RELEVANT LOGICS, MODAL LOGICS AND THEORY CHANGE

André Fuhrmann

A thesis submitted for
the degree of Doctor of Philosophy
of the Australian National University

Department of Philosophy and
Automated Reasoning Project
Research School of Social Sciences
September 1988

Except where otherwise acknowledged,
this thesis is my own work.

Abstract

This thesis is a contribution to applied relevant logics. In Part One relevant logics are presented proof-theoretically and semantically. These logics are then extended to modal logics. Completeness proofs for all of the logics presented in Part One are provided. In Part Two, the logics of Part One are applied to certain problems in philosophical logic and Artificial Intelligence. Deontic and epistemic logics based on relevant logics are presented in chapter three and chapter four contains an extensive investigation of the logic of theory change (or database updating).

Acknowledgements

For all he taught me and for his perfect balance between penetrating criticism and unfailing encouragement, I wish to record my deep gratitude to my supervisor Dr Richard Sylvan. I also thank Prof Frank Jackson of the Department of Philosophy and Dr Michael McRobbie of the Automated Reasoning Project for providing a superb environment for carrrying out the research recorded in this thesis. My thoughts on the topics of this thesis have been shaped in many discussions with past and present members of and visitors to the Department of Philosophy and the Automated Reasoning Project. Among those who have helped me in this way, I wish to thank in particular Peter Lavers, Prof Isaac Levi, Dr Errol Martin, Michaelis Michael, Dr Stephen Read, Dr John Slaney, and Prof Neil Tennant. Some of those mentioned have also been so kind as to comment in writing on various parts of this thesis.

I owe a great deal to the support of my friends during writing this thesis: Steve Buckle, Tim Burgess, Gerald Haberkorn, Peter Lavers, Philippa Mein Smith, Michaelis Michael, Jacques Riche, Kim Sterelny, and Igor Urbas. I shall forever be indebted to Prof Jack Smart and his wife Liz for making me feel at home in Canberra.

My wife Susanne ... *SINE QUA NON*. I would have dedicated this thesis to her, had my parents not a prior claim to this small token of gratitude. So I dedicate this thesis to my parents, Wolfgang and Irmgard Fuhrmann.

Last but not least: thanks to John Barlow for keeping up the Sun.

Contents

Introduction

This dissertation is a contribution to the study of relevant logics. Its emphasis is on applications. Such an emphasis, I believe, is timely. For the purely philosophical debate about the notion of entailment has reached a deadlock. It has issued on the one side in an elaborate classical epicycle[1] and on the other side in a rich fundus of well-investigated alternatives to classical logic.[2] The divide between these two sides is unlikely to become permeable by further reflections on the elusive notion of entailment or introspection of one's linguistic intuitions about *if... then....* Progress, however, can perhaps be made by observing the contenders "in use" rather than *in vacuo*.

Almost coinciding with the decline of the entailment debate within the philosophical community is the increasing interest in non-classical logics among researchers in Artificial Intelligence (AI). It has become plain in recent years that for the solution of many problems in AI, classical logic is either not suited at all or an extremely cumbersome tool to use. Thus, in AI, alternatives to classical logic are now considered and evaluated free from the philosophical prejudices hardened in a seven decades spanning debate about "deviant" logics – non-classical logics suddenly get a "fair go".

The present dissertation attempts to take advantage of the open-minded attitude with which various logics are now considered in AI. Thus, the applications of relevant logics in Part Two of this dissertation are presented with a view to problems in AI. These problems fall under the heading of database theory. Chapter three offers some tools for reasoning about databases in a fixed state; chapter four treats the problem of database updating. In more traditional terms, however, these chapters contain also contributions to philosophical logic: chapter three presents some epistemic and deontic logics based on relevant logics, and chapter four is an exercise in the logic of theory change. The discussion in Part Two will frequently switch between philosophy and AI. Such a transfer of ideas, I believe, is beneficial to both disciplines.

