mean that Cantwell's approach is wrong-headed. Considerations having to do with the trustworthiness of the sources that provide information do play a role in the fixation of beliefs based on this information. On the other hand, people (except for attorneys in American lawsuit television series) usually do not treat the information provided by a source 'as a whole', but rather sort out this information, weighed by the relative trustworthiness of the source. In the rare cases where we are able to assess such a trustworthiness-ordering, a combination of Cantwell's approach and mine might be fruitful.

4.5.3 Inconsistency-Adaptive Logics

Inconsistency-adaptive logics provide a semantical framework for traditional approaches to reasoning from inconsistent information [Batens, manuscript]. Our *belief extractors*, especially the *selective* ones, share some characteristics with the inconsistency-adaptive logics which have been developed by Batens and his associates. Precise relations between their logics and the logics set forth in the present chapter have yet to be established. It would be interesting to study these relations, since, at first sight at least, the semantics of our respective approaches differ wildly.

Next to the mentioned semantical investigations into inconsistency-adaptive logics, proof-theories or tableau-methods have been developed for most of these logics as well. Exploiting the analogy between inconsistency-adaptive logics and our belief extractors, insights gained from the proof-theoretical investigations into inconsistency-adaptive logics may also prove useful for the construction of proof-theories for our belief extractors.

Chapter 5

Conclusion

5.1 Summary of the Results

The introductory chapter served to present the *belief-doubt-belief* model that was first conceived within 19th century American pragmatism, in an attempt to reconcile the theory of knowledge with the insights from Darwin's evolutionary biology. Thus, an alternative for Aristotle's ideal of knowledge was developed, an alternative that can do without the assumption that science needs a secure foundation in the form of a set of evident axioms from which all truths can be deduced. It was shown that attempts, still inspired by the demands of traditional philosophy, to define the concepts of truth and reality in terms of the newly conceived *belief-doubt-belief* model are, to say the least, somewhat premature. Instead, it was proposed to focus our attention on more precise formulations of the model.

Consequently, the logical epistemology of the most prominent contemporary advocate of the *belief-doubt-belief* model, Isaac Levi, has been expounded at length in the third chapter. In the course of this exposition, serious drawbacks of the approach favoured by Levi came to the fore. Levi's claims to construct a normative theory about belief change could not be substantiated, let alone that his theory admits an empirical test. These characteristics are unacceptable. In the face of recent pleas, among others from naturalistic philosophers, for empirical evaluations of philosophical theories, it was argued that the field of belief change urgently needs a reformulation in order to make possible empirical assessments of the merits of the proposed theories. A fruitful way to address this reformulation problem would be an approach that concentrates on eliminating parameters in theories about belief change which have withstood practical assessment so far.

In the fourth chapter, this new program for the field of belief change was carried out in detail. First, by representing the information offered as a *finite state*, we avoided the computational problems that are posed by representations of information by deductively closed sets of sentences. Second, rather than rely on extra-logical considerations in the definition of contraction, the newly proposed system had a purely logical contraction operator. Third, unlike the host of the systems for belief change that have been presented in the literature, the system could cope with inconsistencies, which, if necessary, could be removed one by one. We have not only developed techniques to change a finite state via expansion or contraction, but have indicated methods to extract consistent 'beliefs' from inconsistent finite states as well.[1]

Is this all there is? Obviously, there is much more to belief change than avoiding bewilderment in the presence of contradictions:

> Die Anpassung der Gedanken aneinander erschöpft sich nicht in der Abschleifung der Widersprüche. [...] Das Ökonomisieren, Harmonisieren, Organisieren der Gedanken, welches wir als ein biologisches Bedürfnis fühlen, geht weit über die Forderung der *logischen Widerspruchslosigkeit* hinaus. [Mach, 1905, p. 176]

5.2 Future Work

Since the publication of Quine's programmatic 'Epistemology Naturalized' (1969), in which Quine prompts to undertake *empirical* investigations into the connections between our 'theory of the world' and our sensory stimuli,[2] epistemologists have become more and more interested in something which could be called aptly *metanaturalism*. Rather than concentrate on the construction and evaluation of theories about cognition, the contemporary philosophical discussion has been limited largely to the Kant-style question as to how, in the face of the demise of rationalistic approaches to epistemology, a naturalistic epistemology might still be possible. (In the meantime, cognitive psychologists, neurologists, logicians and researchers within AI do the jobs epistemologists used to do.)

The recent collection of essays *Naturalismus* (2000), edited by the German philosophers Keil and Schnädelbach, is a case in point. In an introductory essay, the editors address the problem of defining the term 'naturalism', discussing seventeen proposals for a definition. Koppelberg, in his contribution, makes clear that naturalism must not be equated to empiricism, fysicalism, scientism, or antimentalism. Moreover, the editors claim "daß sich mit der Opposition gegen den Supranaturalismus heute keine interessante Position mehr markieren läßt"

---

[1] The extractors, which have been defined in the last section of the fourth chapter, fulfill at least one of Nuel Belnap's wishes: "The complete reasoner should, presumably, have some strategy for *giving up* part of what it believes when it finds its beliefs inconsistent. [...] I have never heard of a practical, reasonable, mechanizable strategy for revision of belief in the presence of contradiction" [Belnap, 1977, p. 9].

[2] "The stimulation of his sensory receptors is all the evidence anybody has had to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds?" [Quine, 1969, p. 75.]

[Keil and Schnädelbach, 2000, p. 31]. The present thesis has shown that, although antirationalism might not be worth the trouble of yet another philosophical defense, traces of rationalistic theorizing abound in contemporary epistemology and philosophical logic.

By and large, the discussion in the belief change literature has been focussed largely on technicalities and on *a priori* argumentations in favor of a specific choice of a set of postulates. Lots of different systems have been proposed, lots of different formal desiderata have been formulated and defended, usual on the basis of concise, *a priori* argumentations, that ultimately rely on pre-theoretical intuitions. As a starting phase, this is how it should be. It would be a mistake, however, never to leave the formal considerations for what they are and go on to try to put the proposed systems to an empirical test. Only by testing formal belief change theories empirically, can convincing arguments for or against a specific approach be found, since it would not be the first time in philosophical or mathematical research that exclusive reliance on intuitions led us astray. After all, unlike research in, for example, modal logic, investigations into belief change do purport to *explain* some phenomena in an empirical domain. It would be a grave mistake to conceive of belief change research as a wholly rationalistic (in the philosophical sense of the word) matter.

The investigations that have been conducted in the present thesis have, of course, not been empirical. Nevertheless, they have made an empirical test possible, because of the elimination of extra-logical considerations in the definition of contraction. Hence, the system presented in the fourth chapter ought to be seen as an attempt to 'pragmatize' the field of belief revision,[3] that is, to make the field of belief revision more conducive to empirical considerations than systems advocated by Levi, by Alchourrón, Gärdenfors and Makinson, and by their sympathizers.

Thus, instead of indulging in purely academic debates concerning the proper definition of naturalism, in my opinion, the spirit of the naturalized epistemology programme is better served by a close examination of contemporary theories about cognition. We should detect those elements in these theories that obstruct empirical ways of assessing their merits. We should reformulate theories which possess such elements in order to make empirical evaluations of them possible. But most of all, we should actually carry out these empirical tests and revise our theories according to their outcomes. In the present thesis, the first and second points have been put into practice. The third point, applied to the subject matter of the present thesis, could be implemented, for instance, by the following

[3] "Such a pragmatization brings an already existing theory – a system – into closer contact with the empirical world and with practical life, and this, in turn, can inspire empirical research of different kinds, which may seem irrelevant as long as the theory (for instance, the logical system) is presented in its original objectivistic or solipsist-subjectivistic garb. Perhaps, pragmatizing can not be called empirical research, but we may characterize it as a *proto-empirical* activity." [Barth, 1987, p. 11 – my translation; AMT.]

program:

First and foremost, an empirical test of the fourth chapter's system for belief change has to be carried out. In collaboration with cognitive psychologists, test conditions must be specified in detail. Great care must be taken that the information offered is really considered equally trustworthy by the test subjects. This can be partly realized, for instance, by making sure that the test subjects acknowledge that the order in which the information is offered to them is wholly arbitrary. Moreover, the information must be selected in such a way, that the test subjects have as few preconceived opinions as possible about the relative plausibility of the information offered. Then, prompted by logical theory, strategic questions should be asked relating to the plausibility of selected propositions. The test subjects should answer these questions solely on the basis of logic and the controlled information they received previously. Next, of course, it should be checked whether their answers are more or less in line with the answers provided by our system for belief change.

Furthermore, it would certainly be of interest to generalize the theory of belief change that was presented in the fourth chapter to include not only equable reliable sources, but sources that differ in trustworthiness as well. Though such a generalization would hardly affect the logic of the system, it would, first, facilitate a comparison with belief change systems in the tradition initiated by Levi and by Alchourrón, Gärdenfors and Makinson, since in such a generalized system as well as in traditional belief change systems, extra-logical considerations do play a role. Second, as attributions of trustworthiness do play a role in the real-life evaluation of received information, the generalized theory purports to cover more phenomena than the original theory.

The reintroduction of extra-logical elements would, however, fall short of being an advance, unless well-argued methods are set forth according to which the relative trustworthiness of the sources can be assessed. Here, philosophical, psychological and economical literature on trust will prove relevant. Moreover, a multi-agent perspective seems to be incumbent, since the assessment of the trustworthiness of a source does not amount to the discovery of a timeless and objective property, but will rather be the assessment of a function, that might change over time, of experiences the assessing agent has had with this source and with other sources as well.

Appendix A

A Natural Deduction System for First Degree Entailment

A.1 Introduction

A.1.1 First Degree Entailment

After being inspired by the work of Ackermann [Ackermann, 1956], Anderson and Belnap started their investigation into a theory of implication: if ... then ___ [Anderson and Belnap, 1975]. They developed a number of formal calculi of *entailment*, which later came to be known as *Relevant Logics*.[1] In developing their systems, Anderson and Belnap encountered the difficulty of dealing with nested entailments. Consequently, they considered first degree entailment (fde), in which the antecedent ϕ and consequent ψ in an implicational sentence of the form $\phi \rightarrow \psi$ are truth functional, that is, ϕ and ψ themselves do not contain implications. The purpose of the investigation into fde is, then, to study the truth functional relationship between antecedent and consequent of implicational sentences.

Anderson and Belnap provide a Hilbert-style system and a Gentzen-style system for fde. However, although they give characteristic matrices, Anderson and Belnap do not provide any formal semantics for fde. For this, we had to wait for [Routley and Routley, 1972] and [Dunn, 1976].[2] Routley and Routley provide a two-valued semantics for fde. Although their semantics may be philosophically contentious, it serves as a basis for the semantics for various relevant logics.[3] However, in this Appendix, we are concerned only with Dunn's semantics, which

[1]Or *Relevance Logics*. '*Relevant* logics' is often preferred by Australian relevan*t* logicians. '*Relevance* logics', on the other hand, is preferred by American relevan*ce* logicians.

[2]The paper is included in [Anderson, Belnap, and Dunn, 1992], which is the second volume of [Anderson and Belnap, 1975].

[3]See [Priest and Sylvan, 1992].

is somewhat more intuitive. Together with a tableau system, Dunn presents an ‘intuitive’ formal semantics for fde. Classically, semantic evaluations of sentences are defined to be functions that assign to a formula exactly one truth value. For Dunn, however, evaluations are relations between a truth value and a formula. A formula may, then, take (relate to) no truth value, or may take (relate to) multiple truth values.[4]

One feature of Dunn’s semantics for fde that we should take notice of is that truth and falsity are not mutually complementary. Truth and falsity are considered separately and are independent notions in Dunn’s semantics. This feature plays an important role in developing a natural deduction system for fde later in this Appendix.

A.1.2 Combined Systems

The idea of considering true and false formulas separately can also be found in the study of formal logics for ‘assertion’ and ‘rejection’. Łukasiewicz was, to the best of our knowledge, the first to introduce both a sign for ‘assertion’ and a sign for ‘rejection’ into formal logic. Tracing back the history of the philosophy of logic, Łukasiewicz followed Brentano (1838–1917), who propounded a nonpropositional theory of judgment. Brentano argued that

> As every judgement is based on an idea, the statement expressing a judgement necessarily contains a name [of the idea]. To this, another sign must come, a sign corresponding to the inner state which we call judging, that is, a sign completing the bare name to a sentence. And because this judging can be twofold, *viz.*, asserting or rejecting, the sign indicating it must be twofold too, one for affirmation and one for denial. These signs themselves do not mean anything, but in conjunction with a name, they are the expression of a judgement. Therefore, the most general scheme of a statement is ‘A is’ and ‘A is not’ [translation by the authors].[5]

In the 1921 paper ‘Logika dwuwartościowa’, which was later translated as ‘Two-valued Logic’, Łukasiewicz followed Brentano in adding to Frege’s idea of assertion

[4] Some *paraconsistent logics* are on this idea. Unsurprisingly, fde is often considered to be a paraconsistent logic, as well as a relevant logic.