Chapter one provides a grounding in the proof theory and semantics of relevant logics. We give axiomatic formulations of a group of logics, starting from a very weak system **BM** and proceeding to classical logic **K** via the comparatively strong relevant logics of Anderson and Belnap

[1] See e.g. Jackson (1987).
[2] See e.g. Routley, Meyer et al. (1982).

(1975) and the semi-relevant systems **RM** ("Mingle") and **RM3**. All of these logics will be proved complete with respect to appropriate classes of model structures (frames) of the kind used in Routley, Meyer, et al. (1982). The aim of this chapter is to provide a self-contained completeness argument for all of the major relevant logics (and a few more) as a background to the following chapters. In presenting this argument I have benefited from Dunn's survey article on relevant logics (1986).

In chapter two we shall consider extensions of the systems presented in chapter one in a language including a unary modal operator. The resulting modal systems will be proved sound and complete with respect to two extensions of the semantics introduced in chapter one. The two extensions are, first, a Kripke-style semantics, modelling the modal operator by means of a binary accessibility relation, and, secondly, a Montague-Scott-style semantics in which the modal operator is modelled by means of a so-called neighbourhood function.

In Part Two, we shall put the systems of Part One to use. The modal logics of chapter two will be used in chapter three as a means to represent and reason about the *static* properties of theories of various kinds. We shall consider in some detail two kinds of theories: sets of sentences an agent is committed to accept as true at a particular point of time ("acceptance sets"), and sets of sentences an agent is committed to make true at a particular point of time ("norm sets"). As a result of these considerations, logics of acceptance (or commitment-to-believe) and of obligation will emerge. We shall refrain from enshrining in these logics idealising assumptions about acceptance sets and norm sets; in particular, we shall not assume that such sets are always consistent. The possibility, and indeed actuality, of inconsistent but non-trivial acceptance and norm sets will motivate the move towards epistemic and deontic logics based on a paraconsistent logic. The concern with representing correctly the deductive dependencies within acceptance sets and norm sets will motivate a move towards epistemic and deontic logics based on a relevant logic.

Chapter four focuses on certain *dynamic* aspects of theories. The study of the formal aspects of theory change – though a natural complement to the investigations of Tarski (1930) – has been curiously neglected for a long time. A beginning has only recently been made in the work of Alchourron, Gärdenfors and Makinson. Though squarely based within the framework provided by these three authors, the present contribution to the theory of theory change differs in a number of aspects from their work. First, Alchourron, Gärdenfors and Makinson (AGM) consider changes of theories by one sentence at a time. I consider multiple changes: changes by sets of sentences at a time. Changes by single sentences will emerge as a special case of multiple changes, namely as changes by singleton sets of sentences. Secondly, AGM think of theories as sets of sentences closed under logical consequence; theories are thus rather amorphous objects. I think of theories as sets of sentences generated from a distinguished set of sentences (the base of the theory in question) by means of a logical consequence operation. As I shall argue in chapter four, the base of a theory does play an important role in changing a theory. Thirdly, a central concern for AGM is that changes to theories ought to be minimal: a changed theory should be as big a subset of the original theory as possible under the circumstances. I shall argue that minimality of change is a rule of thumb that may easily be overridden by other constraints on theory change. One such constraint – not recognised in the work of AGM – is that if a sentence B is in a theory just because A is in that theory, then B should not remain in the theory after A has been removed. I call this constraint on theory change 'the filtering condition'. Fourthly, for AGM, theories are closed under a consequence operation provided by classical logic. In view of classical theses like $A \rightarrow . {\sim} A \rightarrow B$ and $A \rightarrow . B \rightarrow A$, the change of inconsistent theories and the removal of logical truths from a theory receive a rather special treatment in AGM's theory. The theory advanced in this dissertation will be more general: any one of the logics of chapter one may provide the consequence operation theories are closed under. However, as I shall argue in chapter four, only if theories are closed under a non-classical, relevant, consequence operation, does a satisfactory account of how inconsistent theories ought to change and how to remove logical truths from a theory emerge.