[5] Brentano writes: “Da jedem Urteil eine Vorstellung zugrunde liegt, so wird die Aussage als Ausdruck des Urteils notwendig einen Namen enthalten. Dazu wird aber noch ein anderes Zeichen kommen müssen, das demjenigen inneren Zustand entspricht, den wir eben Urteilen nennen, d.h. ein Zeichen, das den bloßen Namen zum Satz ergänzt. Und da dieses Urteilen von doppelter Art sein kann, nämlich ein Anerkennen oder Verwerfen, so wird auch das Zeichen dafür ein doppeltes sein müssen, eines für die Bejahung und eines für die Verneinung. Für sich allein bedeuten diese Zeichen nichts [...], aber in Verbindung mit einem Namen sind sie Ausdruck eines Urteils. Das allgemeinste Schema der Aussage lautet daher: A ist (A +) und A ist nicht (A −).” [Brentano, *c*.1887, p. 97–98.]

Brentano's idea of rejection. In his early works, Łukasiewicz argued that a proposition must be rejected if and only if it is false, parallel with Frege's condition for the assertion of a proposition.[6] Later, starting with *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic* (1951), Łukasiewicz redefined the concept of rejection to cover not only false propositions, but propositions which are false under at least one interpretation as well. Furthermore, he introduced syntactical techniques to *derive* all rejected, that is, nontautological, statements. By using the symbol '$\vdash$' for assertion (indicating tautologyhood) and '$\dashv$' for rejection (indicating nontautologyhood), what Łukasiewicz added to classical propositional logic (cpl) is the following:

Axiom	$\dashv p$, where p is a fixed propositional variable.
Detachment	If $\vdash \phi \to \psi$ and $\dashv \psi$, then $\dashv \phi$.
Substitution	If $\dashv \psi$ and ψ can be obtained from ϕ by substitution, then $\dashv \phi$.

The system is first presented in [Łukasiewicz, 1951], where Łukasiewicz also propounded a system of rejection for Aristotle's syllogistic, after some technical problems had been solved by Słupecki. Łukasiewicz also tried to construe systems of rejection for the intuitionistic propositional logic (ipl) and for his own version of modal logic. All these systems share one characteristic: they are all 'combined systems', that is, they all include both a sign for 'assertion' and a sign for 'rejection'.[7]

One of the advantages of combined systems over traditional ones that is worth mentioning in this Appendix is that metatheoretical results can be incorporated in the object language of the system under consideration. For instance, the disjunction property of ipl can be formulated in the object language of a proof system as follows:

$$\frac{\dashv \phi \qquad \dashv \psi}{\dashv \phi \vee \psi}\,.$$

Now, since in many of the combined systems, the systems of Łukasiewicz in particular, $\dashv \phi$ is complemented by the failure of $\vdash \phi$,[8] the concept of rejection contained in these systems is classical. Nonetheless, combined systems, prima facie, take the idea seriously that (possibly) false formulas be considered separately from true formulas. This idea is congenial to Dunn's semantics.

[6] Lukasiewicz writes: "I wish to assert truth and only truth, and to reject falsehood and only falsehood." [Łukasiewicz, 1921, p. 91.]

[7] For a synopsis of the history of theories of rejection for cpl, the reader may have recourse to [Tamminga, 1994].

[8] For a discussion of this feature of combined systems, see [Tamminga, 1994].

A.2 Language and Semantics

A.2.1. Definition. The alphabet of fde consists of the following.

(i)	Propositional variables	$p_1, p_2, p_3, \ldots$
(ii)	Logical symbols	$\neg, \wedge, \vee$
(iii)	Auxiliary symbols	$), ($

$\Box$ denotes an empty sequence. $\mathcal{A}$ denotes the set of propositional variables.

A.2.2. Definition. The set of all formulas of fde, denoted by $\mathcal{F}$, is the least set satisfying the following conditions:

(i) $\mathcal{A} \subset \mathcal{F}$,
(ii) $\phi, \psi \in \mathcal{F} \implies (\phi \wedge \psi), (\phi \vee \psi) \in \mathcal{F}$,
(iii) $\phi \in \mathcal{F} \implies \neg\phi \in \mathcal{F}$.

A.2.3. Definition. Let $\mathcal{M} = \langle \mathcal{F}, \nu \rangle$ be an interpretation for the language where ν is an evaluation such that ν is a function from $\mathcal{A}$ to $\wp(\{0,1\})$. Then ν is extended to an evaluation $\nu_{\mathcal{M}}$ for all formulas ϕ and ψ by the following conditions:

(i) $\nu \subseteq \nu_{\mathcal{M}}$
(ii) $1 \in \nu_{\mathcal{M}}(\phi \wedge \psi) \iff 1 \in \nu_{\mathcal{M}}(\phi)$ and $1 \in \nu_{\mathcal{M}}(\psi)$,
(iii) $0 \in \nu_{\mathcal{M}}(\phi \wedge \psi) \iff 0 \in \nu_{\mathcal{M}}(\phi)$ or $0 \in \nu_{\mathcal{M}}(\psi)$,
(iv) $1 \in \nu_{\mathcal{M}}(\phi \vee \psi) \iff 1 \in \nu_{\mathcal{M}}(\phi)$ or $1 \in \nu_{\mathcal{M}}(\psi)$,
(v) $0 \in \nu_{\mathcal{M}}(\phi \vee \psi) \iff 0 \in \nu_{\mathcal{M}}(\phi)$ and $0 \in \nu_{\mathcal{M}}(\psi)$,
(vi) $1 \in \nu_{\mathcal{M}}(\neg\phi) \iff 0 \in \nu_{\mathcal{M}}(\phi)$,
(vii) $0 \in \nu_{\mathcal{M}}(\neg\phi) \iff 1 \in \nu_{\mathcal{M}}(\phi)$.

A.2.4. Definition. Let $\Pi \subseteq \mathcal{F}$ and $\mathcal{M}$ be an interpretation. Then

(i) $1 \in \nu_{\mathcal{M}}(\Pi) := 1 \in \nu_{\mathcal{M}}(\phi)$ for every $\phi \in \Pi$,
(ii) $0 \in \nu_{\mathcal{M}}(\Pi) := 0 \in \nu_{\mathcal{M}}(\phi)$ for every $\phi \in \Pi$.

We are now in a position to define validity. Validity defined below incorporates the concept of Dunn's semantics for fde. It concerns not only truth but also falsity as in Dunn's semantics.[9]

A.2.5. Definition. [fde Validity] Let $\Pi, \Gamma \subseteq \mathcal{F}$ and $\phi \in \mathcal{F}$. Then

(i) $\Pi; \Gamma \models \phi; \Box \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$, then $1 \in \nu_{\mathcal{M}}(\phi)$,
(ii) $\Pi; \Gamma \models \Box; \phi \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$, then $0 \in \nu_{\mathcal{M}}(\phi)$.

[9]Standardly, validity for fde is defined as in classical logic as follows:

$$\Pi \models \phi \iff \textit{For all } \mathcal{M} : \textit{if } 1 \in \nu_{\mathcal{M}}(\Pi), \textit{then } 1 \in \nu_{\mathcal{M}}(\phi).$$

A.3 A Natural Deduction System

While providing a Hilbert-style system and a Gentzen-style system and natural deduction systems for other relevant logics, Anderson and Belnap do not give any natural deduction system for fde. The first natural deduction system for fde to be formally introduced, other than the system developed in this Appendix, will be included in [Priest, 2001].[10]

In this section, we introduce a natural deduction system ND for fde. The system is developed by amalgamating the concept of Dunn's semantics and that of the combined systems. Instead of taking $\vdash \phi$ to be an assertion of ϕ (a usual policy in combined systems), will it here be semantically interpreted as: ϕ takes 'truth' as *a* truth value. Similarly, $\dashv \phi$ is interpreted as: ϕ takes 'falsity' as *a* truth value.

The system ND is defined as follows.[11]

A.3.1. Definition. Derivations in the system ND are inductively generated as follows.

Basis: The proof tree with a single occurrence of an assumption $\vdash \phi$ or $\dashv \phi$ is a derivation.

Induction Step: Let $\mathcal{D}$, $\mathcal{D}_1$, $\mathcal{D}_2$, $\mathcal{D}_3$ be derivations. Then they can be extended by the following rules:

$$\dfrac{\begin{matrix}\mathcal{D}\\ \vdash \phi\end{matrix}}{\dashv \neg\phi}\neg I_{\vdash} \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \vdash \neg\phi\end{matrix}}{\dashv \phi}\neg E_{\vdash} \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \dashv \phi\end{matrix}}{\vdash \neg\phi}\neg I_{\dashv} \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \dashv \neg\phi\end{matrix}}{\vdash \phi}\neg E_{\dashv}$$

$$\dfrac{\begin{matrix}\mathcal{D}_1\\ \vdash \phi\end{matrix} \qquad \begin{matrix}\mathcal{D}_2\\ \vdash \psi\end{matrix}}{\vdash \phi \wedge \psi}\wedge I_{\vdash} \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \vdash \phi_0 \wedge \phi_1\end{matrix}}{\vdash \phi_i}\wedge E_{\vdash},\ \ i \in \{0,1\}$$

$$\dfrac{\begin{matrix}\\ \mathcal{D}_1\\ \dashv \phi \wedge \psi\end{matrix} \qquad \begin{matrix}[\dashv \phi]^u\\ \mathcal{D}_2\\ X\end{matrix} \qquad \begin{matrix}[\dashv \psi]^v\\ \mathcal{D}_3\\ X\end{matrix}}{X}\wedge E_{\dashv}^{u,v},\ \ \text{where } X = \vdash \chi \text{ or } X = \dashv \chi$$

$$\dfrac{\begin{matrix}\mathcal{D}\\ \dashv \phi_i\end{matrix}}{\dashv \phi_0 \wedge \phi_1}\wedge I_{\dashv},\ \ i \in \{0,1\} \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \vdash \phi_i\end{matrix}}{\vdash \phi_0 \vee \phi_1}\vee I_{\vdash},\ \ i \in \{0,1\}$$

---

[10]After the development of Dunn's semantics, the history of fde is largely anecdotal. For this reason, it is uncertain whether the system provided by Priest will be the first. However, there do not seem to be any published papers that introduce natural deduction systems for fde. This claim was suggested in conversations with Dunn and Priest.

[11]The notational conventions used in the definition of ND are only a slight modification of those of [Troelstra and Schwichtenberg, 1996].

$$\dfrac{\begin{matrix}\mathcal{D}_1\\ \dashv \phi\end{matrix}\qquad \begin{matrix}\mathcal{D}_2\\ \dashv \psi\end{matrix}}{\dashv \phi\vee\psi}\vee I_\dashv \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}\\ \dashv \phi_0\vee\phi_1\end{matrix}}{\dashv \phi_i}\vee E_\dashv,\ \ i\in\{0,1\}$$

$$\dfrac{\begin{matrix}\mathcal{D}_1\\ \vdash \phi\vee\psi\end{matrix}\qquad \begin{matrix}[\vdash\phi]^u\\ \mathcal{D}_2\\ X\end{matrix}\qquad \begin{matrix}[\vdash\psi]^v\\ \mathcal{D}_3\\ X\end{matrix}}{X}\vee E_\vdash^{u,v},\ \ \text{where } X=\vdash\chi \text{ or } X=\dashv\chi$$

where $[\vdash \phi]$ and $[\dashv \phi]$ are assumptions which are discharged by the application of the rules.

A.3.2. LEMMA. *De Morgan rules of the following forms are available in* ND *(double lines indicate that the rules work both ways):*

$$\begin{array}{c}\mathcal{D}\\ \vdash \phi\wedge\psi\\ \hline\hline \dashv \neg\phi\vee\neg\psi\end{array}\ DeM_\vdash \qquad\qquad \begin{array}{c}\mathcal{D}\\ \dashv \phi\wedge\psi\\ \hline\hline \vdash \neg\phi\vee\neg\psi\end{array}\ DeM_\dashv$$

Proof. $DeM_\vdash$:

$$\dfrac{\dfrac{\dfrac{\vdash\phi\wedge\psi}{\vdash\phi}}{\dashv\neg\phi}\qquad \dfrac{\dfrac{\vdash\phi\wedge\psi}{\vdash\psi}\wedge E_\vdash}{\dashv\neg\psi}\neg I_\vdash}{\dashv\neg\phi\vee\neg\psi}\vee I_\dashv$$

$$\dfrac{\dfrac{\dfrac{\dashv\neg\phi\vee\neg\psi}{\dashv\neg\phi}}{\vdash\phi}\qquad \dfrac{\dfrac{\dashv\neg\phi\vee\neg\psi}{\dashv\neg\psi}\vee E_\dashv}{\vdash\psi}\neg E_\dashv}{\vdash\phi\wedge\psi}\wedge I_\vdash$$

$DeM_\dashv$:

$$\dfrac{\dashv\phi\wedge\psi\qquad \dfrac{\dfrac{[\dashv\phi]^u}{\vdash\neg\phi}}{\vdash\neg\phi\vee\neg\psi}\qquad \dfrac{\dfrac{[\dashv\psi]^v}{\vdash\neg\psi}\neg I_\dashv}{\vdash\neg\phi\vee\neg\psi}\vee I_\vdash}{\vdash\neg\phi\vee\neg\psi}\wedge E_\dashv^{u,v}$$

$$\dfrac{\vdash\neg\phi\vee\neg\psi\qquad \dfrac{\dfrac{[\vdash\neg\phi]^u}{\dashv\phi}}{\dashv\phi\wedge\psi}\qquad \dfrac{\dfrac{[\vdash\neg\psi]^v}{\dashv\psi}\neg E_\vdash}{\dashv\phi\wedge\psi}\wedge I_\dashv}{\dashv\phi\wedge\psi}\vee E_\vdash^{u,v}$$

□

A.3.3. DEFINITION. Let $\Pi \subseteq \mathcal{F}$. Then

$$\begin{array}{rlcl} \text{(i)} & \vdash \Pi & := & \{\vdash \phi : \phi \in \Pi\}, \\ \text{(ii)} & \dashv \Pi & := & \{\dashv \phi : \phi \in \Pi\}. \end{array}$$

A.3.4. DEFINITION. [fde Derivability] Let $\Pi, \Gamma \subseteq \mathcal{F}$ and $\phi \in \mathcal{F}$. Then

$$\begin{array}{rlcl} \text{(i)} & \Pi; \Gamma \mapsto \phi; \Box & \iff & \text{There is a derivation in ND of} \vdash \phi \text{ from} \vdash \Pi \cup \dashv \Gamma, \\ \text{(ii)} & \Pi; \Gamma \mapsto \Box; \phi & \iff & \text{There is a derivation in ND of} \dashv \phi \text{ from} \vdash \Pi \cup \dashv \Gamma. \end{array}$$

A.4 Soundness

A.4.1. LEMMA. *Let $\Pi_i, \Gamma_i \subseteq \mathcal{F}$ for $i \in \{1, 2, 3\}$ and $\phi, \psi, \chi \in \mathcal{F}$. Then*

$$\begin{array}{rlcl}
\text{(i)} & \Pi; \Gamma \models \phi; \Box, & \textit{if} & \phi \in \Pi \\
\text{(ii)} & \Pi; \Gamma \models \Box; \phi, & \textit{if} & \phi \in \Gamma \\
\text{(iii)} & \Pi; \Gamma \models \phi; \Box & \Longrightarrow & \Pi; \Gamma \models \Box; \neg\phi \\
\text{(iv)} & \Pi; \Gamma \models \Box; \phi & \Longrightarrow & \Pi; \Gamma \models \neg\phi; \Box \\
\text{(v)} & \Pi; \Gamma \models \neg\phi; \Box & \Longrightarrow & \Pi; \Gamma \models \Box; \phi \\
\text{(vi)} & \Pi; \Gamma \models \Box; \neg\phi & \Longrightarrow & \Pi; \Gamma \models \phi; \Box \\
\text{(vii)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \phi; \Box \\ \Pi_2; \Gamma_2 \models \psi; \Box \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2; \Gamma_1, \Gamma_2 \models \phi \wedge \psi; \Box \\
\text{(viii)} & \Pi; \Gamma \models \Box; \phi & \Longrightarrow & \Pi; \Gamma \models \Box; \phi \wedge \psi \\
\text{(ix)} & \Pi; \Gamma \models \Box; \psi & \Longrightarrow & \Pi; \Gamma \models \Box; \phi \wedge \psi \\
\text{(x)} & \Pi; \Gamma \models \phi \wedge \psi; \Box & \Longrightarrow & \Pi; \Gamma \models \phi; \Box \\
\text{(xi)} & \Pi; \Gamma \models \phi \wedge \psi; \Box & \Longrightarrow & \Pi; \Gamma \models \psi; \Box \\
\text{(xii)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \Box; \phi \wedge \psi \\ \Pi_2; \Gamma_2, \phi \models \chi; \Box \\ \Pi_3; \Gamma_3, \psi \models \chi; \Box \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2, \Pi_3; \Gamma_1, \Gamma_2, \Gamma_3 \models \chi; \Box \\
\text{(xiii)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \Box; \phi \wedge \psi \\ \Pi_2; \Gamma_2, \phi \models \Box; \chi \\ \Pi_3; \Gamma_3, \psi \models \Box; \chi \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2, \Pi_3; \Gamma_1, \Gamma_2, \Gamma_3 \models \Box; \chi \\
\text{(xiv)} & \Pi; \Gamma \models \phi; \Box & \Longrightarrow & \Pi; \Gamma \models \phi \vee \psi; \Box \\
\text{(xv)} & \Pi; \Gamma \models \psi; \Box & \Longrightarrow & \Pi; \Gamma \models \phi \vee \psi; \Box \\
\text{(xvi)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \Box; \phi \\ \Pi_2; \Gamma_2 \models \Box; \psi \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2; \Gamma_1, \Gamma_2 \models \Box; \phi \vee \psi \\
\text{(xvii)} & \Pi; \Gamma \models \Box; \phi \vee \psi & \Longrightarrow & \Pi; \Gamma \models \Box; \phi \\
\text{(xviii)} & \Pi; \Gamma \models \Box; \phi \vee \psi & \Longrightarrow & \Pi; \Gamma \models \Box; \psi \\
\text{(xix)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \phi \vee \psi; \Box \\ \Pi_2, \phi; \Gamma_2 \models \chi; \Box \\ \Pi_3, \psi; \Gamma_3 \models \chi; \Box \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2, \Pi_3; \Gamma_1, \Gamma_2, \Gamma_3 \models \chi; \Box \\
\text{(xx)} & \left.\begin{array}{l} \Pi_1; \Gamma_1 \models \phi \vee \psi; \Box \\ \Pi_2, \phi; \Gamma_2 \models \Box; \chi \\ \Pi_3, \psi; \Gamma_3 \models \Box; \chi \end{array}\right\} & \Longrightarrow & \Pi_1, \Pi_2, \Pi_3; \Gamma_1, \Gamma_2, \Gamma_3 \models \Box; \chi.
\end{array}$$

Proof. (xii) Let $\mathcal{M}$ be an interpretation, such that $1 \in \nu_{\mathcal{M}}(\Pi_1, \Pi_2, \Pi_3)$ and $0 \in \nu_{\mathcal{M}}(\Gamma_1, \Gamma_2, \Gamma_3)$. Then, as $\Pi_1; \Gamma_1 \models \Box; \phi \wedge \psi$, we have $0 \in \nu_{\mathcal{M}}(\phi \wedge \psi)$. Therefore, $0 \in \nu_{\mathcal{M}}(\phi)$ or $0 \in \nu_{\mathcal{M}}(\psi)$. Suppose $0 \in \nu_{\mathcal{M}}(\phi)$. Then, as $\Pi_2; \Gamma_2, \phi \models \chi; \Box$, we have $1 \in \nu_{\mathcal{M}}(\chi)$. Suppose $0 \in \nu_{\mathcal{M}}(\psi)$. Then, as $\Pi_3; \Gamma_3, \psi \models \chi; \Box$, we have $1 \in \nu_{\mathcal{M}}(\chi)$. Hence $1 \in \nu_{\mathcal{M}}(\chi)$. Therefore, $\Pi_1, \Pi_2, \Pi_3; \Gamma_1, \Gamma_2, \Gamma_3 \models \chi; \Box$.

The other cases can be proved analogously. □

A.4.2. Theorem (Soundness of ND). *Let* $\Pi, \Gamma \subseteq \mathcal{F}$ *and* $\phi \in \mathcal{F}$. *Then*

(i) $\Pi; \Gamma \mapsto \phi; \Box \implies \Pi; \Gamma \models \phi; \Box$,
(ii) $\Pi; \Gamma \mapsto \Box; \phi \implies \Pi; \Gamma \models \Box; \phi$.

Proof. The proof is by induction on the depth of derivation. All that needs to be checked is that the rules preserve truth and falsity in the appropriate way. This can be shown using Lemma A.4.1. □

A.5 Completeness

We now prove the completeness theorem for ND. [Priest, 2001] demonstrates techniques to prove completeness theorems for natural deduction systems for various relevant and paraconsistent logics. Although Priest defines validity and derivability in a standard way, his techniques provide some insights into the structure of the proof for the theorem. Here we adapt his techniques in our proof.

A.5.1. Definition. Let $\Pi, \Gamma \subseteq \mathcal{F}$. Then $\langle \Pi; \Gamma \rangle$ is a theory, if $\langle \Pi; \Gamma \rangle$ is closed under deducibility, that is, if both

(i) $\Pi; \Gamma \mapsto \phi; \Box \implies \phi \in \Pi$,
(ii) $\Pi; \Gamma \mapsto \Box; \phi \implies \phi \in \Gamma$.

A.5.2. Definition. Let $\langle \Pi; \Gamma \rangle$ be a theory. Then $\langle \Pi; \Gamma \rangle$ is dual prime, if $\langle \Pi; \Gamma \rangle$ has both the disjunction property and the conjunction property, that is, if both

(i) $\phi \vee \psi \in \Pi \implies \phi \in \Pi$ or $\psi \in \Pi$,
(ii) $\phi \wedge \psi \in \Gamma \implies \phi \in \Gamma$ or $\psi \in \Gamma$.

A.5.3. Lemma. *Let* $\Pi, \Gamma \subseteq \mathcal{F}$ *and* $\phi, \psi \in \mathcal{F}$. *Let* $\langle \Pi; \Gamma \rangle$ *be a dual prime theory. Then*

(i) $\phi \wedge \psi \in \Pi \iff \phi \in \Pi$ *and* $\psi \in \Pi$,
(ii) $\phi \wedge \psi \in \Gamma \iff \phi \in \Gamma$ *or* $\psi \in \Gamma$,
(iii) $\phi \vee \psi \in \Pi \iff \phi \in \Pi$ *or* $\psi \in \Pi$,
(iv) $\phi \vee \psi \in \Gamma \iff \phi \in \Gamma$ *and* $\psi \in \Gamma$,
(v) $\phi \in \Pi \iff \neg\phi \in \Gamma$,
(vi) $\neg\phi \in \Pi \iff \phi \in \Gamma$.

Proof.

(i) Suppose $\phi \wedge \psi \in \Pi$. Then $\Pi; \Gamma \mapsto \phi \wedge \psi; \Box$. So $\Pi; \Gamma \mapsto \phi; \Box$ and $\Pi; \Gamma \mapsto \psi; \Box$ by $\wedge E_{\vdash}$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \in \Pi$ and $\psi \in \Pi$. Suppose $\phi \in \Pi$ and $\psi \in \Pi$. By $\wedge I_{\vdash}$, $\Pi; \Gamma \mapsto \phi \wedge \psi; \Box$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \wedge \psi \in \Pi$.

(ii) Suppose $\phi \wedge \psi \in \Gamma$. By dual primeness, $\phi \in \Gamma$ or $\psi \in \Gamma$. Suppose $\phi \in \Gamma$ or $\psi \in \Gamma$. By $\wedge I_{\dashv}$, $\Pi; \Gamma \mapsto \Box; \phi \wedge \psi$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \wedge \psi \in \Gamma$.

(iii) Suppose $\phi \vee \psi \in \Pi$. By dual primeness, $\phi \in \Pi$ or $\psi \in \Pi$. Suppose $\phi \in \Pi$ or $\psi \in \Pi$. By $\vee I_{\vdash}$, $\Pi; \Gamma \mapsto \phi \vee \psi; \Box$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \vee \psi \in \Pi$.

(iv) Suppose $\phi \vee \psi \in \Gamma$. By $\vee E_{\dashv}$, $\Pi; \Gamma \mapsto \Box; \phi$ and $\Pi; \Gamma \mapsto \Box; \psi$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \in \Gamma$ and $\psi \in \Gamma$. Suppose $\phi \in \Gamma$ and $\psi \in \Gamma$. By $\vee I_{\dashv}$, $\Pi; \Gamma \mapsto \Box; \phi \vee \psi$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \vee \psi \in \Gamma$.

(v) Suppose $\phi \in \Pi$. By $\neg I_{\vdash}$, $\Pi; \Gamma \mapsto \Box; \neg\phi$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\neg\phi \in \Gamma$. Suppose $\neg\phi \in \Gamma$. By $\neg E_{\dashv}$, $\Pi; \Gamma \mapsto \phi; \Box$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \in \Pi$.

(vi) Suppose $\neg\phi \in \Pi$. By $\neg E_{\vdash}$, $\Pi; \Gamma \mapsto \Box; \phi$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\phi \in \Gamma$. Suppose $\phi \in \Gamma$. By $\neg I_{\dashv}$, $\Pi; \Gamma \mapsto \neg\phi; \Box$. Since $\langle \Pi; \Gamma \rangle$ is a theory, $\neg\phi \in \Pi$.