The chapters of Part Two complement each other in a straightforward sense: while chapter three provides a formal framework for reasoning about theories at a particular point of time, theories as they "move" along a time axis are the subject of formal investigations in chapter four. The formal tools employed in these chapters are, however, quite distinct. Whereas modal logics provide the background for chapter three, Tarski's theory of consequence operations is the unifying theory behind the considerations in chapter four. In the final section of this dissertation, an outlook on one way of bringing to bear modal logic on the theory of theory change will be given by employing the resources of dynamic logic in order to formulate a logic of theory change.

Part One
Relevant Logics and Modal Logics

Chapter I

Relevant logics determined by R*-models

1. Language

By a *propositional language*, **PL**, we mean a triple $\langle \mathbf{At},\mathbf{Op},r\rangle$, where **At** is a set (of propositional *atoms*), **Op** is a set (of propositional *connectives*), and r is a function (the *rank* function) from **Op** to the set of natural numbers **N**. We require that **At** and **Op** are denumerable and disjoint sets. We shall use p,q,... (occasionally subscripted with numerals) as variables ranging over members of **At**, and ϕ will stand variably for members of **Op**.

The function r assigns a rank to every connective. If $r(\phi)=0$ (=1, =2, ...) then we shall say that ϕ is a nullary (unary, binary, ...) connective. Nullary connectives will also be referred to as *propositional constants*.

An *expression* of **PL** is any nonempty finite sequence of members of $\mathbf{At}\cup\mathbf{Op}$. The set **Wff** of *well-formed formulae* of **PL** is inductively defined as follows.

(i) $\mathbf{At}\subseteq\mathbf{Wff}$,

(ii) for each $\phi\in\mathbf{Op}$:
if $r(\phi)=n$ and $A_1,\ldots,A_n\in\mathbf{Wff}$, then $\phi A_1\cdots A_n\in\mathbf{Wff}$.

We shall use A,B,C,... (occasionally subscripted with numerals) as variables ranging over the set of well-formed formulae.

According to our last definition, formulae are written down in what is known as *Polish notation*, that is, with connectives prefixed to formulae, thereby dispensing with the need for some device, such as brackets, delimiting the scope of the connectives. Despite the formal elegance of the prefix notation, it has never enjoyed widespread popularity; the longer the formula, the more effort has to be spent on figuring out the subformulae of which it is composed. By contrast, using brackets often allows to grasp the "meaning" of a formula at a glance. We shall adopt a device here that allows for economy of primitive symbols while making formulae more readable. We introduce brackets into the metalanguage in which we write about formulae of **PL** by means of the following notational convention:

(Let ϕ^n be a connective of rank n.)
If $A = \phi^n A_1 \cdots A_n$ is a formula of **L**, then A will be represented in the text as $\phi^n(A_1 \cdots A_n)$.

Moreover, for binary connectives, we shall help ourselves to infix notation:

$\phi^2(AB) := (A\phi B)$.

Brackets will be dropped as long as no confusion can arise. We shall also mix brackets with interpunctuation along the conventions of Church (1956), adopted and explained in Anderson and Belnap (1975), p.6. And when fixing the connective set of particular languages, we shall grade the power of connectives to bind propositional variables as usual. In sum: we shall mix all of the better known delimiting devices to make reading formulae as easy as possible.

A language $\mathbf{PL}_1 = \langle \mathbf{At}_1, \mathbf{Op}_1, r_1 \rangle$ is an *extension* of $\mathbf{PL}_2 = \langle \mathbf{At}_2, \mathbf{Op}_2, r_2 \rangle$ if and only if

$\mathbf{At}_1 \subseteq \mathbf{At}_2$,
$\mathbf{Op}_1 \subseteq \mathbf{Op}_2$, and
$r_1 = r