□

A.5.4. Definition. Let $\Pi, \Gamma, \Delta, \Sigma \subseteq \mathcal{F}$. Then

(i)	$\Pi; \Gamma \mapsto \Delta; \Box$	$\iff$	There are $\delta_1, \ldots, \delta_n \in \Delta$ such that $\Pi; \Gamma \mapsto \delta_1 \vee \cdots \vee \delta_n; \Box$,
(ii)	$\Pi; \Gamma \mapsto \Box; \Sigma$	$\iff$	There are $\sigma_1, \ldots, \sigma_n \in \Sigma$ such that $\Pi; \Gamma \mapsto \Box; \sigma_1 \wedge \cdots \wedge \sigma_n$.

A.5.5. Lemma. *Let* $\Pi, \Gamma, \Delta \subseteq \mathcal{F}$ *such that* $\Pi; \Gamma \not\mapsto \Delta; \Box$. *Then there are sets* $\Pi^* \supseteq \Pi$, $\Gamma^* \supseteq \Gamma$ *and* $\Delta^* \supseteq \Delta$ *such that*

(i) $\Pi^*; \Gamma^* \not\mapsto \Delta^*; \Box$,
(ii) $\langle \Pi^*; \Gamma^* \rangle$ *is a theory,*
(iii) $\langle \Pi^*; \Gamma^* \rangle$ *is dual prime.*

Proof. Assume that $\Pi; \Gamma \not\mapsto \Delta; \Box$ for $\Pi, \Gamma, \Delta \subseteq \mathcal{F}$. Let $\chi_0, \chi_2, \chi_4, \ldots$ be an enumeration of $\mathcal{F}$. Let $m \in \{0, 2, 4, \ldots\}$. We define by recursion the sequence $\langle \Pi_n; \Gamma_n; \Delta_n \rangle$ $(n \in \omega)$ as follows:

$$\langle \Pi_0; \Gamma_0; \Delta_0 \rangle \quad := \quad \langle \Pi; \Gamma; \Delta \rangle$$

$$\langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1} \rangle \quad := \begin{cases} \langle \Pi_m, \chi_m; \Gamma_m; \Delta_m \rangle, & \textit{if} \\ \quad \Pi_m, \chi_m; \Gamma_m \not\mapsto \Delta_m; \Box \\ \langle \Pi_m; \Gamma_m; \Delta_m, \chi_m \rangle, & \textit{if} \\ \quad \Pi_m, \chi_m; \Gamma_m \mapsto \Delta_m; \Box. \end{cases}$$

$$\langle \Pi_{m+2}; \Gamma_{m+2}; \Delta_{m+2} \rangle \quad := \begin{cases} \langle \Pi_{m+1}; \Gamma_{m+1}, \chi_m; \Delta_{m+1} \rangle, & \textit{if} \\ \quad \Pi_{m+1}; \Gamma_{m+1}, \chi_m \not\mapsto \Delta_{m+1}; \Box \\ \langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1}, \neg\chi_m \rangle, & \textit{if} \\ \quad \Pi_{m+1}; \Gamma_{m+1}, \chi_m \mapsto \Delta_{m+1}; \Box. \end{cases}$$

We define the following by means of the sequence defined above thus:

$$\langle \Pi^*; \Gamma^*; \Delta^* \rangle \quad := \quad \langle \textstyle\bigcup_{n\in\omega} \Pi_n; \bigcup_{n\in\omega} \Gamma_n; \bigcup_{n\in\omega} \Delta_n \rangle.$$

(i) We show that $\Pi^*; \Gamma^* \not\mapsto \Delta^*; \Box$ by induction on the construction of $\langle \Pi^*; \Gamma^*; \Delta^* \rangle$.

Basis: $n = 0$. Then $\Pi_0; \Gamma_0 \not\mapsto \Delta_0; \Box$ by assumption.

Induction Hypothesis: $\Pi_n; \Gamma_n \not\mapsto \Delta_n; \Box$.

Induction Step: We must show that $\Pi_{n+1}; \Gamma_{n+1} \not\mapsto \Delta_{n+1}; \Box$. There are two cases: (a) $n+1 = m+1$ for some $m \in \{0, 2, 4, \ldots\}$, and (b) $n+1 = m+2$ for some $m \in \{0, 2, 4, \ldots\}$.

(a) Suppose that $n+1 = m+1$ for some $m \in \{0, 2, 4, \ldots\}$. Then there are two cases based on the construction of $\langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1} \rangle$ from $\langle \Pi_m; \Gamma_m; \Delta_m \rangle$.

(a') $\Pi_m, \chi_m; \Gamma_m \not\mapsto \Delta_m; \Box$. Then, by construction, $\langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1} \rangle = \langle \Pi_m, \chi_m; \Gamma_m; \Delta_m \rangle$. Hence, $\Pi_{m+1}; \Gamma_{m+1} \not\mapsto \Delta_{m+1}; \Box$. Therefore, it must be that $\Pi_{n+1}; \Gamma_{n+1} \not\mapsto \Delta_{n+1}; \Box$.

(a'') $\Pi_m, \chi_m; \Gamma_m \mapsto \Delta_m; \Box$. Then, by construction, $\langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1} \rangle = \langle \Pi_m; \Gamma_m; \Delta_m, \chi_m \rangle$. Suppose $\Pi_{m+1}; \Gamma_{m+1} \mapsto \Delta_{m+1}; \Box$. Then $\Pi_m; \Gamma_m \mapsto \Delta_m, \chi_m; \Box$. By an application of $\vee E_\vdash$ and, if necessary, applications of $\vee I_\vdash$ we have that $\Pi_m; \Gamma_m \mapsto \Delta_m; \Box$, that is, $\Pi_n; \Gamma_n \mapsto \Delta_n; \Box$, contrary to the Induction Hypothesis.

(b) Suppose that $n+1 = m+2$ for some $m \in \{0, 2, 4, \ldots\}$. Then, obviously, there are two cases based on the construction of $\langle \Pi_{m+2}; \Gamma_{m+2}; \Delta_{m+2} \rangle$ from $\langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1} \rangle$.

(b') $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \not\mapsto \Delta_{m+1}; \Box$. By construction, $\langle \Pi_{m+2}; \Gamma_{m+2}; \Delta_{m+2} \rangle = \langle \Pi_{m+1}; \Gamma_{m+1}, \chi_m; \Delta_{m+1} \rangle$. Thus, $\Pi_{m+2}; \Gamma_{m+2} \not\mapsto \Delta_{m+2}; \Box$. Therefore, it must be that $\Pi_{n+1}; \Gamma_{n+1} \not\mapsto \Delta_{n+1}; \Box$.

(b'') $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \mapsto \Delta_{m+1}; \Box$. By construction, $\langle \Pi_{m+2}; \Gamma_{m+2}; \Delta_{m+2} \rangle = \langle \Pi_{m+1}; \Gamma_{m+1}; \Delta_{m+1}, \neg\chi_m \rangle$. Suppose $\Pi_{m+2}; \Gamma_{m+2} \mapsto \Delta_{m+2}; \Box$. Then it holds that $\Pi_{m+1}; \Gamma_{m+1} \mapsto \Delta_{m+1}, \neg\chi_m; \Box$. By applications of $\neg E_\vdash$, $\vee E_\vdash$, and, if necessary, applications of $\vee I_{\vdash}$, it must be that $\Pi_{m+1}; \Gamma_{m+1} \mapsto \Delta_{m+1}; \Box$. Therefore, $\Pi_n; \Gamma_n \mapsto \Delta_n; \Box$, contrary to the Induction Hypothesis.

By (*a*) and (*b*), $\Pi_{n+1}; \Gamma_{n+1} \not\mapsto \Delta_{n+1}; \Box$. Hence $\Pi_n; \Gamma_n \not\mapsto \Delta_n; \Box$ for all n by induction. Therefore, $\Pi^*; \Gamma^* \not\mapsto \Delta^*; \Box$.

(ii) We show that $\langle \Pi^*; \Gamma^* \rangle$ is a theory. Assume that $\Pi^*; \Gamma^* \mapsto \phi; \Box$. Now suppose that $\phi \notin \Pi^*$. Then by the construction, for some $m \in \{0, 2, 4, \ldots\}$ where $\phi = \chi_m$, it is not the case that $\Pi_m, \chi_m; \Gamma_m \not\mapsto \Delta_m; \Box$. Hence, it must be that $\Pi_m, \chi_m; \Gamma_m \mapsto \Delta_m; \Box$. Hence $\phi \in \Delta_{m+1} \subseteq \Delta^*$. Thus $\Pi^*; \Gamma^* \mapsto \Delta^*; \Box$, contrary to (i) proved above.

Assume that $\Pi^*; \Gamma^* \mapsto \Box; \phi$, or equivalently, $\Pi^*; \Gamma^* \mapsto \neg\phi; \Box$ by $\neg I_{\dashv}$. Now suppose that $\phi \notin \Gamma^*$. Then by the construction, for some $m \in \{0, 2, 4, \ldots\}$ where $\phi = \chi_m$, it is not the case that $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \not\mapsto \Delta_{m+1}; \Box$. So $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \mapsto \Delta_{m+1}; \Box$. Hence $\neg\phi \in \Delta_{m+2} \subseteq \Delta^*$. Thus $\Pi^*; \Gamma^* \mapsto \Delta^*; \Box$, contrary to (i) proved above.

(iii) We show that $\langle \Pi^*; \Gamma^* \rangle$ is dual prime. Assume that $\phi \vee \psi \in \Pi^*$. Then $\Pi^*; \Gamma^* \mapsto \phi \vee \psi; \Box$. Now suppose that $\phi \notin \Pi^*$ and $\psi \notin \Pi^*$. By the construction, for some $m \in \{0, 2, 4, \ldots\}$ where $\phi = \chi_m$ and $n \in \{0, 2, 4, \ldots\}$ where $\psi = \chi_n$, it is not the case that $\Pi_m, \chi_m; \Gamma_m \not\mapsto \Delta_m; \Box$, nor that $\Pi_n, \chi_n; \Gamma_n \not\mapsto \Delta_n; \Box$. So $\Pi_m, \chi_m; \Gamma_m \mapsto \Delta_m; \Box$ and $\Pi_n, \chi_n; \Gamma_n \mapsto \Delta_n; \Box$. Hence $\phi \in \Delta_{m+1} \subseteq \Delta^*$ and $\psi \in \Delta_{n+1} \subseteq \Delta^*$. Therefore $\Pi^*; \Gamma^* \mapsto \Delta^*; \Box$, contrary to (i) proved above.

Assume that $\phi \wedge \psi \in \Gamma^*$. Then $\Pi^*; \Gamma^* \mapsto \Box; \phi \wedge \psi$, or equivalently, $\Pi^*; \Gamma^* \mapsto \neg\phi \vee \neg\psi; \Box$ by $DeM_{\dashv}$. Now suppose that $\phi \notin \Gamma^*$ and $\psi \notin \Gamma^*$. By the construction, for some $m \in \{0, 2, 4, \ldots\}$ where $\phi = \chi_m$ and $n \in \{0, 2, 4, \ldots\}$ where $\psi = \chi_n$, it is not the case that $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \not\mapsto \Delta_{m+1}; \Box$, nor that $\Pi_{n+1}; \Gamma_{n+1}, \chi_n \not\mapsto \Delta_{n+1}; \Box$. So $\Pi_{m+1}; \Gamma_{m+1}, \chi_m \mapsto \Delta_{m+1}; \Box$ and $\Pi_{n+1}; \Gamma_{n+1}, \chi_n \mapsto \Delta_{n+1}; \Box$. Hence $\neg\phi \in \Delta_{m+2} \subseteq \Delta^*$ and $\neg\psi \in \Delta_{n+2} \subseteq \Delta^*$. Therefore $\Pi^*; \Gamma^* \mapsto \Delta^*; \Box$, contrary to (i) proved above.

□

A.5.6. Lemma. *Let* $\Pi, \Gamma, \Sigma \subseteq \mathcal{F}$ *such that* $\Pi; \Gamma \not\mapsto \Box; \Sigma$. *Then there are sets* $\Pi^* \supseteq \Pi$, $\Gamma^* \supseteq \Gamma$ *and* $\Sigma^* \supseteq \Sigma$ *such that*

(i) $\Pi^*; \Gamma^* \not\mapsto \Box; \Sigma^*$,
(ii) $\langle \Pi^*; \Gamma^* \rangle$ *is a theory,*
(iii) $\langle \Pi^*; \Gamma^* \rangle$ *is dual prime.*

Proof. Assume that $\Pi; \Gamma \not\mapsto \Box; \Sigma$ for $\Pi, \Gamma, \Sigma \subseteq \mathcal{F}$. Let $\chi_0, \chi_2, \chi_4, \ldots$ be an enumeration of $\mathcal{F}$. Let $m \in \{0, 2, 4, \ldots\}$. We define by recursion the sequence $\langle \Pi_n; \Gamma_n; \Sigma_n \rangle$ $(n \in \omega)$ as follows:

$$\langle \Pi_0; \Gamma_0; \Sigma_0 \rangle \quad := \quad \langle \Pi; \Gamma; \Sigma \rangle$$

$$\langle \Pi_{m+1}; \Gamma_{m+1}; \Sigma_{m+1} \rangle \quad := \begin{cases} \langle \Pi_m, \chi_m; \Gamma_m; \Sigma_m \rangle, \quad \textit{if} \\ \qquad \Pi_m, \chi_m; \Gamma_m \not\mapsto \Box; \Sigma_m \\ \\ \langle \Pi_m; \Gamma_m; \Sigma_m, \neg\chi_m \rangle, \quad \textit{if} \\ \qquad \Pi_m, \chi_m; \Gamma_m \mapsto \Box; \Sigma_m. \end{cases}$$

$$\langle \Pi_{m+2}; \Gamma_{m+2}; \Sigma_{m+2} \rangle \quad := \begin{cases} \langle \Pi_{m+1}; \Gamma_{m+1}, \chi_m; \Sigma_{m+1} \rangle, \quad \textit{if} \\ \qquad \Pi_{m+1}; \Gamma_{m+1}, \chi_m \not\mapsto \Box; \Sigma_{m+1} \\ \\ \langle \Pi_{m+1}; \Gamma_{m+1}; \Sigma_{m+1}, \chi_m \rangle, \quad \textit{if} \\ \qquad \Pi_{m+1}; \Gamma_{m+1}, \chi_m \mapsto \Box; \Sigma_{m+1}. \end{cases}$$

Then (i), (ii) and (iii) can be proved as in Lemma A.5.5. □

A.5.7. LEMMA. *Let* $\Pi, \Gamma \subseteq \mathcal{F}$ *and* $\phi \in \mathcal{F}$. *Then*

$$\Pi; \Gamma \models \phi; \Box \quad \Longrightarrow \quad \Pi; \Gamma \mapsto \phi; \Box.$$

Proof. Suppose that $\Pi; \Gamma \not\mapsto \phi; \Box$ for $\Pi, \Gamma \subseteq \mathcal{F}$ and $\phi \in \mathcal{F}$. By applying Lemma A.5.5 with $\{\phi\}$ as Δ, there is a dual prime theory $\langle \Pi^*; \Gamma^* \rangle$ for $\Pi^* \supseteq \Pi$ and $\Gamma^* \supseteq \Gamma$ and $\Delta^* \supseteq \Delta$, such that $\Pi^*; \Gamma^* \not\mapsto \Delta^*; \Box$.

Let $\mathcal{M}$ $(= \langle \mathcal{F}, \nu \rangle)$ be an interpretation and $p \in \mathcal{A}$. We define an evaluation ν as:

$$\begin{aligned} 1 \in \nu_{\mathcal{M}}(p) &\iff p \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(p) &\iff p \in \Gamma^*. \end{aligned}$$

It is then asserted that the above conditions extend to all formulas:

$$\begin{aligned} 1 \in \nu_{\mathcal{M}}(\phi) &\iff \phi \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\phi) &\iff \phi \in \Gamma^*. \end{aligned}$$

The assertion is proved by structural induction on ϕ.

Basis: By assumption, it holds for $\phi \in \mathcal{A}$:

$$\begin{aligned} 1 \in \nu_{\mathcal{M}}(\phi) &\iff \phi \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\phi) &\iff \phi \in \Gamma^*. \end{aligned}$$

Induction Hypothesis: For all ψ with fewer logical operators than ϕ:

$$\begin{aligned} 1 \in \nu_{\mathcal{M}}(\psi) &\iff \psi \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\psi) &\iff \psi \in \Gamma^*. \end{aligned}$$

Induction Step: There are six cases based on the connectives in ϕ.

$$\begin{aligned} 1 \in \nu_{\mathcal{M}}(\psi_1 \wedge \psi_2) &\iff 1 \in \nu_{\mathcal{M}}(\psi_1) \text{ and } 1 \in \nu_{\mathcal{M}}(\psi_2) && \text{by Definition A.2.3} \\ &\iff \psi_1 \in \Pi^* \text{ and } \psi_2 \in \Pi^* && \text{by Induction Hypothesis} \\ &\iff \psi_1 \wedge \psi_2 \in \Pi^* && \text{by Lemma A.5.3,} \end{aligned}$$

$$\begin{aligned} 0 \in \nu_{\mathcal{M}}(\psi_1 \wedge \psi_2) &\iff 0 \in \nu_{\mathcal{M}}(\psi_1) \text{ or } 0 \in \nu_{\mathcal{M}}(\psi_2) && \text{by Definition A.2.3} \\ &\iff \psi_1 \in \Gamma^* \text{ or } \psi_2 \in \Gamma^* && \text{by Induction Hypothesis} \\ &\iff \psi_1 \wedge \psi_2 \in \Gamma^* && \text{by Lemma A.5.3.} \end{aligned}$$

Similarly, we have

$$\begin{array}{lcl} 1 \in \nu_{\mathcal{M}}(\psi_1 \vee \psi_2) & \iff & \psi_1 \vee \psi_2 \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\psi_1 \vee \psi_2) & \iff & \psi_1 \vee \psi_2 \in \Gamma^*, \\ 1 \in \nu_{\mathcal{M}}(\neg\psi) & \iff & \neg\psi \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\neg\psi) & \iff & \neg\psi \in \Gamma^*. \end{array}$$

Hence the evaluation conditions defined above hold for all formulas by induction. Since $\Pi^*; \Gamma^* \not\mapsto \Delta^*; \square$, we have that $\phi \notin \Pi^*$. By the above conditions then, $1 \notin \nu_{\mathcal{M}}(\phi)$. But $1 \in \nu_{\mathcal{M}}(\psi)$ and $0 \in \nu_{\mathcal{M}}(\chi)$ for all $\psi \in \Pi^*$ and $\chi \in \Gamma^*$. Hence $\Pi^*; \Gamma^* \not\models \phi; \square$. Therefore, $\Pi; \Gamma \not\models \phi; \square$. □

A.5.8. Lemma. *Let* $\Pi, \Gamma \subseteq \mathcal{F}$ *and* $\phi \in \mathcal{F}$. *Then*

$$\Pi; \Gamma \models \square; \phi \implies \Pi; \Gamma \mapsto \square; \phi.$$

Proof. Suppose that $\Pi; \Gamma \not\mapsto \square; \phi$ for $\Pi, \Gamma \subseteq \mathcal{F}$ and $\phi \in \mathcal{F}$. By applying Lemma A.5.6 with $\{\phi\}$ as Σ, there is a dual prime theory $\langle \Pi^*; \Gamma^* \rangle$ for $\Pi^* \supseteq \Pi$ and $\Gamma^* \supseteq \Gamma$ and $\Sigma^* \supseteq \Sigma$, such that $\Pi^*; \Gamma^* \not\mapsto \square; \Sigma^*$.

Let $\mathcal{M}$ $(= \langle \mathcal{F}, \nu \rangle)$ be an interpretation and $p \in \mathcal{A}$. We define an evaluation ν as:

$$\begin{array}{lcl} 1 \in \nu_{\mathcal{M}}(p) & \iff & p \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(p) & \iff & p \in \Gamma^*. \end{array}$$

It is then asserted that the above conditions extend to all formulas:

$$\begin{array}{lcl} 1 \in \nu_{\mathcal{M}}(\phi) & \iff & \phi \in \Pi^*, \\ 0 \in \nu_{\mathcal{M}}(\phi) & \iff & \phi \in \Gamma^*. \end{array}$$

This assertion is proved as in Lemma A.5.7. Since $\Pi^*; \Gamma^* \not\mapsto \square; \Sigma^*$, we have that $\phi \notin \Gamma^*$. By the above conditions, then, $0 \notin \nu_{\mathcal{M}}(\phi)$. But $1 \in \nu_{\mathcal{M}}(\psi)$ and $0 \in \nu_{\mathcal{M}}(\chi)$ for all $\psi \in \Pi^*$ and $\chi \in \Gamma^*$. Hence $\Pi^*; \Gamma^* \not\models \square; \phi$. Therefore, $\Pi; \Gamma \not\models \square; \phi$. □

A.5.9. Theorem (Completeness of ND). *Let* $\Pi, \Gamma \subseteq \mathcal{F}$ *and* $\phi \in \mathcal{F}$. *Then*

(i) $\Pi; \Gamma \models \phi; \square \iff \Pi; \Gamma \mapsto \phi; \square$,
(ii) $\Pi; \Gamma \models \square; \phi \iff \Pi; \Gamma \mapsto \square; \phi$.

Proof. Directly from Theorem A.4.2, Lemma A.5.7 and Lemma A.5.8. □

A.6 Rejection Eliminated?

Although the system ND captures the underlying idea of Dunn's semantics, one might argue that the introduction of rejected formulas is theoretically redundant. The argument runs as follows. ND takes $\dashv$ as a falsity operator understood semantically. So stating $\dashv \phi$ amounts to stating that ϕ is false. But then $\dashv \phi$ is just $\vdash \neg\phi$. Hence '$\dashv$' may be replaced by '$\vdash \neg$'. Once we have adopted this convention, '$\vdash$' can be dropped from the system, since we need not indicate the status (asserted or rejected) of a formula anymore. For example, the rule $\neg I_{\vdash}$ becomes

$$\dfrac{\begin{matrix}\mathcal{D}\\ \phi\end{matrix}}{\neg\neg\phi}\ \neg\neg I$$

and $\wedge E_{\dashv}^{u,v}$ becomes

$$\dfrac{\begin{matrix}\\ \mathcal{D}_1\\ \neg(\phi\wedge\psi)\end{matrix}\quad\begin{matrix}[\neg\phi]^u\\ \mathcal{D}_2\\ \chi\end{matrix}\quad\begin{matrix}[\neg\psi]^v\\ \mathcal{D}_3\\ \chi\end{matrix}}{\chi}\ \neg\wedge E^{u,v}$$

Moreover, if we add the De Morgan rules as primitive, there will be some rules of inference which are redundant. For example, $\neg\wedge E^{u,v}$ in the new system will be a special case of $\vee E^{u,v}$. The resulting system will then be that of [Priest, 2001], as can easily be checked.[12]

These changes give rise to changes to the definitions of validity and derivability as well. Since every (rejected) formula in Γ in our definition of validity, that is, Definition A.2.5, can be incorporated into Π by placing '$\neg$' in front of the formulas under consideration, validity is defined standardly. Similarly, derivability is defined standardly. Then soundness and completeness can be established as in [Priest, 2001].

The fact that ND collapses under the proposed substitution to a standard system, such as Priest's, however, does not imply the inferiority of the system presented in this Appendix, as there are some obvious advantages of our combined system over the standard ones. First, ND visually reflects the underlying idea of Dunn's semantics: truth and falsity are evaluated separately. Second, because of the introduction of both asserted and rejected formulas in our proof system, our system, contrary to Priest's, does not have any rules for combinations of logical operators: each operator has two introduction rules and two elimination rules, according to the status (asserted or rejected) of the formula which serves as a premise in the application of a rule. Rules which necessitate combinations of operators obscure the meanings of the operators. In constructing a proof tree in our system, at each step only the principal operator needs to be considered. This procedure makes the construction of proofs intuitive and mechanical, which is the main purpose of formal logics.

Third, ND has conjunction elimination rules which have the same forms as disjunction elimination rules. Standardly, the disjunction elimination rule includes subproof trees, while the conjunction elimination rule does not. So they have different forms. In ND, the conjunction elimination rule, $\wedge E_{\vdash}$, has the same form as the disjunction elimination rule, $\vee E_{\dashv}$, and $\wedge E_{\dashv}$ does the same as $\vee E_{\vdash}$. Thus the elimination rules for conjunction and disjunction are dual. This feature of the system, therefore, provides symmetric proofs which capture the semantics in a natural way without any technical complications.

[12] [Smullyan, 1968] shows a similar result for a classical tableaux system.

Finally, our definition of validity may be extended to capture more general consequence relations as follows:

A.6.1. Definition. Let $\Pi, \Gamma, \Sigma, \Delta \subseteq \mathcal{F}$ and $\phi \in \mathcal{F}$. Then

(i) $\Pi; \Gamma; \Sigma; \Delta \models \phi; \square; \square; \square \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$ and $1 \notin \nu_{\mathcal{M}}(\Sigma)$ and $0 \notin \nu_{\mathcal{M}}(\Delta)$, then $1 \in \nu_{\mathcal{M}}(\phi)$,

(ii) $\Pi; \Gamma; \Sigma; \Delta \models \square; \phi; \square; \square \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$ and $1 \notin \nu_{\mathcal{M}}(\Sigma)$ and $0 \notin \nu_{\mathcal{M}}(\Delta)$, then $0 \in \nu_{\mathcal{M}}(\phi)$,

(iii) $\Pi; \Gamma; \Sigma; \Delta \models \square; \square; \phi; \square \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$ and $1 \notin \nu_{\mathcal{M}}(\Sigma)$ and $0 \notin \nu_{\mathcal{M}}(\Delta)$, then $1 \notin \nu_{\mathcal{M}}(\phi)$,

(iv) $\Pi; \Gamma; \Sigma; \Delta \models \square; \square; \square; \phi \iff$ For all $\mathcal{M}$: if $1 \in \nu_{\mathcal{M}}(\Pi)$ and $0 \in \nu_{\mathcal{M}}(\Gamma)$ and $1 \notin \nu_{\mathcal{M}}(\Sigma)$ and $0 \notin \nu_{\mathcal{M}}(\Delta)$, then $0 \notin \nu_{\mathcal{M}}(\phi)$.

Proof-theoretical characterizations of the above consequence relations have yet to be investigated. However, it does not seem impossible to give a proof theory in the style of [Konikowska, 1990]. Moreover, these general consequence relations may be studied in the context of many logics other than fde as well.

Bibliography

[Ackermann, 1956] Wilhelm Ackermann. Begründung einer strengen Implikation, *Journal of Symbolic Logic*, 21: 113–128, 1956.

[Alchourrón, Gärdenfors, and Makinson, 1985] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions, *Journal of Symbolic Logic*, 50: 510–530, 1985.

[Alchourrón and Makinson, 1982] Carlos E. Alchourrón and David Makinson. On the Logic of Theory Change: Contraction Functions and Their Associated Revision Functions, *Theoria*, 48: 14–37, 1982.

[Alchourrón and Makinson, 1985] Carlos E. Alchourrón and David Makinson. On the Logic of Theory Change: Safe Contraction, *Studia Logica*, 44: 405–422, 1985.

[Anderson and Belnap, 1962] Alan Ross Anderson and Nuel D. Belnap Jr. Tautological Entailments, *Philosophical Studies*, 13: 9–24, 1962.

[Anderson and Belnap, 1975] Alan Ross Anderson and Nuel D. Belnap Jr. *Entailment: The Logic of Relevance and Necessity*, volume 1, Princeton University Press, Princeton 1975.

[Anderson, Belnap, and Dunn, 1992] Alan Ross Anderson, Nuel D. Belnap Jr., and J. Michael Dunn. *Entailment: The Logic of Relevance and Necessity*, volume 2, Princeton University Press, Princeton 1992.

[Barth, 1987] Else M. Barth. *Empiristische en empirische logica*, Noord-Hollandse Uitgevers Maatschappij, Amsterdam 1987.

[Batens, 1980] Diderik Batens. Paraconsistent Extensional Propositional Logics, *Logique & Analyse*, 90–91: 195–234, 1980.

[Batens, 1986] Diderik Batens. Dialectical Dynamics within Formal Logics, *Logique & Analyse*, 114: 161–173, 1986.

[Batens, 1989] Diderik Batens. Dynamic Dialectical Logics, in Priest, Routley, and Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 187–217, Philosophia Verlag, Munich 1989.

[Batens, 1992] Diderik Batens. *Menselijke kennis*, Garant, Leuven 1992.

[Batens, forthcoming] Diderik Batens. In Defence of a Programme for Handling Inconsistencies, in Meheus, editor, *Inconsistency in Science*, Kluwer, Dordrecht, forthcoming.

[Batens, manuscript] Diderik Batens. Towards the Unification of Inconsistency Handling Mechanisms, manuscript.

[Belnap, 1977] Nuel Belnap. A Useful Four-Valued Logic, in Dunn and Epstein, *Modern Uses of Multiple-Valued Logic*, pages 8–37, Reidel Publishing Company, Dordrecht 1977.

[Benferhat, Dubois, and Prade, 1997] Salem Benferhat, Didier Dubois, and Henri Prade. Some Syntactic Approaches to the Handling of Inconsistent Knowledge Bases: A Comparative Study. Part 1: The Flat Case, *Studia Logica*, 58: 17–45, 1997.

[Beth, 1950] Evert Willem Beth. *Wijsgerige ruimteleer*, Standaard-Boekhandel, Antwerp 1950.

[Beth, 1959] Evert Willem Beth. *The Foundations of Mathematics*, 2d edition, North-Holland Publishing Company, Amsterdam 1965.

[Beth, 1964] Evert Willem Beth. *Door wetenschap tot wijsheid*, Van Gorcum, Assen 1964.

[Brentano, *c.*1887] Franz Brentano. *Die Lehre vom richtigen Urteil*, Francke Verlag, Bern 1956.

[Burke, 1994] Tom Burke. *Dewey's New Logic. A Reply to Russell*, The University of Chicago Press, Chicago 1994.

[Cantwell, 1998] John Cantwell. Resolving Conflicting Information, *Journal of Logic, Language and Information*, 7: 191–220, 1998.

[Dunn, 1976] J. Michael Dunn. Intuitive Semantics for First-Degree Entailments and 'Coupled Trees', *Philosophical Studies*, 29: 149–168, 1976.

[Fermé and Hansson, 1999] Eduardo L. Fermé and Sven Ove Hansson. Selective Revision, *Studia Logica*, 63: 331–342, 1999.

[Van Fraassen, 1966] Bas C. van Fraassen. Singular Terms, Truthvalue Gaps, and Free Logic, in Lambert, editor, *Philosophical Applications of Free Logic*, pages 82–97, Oxford University Press, New York 1991.

[Friedman and Halpern, 1999] Nir Friedman and Joseph Y. Halpern. Belief Revision: A Critique, *Journal of Logic, Language, and Information*, 8: 401–420, 1999.

[Fuhrmann, 1991] André Fuhrmann. Theory Contraction Through Base Contraction, *Journal of Philosophical Logic*, 20: 175–203, 1991.

[Gärdenfors, 1988] Peter Gärdenfors. *Knowledge in Flux*, The MIT Press, Cambridge (Mass.) 1988.

[Gärdenfors and Makinson, 1988] Peter Gärdenfors and David Makinson. Revisions of Knowledge Systems Using Epistemic Entrenchment, in Vardi, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 83–95, Morgan Kaufmann, Los Altos (California), 1988.

[Hansson, 1993] Sven Ove Hansson. Reversing the Levi Identity, *Journal of Philosophical Logic*, 22: 637–669, 1993.

[Hansson, 1998] Sven Ove Hansson. Editorial: Belief Revision Theory Today, *Journal of Logic, Language, and Information*, 7: 123–126, 1998.

[Hansson, 1999a] Sven Ove Hansson. Recovery and Epistemic Residue, *Journal of Logic, Language, and Information*, 8: 421–428, 1999.

[Hansson, 1999b] Sven Ove Hansson. A Survey of Non-Prioritized Belief Revision, *Erkenntnis*, 50: 413–427, 1999.

[Hansson, 1999c] Sven Ove Hansson. *A Textbook of Belief Dynamics*, Kluwer, Dordrecht 1999.

[Hansson and Olsson, 1995] Sven Ove Hansson and Erik J. Olsson. Levi Contractions and AGM Contractions: A Comparison, *Notre Dame Journal of Formal Logic*, 36: 103–119, 1995.

[Hilpinen, 1982] Risto Hilpinen. On C.S. Peirce's Theory of the Proposition: Peirce as a Precursor of Game-Theoretical Semantics, *The Monist*, 65: 182–188, 1982.

[Hookway, 1985] Christopher Hookway. *Peirce*, Routledge & Kegan Paul, London 1985.

[James, 1907] William James. *Pragmatism*, Harvard University Press, Cambridge (Mass.) 1975.

[Jardine, 1986] Nicholas Jardine. *The Fortunes of Inquiry*, Clarendon Press, Oxford 1986.

[Keil and Schnädelbach, 2000] Geert Keil and Herbert Schnädelbach. *Naturalismus*, Suhrkamp, Frankfurt am Main 2000.

[Kleene, 1952] Stephen Cole Kleene. *Introduction to Metamathematics*, Wolters-Noordhoff, Groningen 1952.

[Kneale, 1949] William Kneale. *Probability and Induction*, Clarendon Press, Oxford 1966.

[Kneale and Kneale, 1962] William Kneale and Martha Kneale. *The Development of Logic*, Clarendon Press, Oxford 1971.

[Konikowska, 1990] Beata Konikowska. A Two-Valued Logic for Reasoning about Different Types of Consequence in Kleene's Three-Valued Logic, *Studia Logica*, 49: 541–555, 1990.

[Kratzer, 1981] Angelika Kratzer. Partition and Revision: The Semantics of Counterfactuals, *Journal of Philosophical Logic* 10: 201–216, 1981.

[Kraus, Lehmann, and Magidor, 1990] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics, *Artificial Intelligence*, 44: 167–207, 1990.

[Kripke, 1975] Saul Kripke. Outline of a Theory of Truth, *Journal of Philosophy*, 72: 690–716, 1975.

[Kuklick, 1977] Bruce Kuklick. *The Rise of American Philosophy*, Yale University Press, New Haven 1977.

[Levi, 1967a] Isaac Levi. *Gambling with Truth*, Alfred A. Knopf, New York 1967.

[Levi, 1967b] Isaac Levi. Information and Inference, *Synthese*, 17: 369–391, 1967.

[Levi, 1976] Isaac Levi. Acceptance Revisited, in Bogdan, editor, *Local Induction*, pages 1–71, Reidel, Dordrecht 1976.

[Levi, 1980] Isaac Levi. *The Enterprise of Knowledge*, The MIT Press, Cambridge (Mass.) 1980.

[Levi, 1991] Isaac Levi. *The Fixation of Belief and Its Undoing*, Cambridge University Press, Cambridge 1991.

[Levi, 1996] Isaac Levi. *For the Sake of the Argument*, Cambridge University Press, Cambridge 1996.

[Levi, 1998] Isaac Levi. Contraction and Informational Value, 7th version, manuscript.

[Lewis 1973] David Lewis. *Counterfactuals*, Basil Blackwell, Oxford 1973.

[Łukasiewicz, 1921] Jan Łukasiewicz. Two-Valued Logic, in Borkowski, editor, *Jan Łukasiewicz. Selected Works*, pages 89–109, North-Holland Publishing Company, Amsterdam 1970.

[Łukasiewicz, 1951] Jan Łukasiewicz. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, 2d edition, Clarendon Press, Oxford 1957.

[Mach, 1905] Ernst Mach. *Erkenntnis & Irrtum*, Verlag von Johann Ambrosius Barth, Leipzig 1906.

[Makinson, 1987] David Makinson. On the Status of the Postulate of Recovery in the Logic of Theory Change, *Journal of Philosophical Logic*, 16: 383–394, 1987.

[Malcolm, 1952] Norman Malcolm. Knowledge and Belief, in Phillips Griffiths, editor, *Knowledge and Belief*, pages 69–81, Oxford University Press, Oxford 1967.

[Mares, 1998] Edward D. Mares. Who's Afraid of Impossible Worlds?, *Notre Dame Journal of Formal Logic*, 38: 516–526, 1997.

[Misak, 1991] Cheryl J. Misak. *Truth and the End of Inquiry: A Peircean Account of Truth*, Clarendon Press, Oxford 1991.

[Misak, 1998] Cheryl J. Misak. Deflating Truth: Relativism *vs.* Minimalism, *The Monist*, 81: 407–425, 1998.

[Muller, 1913] Tobias B. Muller. *De kennisleer van het Anglo-Amerikaansch pragmatisme*, H.P. de Swart & Zoon, The Hague 1913.

[Nagel, 1982] Ernest Nagel. Charles Peirce's Place in Philosophy, *Historia Mathematica*, 9: 302–310, 1982.

[Pais and Jackson, 1992] John Pais and Peter Jackson. Partial Monotonicity and a New Version of the Ramsey Test, *Studia Logica*, 51: 21–47, 1992.

[Peirce, *CP*] Charles Sanders Peirce. *Collected Papers of Charles Sanders Peirce*, Hartshorne and Weiss, editors of volumes 1–6, Burks, editor of volumes 7–8, Belknap Press, Cambridge (Mass.) 1931–1958.

[Peirce, *CW*] Charles Sanders Peirce. *Writings of Charles S. Peirce. A Chronological Edition*, Fisch, editor, Indiana University Press, Bloomington 1982–1993.

[Peirce, *NE*] Charles Sanders Peirce. *The New Elements of Mathematics*, Eisele, editor, Mouton, The Hague 1976.

[Priest, 1979] Graham Priest. The Logic of Paradox, *Journal of Philosophical Logic*, 8: 219–241, 1979.

[Priest, 2001] Graham Priest. Paraconsistent Logic, in Gabbay and Guenthner, editors, *Handbook of Philosophical Logic*, 2d edition, volume 8, pages 287–393, Kluwer Academic Publishers, Dordrecht 2001.

[Priest and Sylvan, 1992] Graham Priest and Richard Sylvan. Simplified Semantics for Basic Relevant Logics, *Journal of Philosophical Logic*, 21: 217–232, 1992.

[Putnam, 1981] Hilary Putnam. *Reason, Truth and History*, Cambridge University Press, Cambridge 1981.

[Putnam, 1982] Hilary Putnam. Peirce the Logician, in Hilary Putnam, *Realism with a Human Face*, pages 252–260, Harvard University Press, Cambridge (Mass.) 1990.

[Putnam, 1983] Hilary Putnam. *Realism and Reason: Philosophical Papers, Volume 3*, Cambridge University Press, Cambridge 1983.

[Putnam, 1987] Hilary Putnam. *The Many Faces of Realism*, Open Court, La Salle (Ill.) 1987.

[Putnam, 1990] Hilary Putnam. *Realism with a Human Face*, Harvard University Press, Cambridge (Mass.) 1990.

[Quine, 1953] Willard Van Orman Quine. *From a Logical Point of View*, Harper & Row, New York 1963.

[Quine, 1960] Willard Van Orman Quine. *Word and Object*, The MIT Press, Cambridge (Mass.) 1960.

[Quine, 1969] Willard Van Orman Quine. *Ontological Relativity and Other Essays*, Columbia University Press, New York 1969.

[Quine, 1981] Willard Van Orman Quine. *Theories and Things*, Belknap Press, Cambridge (Mass.) 1994.

[Ramsey, 1929] Frank P. Ramsey. General Propositions and Causality, in Ramsey, *Foundations. Essays in Philosophy, Logic, Mathematics and Economics*, pages 133–151, Humanities Press, Atlantic Highlands (N.J.) 1978.

[Reichenbach 1948] Hans Reichenbach. Rationalism and Empiricism. An Enquiry into the Roots of Philosophical Error, in Reichenbach, *Modern Philosophy of Science*, pages 135–150, Routledge & Kegan Paul, London 1959.

[Restall and Slaney, 1995] Greg Restall and John Slaney. Realistic Belief Revision, Technical Report TR-ARP-2-95, Research School of Information Sciences and Engineering and Centre for Information Science Research, Australian National University, Canberra, Australia, 1995.

[Rescher, 1964] Nicholas Rescher. *Hypothetical Reasoning*, North-Holland Publishing Company, Amsterdam 1964.

[Reynolds, 2000] Andrew Reynolds. Statistical Method and the Peircean Account of Truth, *Canadian Journal of Philosophy*, 30: 287–314, 2000.

[De Rijk, 1988] L.M. de Rijk. *Waarom? – daarom. De lotgevallen van de aristotelische bewijstheorie in de laat-middeleeuwse metafysica*, Rijksuniversiteit Leiden, Leyden 1988.

[Rott, 1989] Hans Rott. Conditionals and Theory Change: Revisions, Expansions, and Additions, *Synthese*, 81: 91–113, 1989.

[Rott, 1991] Hans Rott. Two Methods of Constructing Contractions and Revisions of Knowledge Systems, *Journal of Philosophical Logic*, 20: 149–173, 1991.

[Rott, 1992] Hans Rott. On the Logic of Theory Change: More Maps between Different Kinds of Contraction Function, in Gärdenfors, editor, *Belief Revision*, pages 122–141, Cambridge University Press, Cambridge 1992.

[Rott, 1999] Hans Rott. Moody Conditionals: Hamburgers, Switches, and the Tragic Death of an American President, in Gerbrandy, Marx, de Rijke, and Venema, editors, *JFAK. Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*, Vossiuspers AUP, Amsterdam 1999.

[Rott and Pagnucco, 1999] Hans Rott and Maurice Pagnucco. Severe Withdrawal (and Recovery), *Journal of Philosophical Logic*, 28: 501–547, 1999.

[Routley and Routley, 1972] Richard Routley and Val Routley. The Semantics of First Degree Entailment, *Noûs*, 6: 335–359, 1972.

[Russell, 1917] Bertrand Russell. *Mysticism and Logic*, George Allen & Unwin, London 1949.

[Smullyan, 1968] Raymond M. Smullyan. *First-Order Logic*, Springer-Verlag, Berlin 1968.

[Stalnaker, 1968] Robert Stalnaker. A Theory of Conditionals, in Rescher, editor, *Studies in Logical Theory*, pages 98–112, Basil Blackwell, Oxford 1968.

[Stalnaker, 1992] Robert Stalnaker. Notes on Conditional Semantics, in Moses, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 316–327, Morgan Kaufmann, San Mateo (California) 1992.

[Tamminga, 1994] Allard Tamminga. Logics of Rejection: Two Systems of Natural Deduction, *Logique & Analyse*, 146: 169–208, 1994 (appeared in 1996).

[Tamminga and Tanaka, 1999] Allard Tamminga and Koji Tanaka. A Natural Deduction System for First Degree Entailment, *Notre Dame Journal of Formal Logic*, 40: 258–272, 1999 (appeared in 2001).

[Tichý, 1976] Pavel Tichý. A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals, *Philosophical Studies*, 29: 271–273, 1976.

[Troelstra and Schwichtenberg, 1996] Anne S. Troelstra and Helmut Schwichtenberg. *Basic Proof Theory*, Cambridge University Press, Cambridge 1996.

[Veltman, 1976] Frank Veltman. Prejudices, Presuppositions and the Theory of Counterfactuals, in Groenendijk and Stokhof, editors, *Amsterdam Papers in Formal Grammar*, volume 1, pages 248–281, Centrale Interfaculteit, Universiteit van Amsterdam, Amsterdam 1976.

[Veltman, 1996] Frank Veltman. Defaults in Update Semantics, *Journal of Philosophical Logic*, 25: 221–261, 1996.

[Wasserman, 2000] Renata Wassermann. *Resource-Bounded Belief Revision*, Ph.D. thesis, Universiteit van Amsterdam, ILLC Dissertation Series 2000-01, Amsterdam 2000.

[Wright, 1992] Crispin Wright. *Truth and Objectivity*, Harvard University Press, Cambridge (Mass.) 1992.

Samenvatting

Overzicht

In dit proefschrift leg ik verslag van een studie naar redelijke veranderingen van mening. In drie samenhangende, maar onafhankelijke hoofdstukken verken ik een veelbelovend alternatief voor gangbare typen van kennistheorie, waarin men vooral beoogt grondslagen voor onze kennis te leveren. De door de traditionele epistemologie voorgeschreven onderzoeksagenda, waarop de jacht op voorwaarden waaronder onze kennis gerechtvaardigd of zelfs waar is, bovenaan staat, wordt niet door mij onderschreven. In plaats daarvan volg ik in de onderhavige studie het door Amerikaanse pragmatisten als Charles Sanders Peirce en John Dewey gebaande pad en richt ik me in de eerste plaats op de articulatie en verdediging van criteria volgens welke een verandering van mening tevens een verbetering is.

Het werk van twee pragmatistische filosofen, Charles Sanders Peirce en Isaac Levi, komt in het tweede en het derde hoofdstuk van dit proefschrift uitgebreid aan bod, opdat we goed grip krijgen op hun pogingen een model op te stellen dat beoogt optimale strategieën voor een verandering van mening te formuleren. Dit *belief-doubt-belief*-model is de verbindende schakel tussen de hoofdstukken van het proefschrift.

Tijdens de behandeling van het *belief-doubt-belief*-model blijkt een groot aantal logische overwegingen een rol te spelen, zodat een technisch-logische beschouwing van formele systemen voor *belief change* onontkomelijk wordt. We komen tot de slotsom dat de pragmatistische herkomst van de hedendaagse systemen voor *belief change*, hun naturalistische oorsprong ten spijt, in de loop van hun ontwikkeling steeds meer in de verdrukking is geraakt en veld heeft moeten ruimen voor een benadering van logisch-filosofisch onderzoek waarin voor empirische overwegingen nauwelijks plaats meer is. Met een pleidooi voor de eliminatie van elementen in systemen voor *belief change* die de praktische toepassing en empirische toetsing van deze systemen verhinderen, tracht ik het verband tussen logische theorievorming en cognitieve praxis te herstellen.

In het vierde hoofdstuk wordt een nieuw systeem voor *belief change* aangeboden dat (1) epistemische toestanden met eindige middelen representeert, (2) inconsistente informatie aankan, (3) eindige veranderingsoperaties kent, (4) buitenlogische elementen niet behoeft, maar ook niet uitsluit, en (5) altijd aanleiding geeft tot een consistente verzameling opvattingen. Een Appendix, waarin de onderliggende logica van het systeem voor *belief change* vanuit een bewijstheoretisch oogpunt wordt bestudeerd, vormt het sluitstuk van het proefschrift.

Samenvatting

Zinnige logische en kentheoretische onderzoekingen binnen de traditie van het wijsgerige pragmatisme worden bemoeilijkt doordat het leeuwendeel der pragmatistische filosofen (en niet alleen zij) voornamelijk belang stelt in oplossingen voor traditioneel-filosofische vraagstukken over wat waar en werkelijk is. Dergelijke pogingen noodzakelijke en voldoende voorwaarden voor de waarheid van een uitspraak te vatten of, anders, een werkelijkheidsbegrip te verdedigen dat beantwoordt aan een reeds verkozen invulling van het waarheidspredikaat, hebben wijsgerig onderzoek te lang beperkt tot vragen die zich reeds eeuwen steeds opnieuw aan voorgestelde antwoorden hebben ontworsteld. In het openingshoofdstuk stel ik derhalve voor deze metafysische voetangels en klemmen op te ruimen om zo vruchtbaarder onderzoekingen binnen het pragmatisme ruim baan te bieden.

Met dit doel voor ogen maak ik in het tweede hoofdstuk aannemelijk dat Peirce' metafysische streven naar definities van waarheid en werkelijkheid in termen van een theorie van wetenschappelijk onderzoek binnen een pragmatistische kenleer niet op coherente wijze kan worden bereikt. Om vat te krijgen op Peirce' definitie van waarheid als de ideële limiet van wetenschappelijk onderzoek, opent het hoofdstuk met een uiteenzetting van Peirce' onderzoekstheorie en zijn filosofische logica. Deductieve, inductieve en abductieve redeneervormen en hun onderlinge relaties komen aan bod, opdat Peirce' vroege opvattingen omtrent het pragmatistische *belief-doubt-belief*-model geschetst kunnen worden. Met dit model beoogde Peirce de 'wetenschappelijke methode' te formuleren, een methode die ons in staat stelt ons huidige verwachtingspatroon in het geval van een onverwachte of zelfs strijdige ervaring aan te passen en te verbeteren. Peirce hield staande dat indien we maar consciëntieus de wetenschappelijke methode zouden blijven volgen, we uiteindelijk een verwachtingspatroon zouden bereiken dat nooit door een toekomstige ervaring zou kunnen worden weersproken. Aangezien een dergelijk verwachtingspatroon precies dezelfde praktische consequenties heeft als een *ware* 'theorie van de wereld', kan op grond van Peirce' pragmatistische betekenistheorie het predikaat 'waar' *niet* aan dat stabiele verwachtingspatroon worden onthouden. Zo omschrijft Peirce waarheid.

In de rest van het tweede hoofdstuk worden de argumenten van Cheryl Misak, de voornaamste representant van een groep van hedendaagse pleitbezorgers van Peirce' waarheidsopvatting, aan een kritische beschouwing onderworpen. Haar argumenten blijken verre van afdoende, daar het principe van tweewaardigheid op twijfelachtige wijze wordt verdedigd en omdat haar argumenten voor een Peirceaanse waarheidsopvatting onverenigbaar zijn met een consequent pragmatistische kenleer. Een Peirceaanse waarheidsopvatting is derhalve een doodlopende weg. Een klakkeloze aanvaarding van het probleemdomein van de traditionele wijsbegeerte zou met zich meebrengen dat we, nu een Peirceaanse invulling van het waarheidsbegrip een onmogelijkheid gebleken is, weliswaar iets wijzer zijn dan voorheen, maar weer onversaagd verder op zoek moeten gaan naar een andere interpretatie van het waarheidsbegrip. Ik zal echter pleiten voor een bescheidener aanpak en stel derhalve voor zo veel als mogelijk van Peirce' dynamische theorie van wetenschappelijk onderzoek – zijn *belief-doubt-belief*-model – te behouden, en, ongehinderd door een of andere wijsgerige queeste naar waarheid, onze aandacht te richten op de technische bijzonderheden van hedendaagse theorieën voor *belief change* – de ware erfgenamen van Peirce' *belief-doubt-belief*-model.

In het derde hoofdstuk wordt een begin gemaakt met de uitvoering van dit voorstel. Een kritische uiteenzetting van de logische kenleer van Isaac Levi, de Amerikaanse pragmatistische filosoof die het voortouw nam bij de afbakening van het veld van logisch-filosofisch onderzoek dat bekend is geworden onder de naam *belief change*, laat zien hoe het logisch onderzoek op het vlak der *belief change* is ontsproten aan methodologische overwegingen van klassieke Amerikaanse pragmatisten. Levi stelt zich dan ook ten doel Peirce's en Dewey's *belief-doubt-belief*-model te formaliseren met gebruikmaking van klassieke logica, beslis- en waarschijnlijkheidstheorie.

Ter inleiding plaats ik Levi's wijsgerige grondhouding binnen de context van het Amerikaans pragmatisme. Zoals in het tweede hoofdstuk reeds is vermeld, nam Peirce van de negentiende-eeuwse Engelse logische school onder meer de belangstelling voor niet-deductieve redeneervormen over, en was hij, bovendien, geïnteresseerd in een systematische theorie waarmee redelijke meningsveranderingen zouden kunnen worden gevat. Hoewel Levi deze van Peirce' voorkeuren deelt, wijst hij, anders dan Peirce, een invulling van het waarheidsbegrip met behulp van een pragmatistische kenleer van de hand.

Hiervoor behoeft Levi een medium om iemands opvattingen op een bepaald tijdstip formeel te representeren. Levi kiest voor *corpora*: onder deductie gesloten verzamelingen zinnen. Ik draag twee historische en een systematische reden aan voor deze keuze. Doordat Levi de opvattingen van een actor met corpora representeert, kan hij veranderingen van mening opvatten als transities van één corpus naar een ander. Levi beweert dat alle transities tussen corpora kunnen worden verklaard met twee typen basale transities: expansies en contracties.

Een *expansie* voegt een zin aan een actors huidige corpus toe. Met aan de beslistheorie ontleende argumenten verdedigt Levi een criterium waaronder een dergelijke toevoeging gerechtvaardigd is. Dit criterium schrijft ons voor de informatieve waarde van de toe te voegen zin af te wegen tegen de plausibiliteit van deze zin. Levi legt tevens de logische uitvoering vast van een eenmaal legitiem bevonden expansie. Vervolgens ontwikkelt Levi een geïntegreerde theorie voor *contracties*. Een contractie verwijdert een zin uit het huidige corpus van een actor. Parallel aan zijn uiteenzetting inzake expansie stelt Levi criteria voor waaronder een contractie legitiem is of zelfs, in bijzondere gevallen, geboden. Anders dan bij een expansie is het corpus dat resulteert uit een contractie niet onmiddellijk – dat wil zeggen, aan de hand van logica alléén – gegeven, zodra een contractie van het huidige corpus van een actor met een zin eenmaal legitiem bevonden is. Verschillende corpora, deelverzamelingen van het te contraheren corpus, voldoen aan de logische vereisten voor de beoogde contractie. Levi raadt ons aan dát corpus te kiezen dat niet alleen vanuit logisch oogpunt voldoet, maar tevens het verlies aan informatieve waarde (er worden immers zinnen opgegeven!) zo klein mogelijk houdt. Een toepassing van Levi's theorie voor *belief change* op twee centrale thema's uit de filosofsiche logica – modale en conditionele uitspraken – sluit mijn uiteenzetting van Levi's kenleer af.

De rest van het derde hoofdstuk bestaat uit een kritische evaluatie van Levi's logische kenleer. Als eerste nemen we Levi's aanname dat de opvattingen van een actor met een corpus kunnen worden gerepresenteerd onder de loep. Bij nader inzien blijkt het ondoenlijk eenduidig te bepalen welke zinnen tot een corpus behoren en welke niet. Een tweede punt van kritiek richt zich op Levi's idealiserende uitgangspunt dat een voor de toepassing van zijn criteria benodigd stelsel van parameters reeds vooraf is gegeven. Daar zowel zijn overwegingen inzake expansie als zijn overwegingen inzake contractie stoelen op een eerdere bepaling van de parameter die informatieve waarden vastlegt, probeer ik na te gaan hoe een dergelijke bepaling in haar werk gaat. Eenzelfde strategie wordt gevolgd bij de evaluatie van de 'mate van lef', een parameter die een rol speelt bij de toepassing van het criterium voor gerechtvaardigde expansies. Het wekt bevreemding dat het gewenste resulterende corpus uiteindelijk de doorslaggevende factor is bij de schatting van de besproken parameters.

Het bovenstaande noopt ons, Levi's oogmerk normatieve criteria voor *belief change* te leveren ten spijt, tot de conclusie dat, zolang het onduidelijk blijft hoe de beginwaarden van de in Levi's systeem optredende parameters op een overtuigende en betrouwbare wijze kunnen worden vastgesteld, Levi's theorie elke normatieve kracht ontbeert. Zelfs als Levi's systeem opgevat wordt als een *beschrijving* van hoe we werkelijk redeneren, weerstaat Levi's theorie empirische toetsing, zolang de parameters niet kunnen worden bepaald zonder een beroep op het gewenste resultaat van een verandering van mening. Tot slot bepleit ik een theorie voor *belief change* die het zonder ongrijpbare buiten-logische overwegingen kan stellen, om zo, aan de ene kant, de voor het onderzoek op het vlak der *belief change* tot nu toe kenmerkende rationalistische argumentaties te vermijden en, anderzijds, onderzoek naar de empirische adequatie van voorgestelde theorieën voor *belief change* überhaupt mogelijk te maken.

Ik zet in het vierde hoofdstuk een theorie voor *belief change* uiteen die, ten eerste, anders dan standaardsystemen, geen beroep doet op buiten-logische overwegingen en, ten tweede, inconsistente informatie aankan. Niettemin zijn de opvattingen die worden voortgebracht door de gegeven informatie, ook als deze laatste inconsistent is, altijd consistent. Mijn theorie kent twee niveaus: een eerste niveau waarop de dynamiek van mogelijk inconsistente informatie wordt behandeld, en een tweede niveau waarop een verzameling consistente opvattingen uit het eerste niveau wordt geëxtraheerd.

Op het eerste niveau beschrijf ik, met first degree entailment als onderliggende logica, een techniek om mogelijk inconsistente informatie te representeren met een *eindige toestand*, die precies uit alle valuaties bestaat die minimaal vereist zijn om de gegeven informatie waar te maken. Een algoritme dat de eindige toestand van een gegeven formule levert, volgt. Daarna wordt een operatie gedefinieerd aan de hand waarvan een eindige toestand met een formule kan worden geëxpandeerd. Een representatiestelling, die aantoont dat mijn expansie-operatie aan de intuïtieve vereisten voor een expansie voldoet, karakteriseert de expansie-operatie. Vervolgens definieer ik een contractie-operatie, aan de hand waarvan een formule uit een eindige toestand kan worden geschrapt. Opmerkelijk genoeg onderstelt deze contractie-operatie geen buiten-logisch element, zoals een keuzefunctie of een ordening over (verzamelingen van) zinnen. (Een dergelijke veronderstelling is alomtegenwoordig in de op het werk van Levi en van Alchourrón, Gärdenfors en Makinson geënte systemen voor *belief change*.) Bovendien zijn mijn expansie- en contractie-operatoren elkaars dualen. Mijn bespreking van de contractie-operator mondt uit in een representatiestelling die de operator karakteriseert in termen van een verzameling postulaten. Deze postulaten worden geformuleerd in termen van het *informatie-opspansel* van een eindige toestand, een eigenschap van eindige toestanden die geen pendant heeft in de standaard mogelijke-werelden-semantiek.

Op het tweede niveau bestudeer ik vier *extractors*, operaties die een consistente verzameling 'plausibele' opvattingen ontlenen aan een eindige toestand. Indien de eindige toestand consistent is, leveren alle extractors hetzelfde resultaat. In het geval een eindige toestand inconsistent is, vallen de uit de toepassingen van de vier extractors resulterende verzamelingen 'plausibele' opvattingen in het algemeen *niet* samen.

Aanvankelijk worden eerst twee extractors besproken. De eerste extractor is een vertaling van een voorstel van Restall en Slaney naar de huidige context van eindige toestanden. De tweede extractor komt neer op een contractie van de eindige toestand met een formule die aangeeft welke propositieletters zich inconsistent in deze eindige toestand gedragen. Aangezien de extractors aanleiding geven tot niet-monotone gevolgtrekkingsoperaties, worden enkele bewijstheoretische eigenschappen van de twee genoemde extractors besproken. Vervolgens wordt een selectiefunctie voorgesteld die de meest consistente valuaties uit een eindige toestand kiest. Deze selectiefunctie kan worden gebruikt om de eindige toestand te bewerken alvorens een van de twee bovengenoemde extractors toe te passen. Zo zijn dus vier verschillende extractors gedefinieerd.

De Appendix – een in samenwerking met Koji Tanaka van Macquarie University in Sidney uitgevoerd onderzoek – bevat een oorspronkelijke bewijstheoretische studie van de relevantielogica first degree entailment (fde), die, zoals vermeld, de onderliggende logica was van het in het vierde hoofdstuk aangeboden systeem voor *belief change*. Deze logica kan worden beschouwd als een generalisering van de klassieke propositielogica (cpl), omdat ze niet alleen, zoals cpl, totale en consistente valuaties toestaat, maar tevens partiële of inconsistente valuaties. Er geldt dus dat alle gevolgtrekkingen die geldig zijn in fde eveneens geldig zijn in cpl. Het omgekeerde geldt echter niet: fde kent bijvoorbeeld geen tautologieën en bovendien gaan noch het disjunctieve syllogisme noch de *ex falso quodlibet*-regel van cpl op in fde.

Een beknopte geschiedenis van fde opent de Appendix. Voor fde zijn twee logische semantieken ontwikkeld, een tweewaardige semantiek door Routley en Routley, en een vierwaardige door Dunn. In de Appendix bedienen we ons van de vierwaardige semantiek. Er volgt een bondige bespreking van op het bewijstheoretische werk van Łukasiewicz gestoelde 'gecombineerde systemen'. In deze gecombineerde systemen kunnen zowel aanvaarde als verworpen formules worden afgeleid, die in de afleidingen van elkaar worden onderscheiden doordat ze worden voorafgegaan door een teken ('$\vdash$' voor aanvaarding, '$\dashv$' voor verwerping) dat hun status aangeeft. Dit idee zal worden toegepast bij de ontwikkeling van een systeem van natuurlijke deductie voor fde. Na definities van de taal en de vierwaardige semantiek voor fde volgt een systeem van natuurlijke deductie voor fde in Gentzen-vorm. De correctheid van dit systeem wordt op de gebruikelijke wijze aangetoond. Bij het volledigheidsbewijs gebruiken we Henkin's methode, hoewel de voor het inbeddingslemma gebruikte constructie aanpassing behoeft om aan de vereisten van het modelexistentielemma te kunnen voldoen, waar immers een vierwaardig tegenvoorbeeld is vereist.

Titles in the ILLC Dissertation Series:

ILLC DS-1996-01: **Lex Hendriks**
Computations in Propositional Logic

ILLC DS-1996-02: **Angelo Montanari**
Metric and Layered Temporal Logic for Time Granularity

ILLC DS-1996-03: **Martin H. van den Berg**
Some Aspects of the Internal Structure of Discourse: the Dynamics of Nominal Anaphora

ILLC DS-1996-04: **Jeroen Bruggeman**
Formalizing Organizational Ecology

ILLC DS-1997-01: **Ronald Cramer**
Modular Design of Secure yet Practical Cryptographic Protocols

ILLC DS-1997-02: **Nataša Rakić**
Common Sense Time and Special Relativity

ILLC DS-1997-03: **Arthur Nieuwendijk**
On Logic. Inquiries into the Justification of Deduction

ILLC DS-1997-04: **Atocha Aliseda-LLera**
Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence

ILLC DS-1997-05: **Harry Stein**
The Fiber and the Fabric: An Inquiry into Wittgenstein's Views on Rule-Following and Linguistic Normativity

ILLC DS-1997-06: **Leonie Bosveld - de Smet**
On Mass and Plural Quantification. The Case of French 'des'/'du'-NP's.

ILLC DS-1998-01: **Sebastiaan A. Terwijn**
Computability and Measure

ILLC DS-1998-02: **Sjoerd D. Zwart**
Approach to the Truth: Verisimilitude and Truthlikeness

ILLC DS-1998-03: **Peter Grunwald**
The Minimum Description Length Principle and Reasoning under Uncertainty

ILLC DS-1998-04: **Giovanna d'Agostino**
Modal Logic and Non-Well-Founded Set Theory: Translation, Bisimulation, Interpolation

ILLC DS-1998-05: **Mehdi Dastani**
Languages of Perception

ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke

ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation

ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz

ILLC DS-1999-04: **Barbara Terhal**
Quantum Algorithms and Quantum Entanglement

ILLC DS-2000-01: **Renata Wasserman**
Resource Bounded Belief Revision

ILLC DS-2000-02: **Jaap Kamps**
A Logical Approach to Computational Theory Building (with applications to sociology)

ILLC DS-2000-03: **Marco Vervoort**
Games, Walks and Grammars: Problems I've Worked On

ILLC DS-2000-04: **Paul van Ulsen**
E.W. Beth als logicus

ILLC DS-2000-05: **Carlos Areces**
Logic Engineering. The Case of Description and Hybrid Logics

ILLC DS-2000-06: **Hans van Ditmarsch**
Knowledge Games

ILLC DS-2000-07: **Egbert L.J. Fortuin**
Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology.

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability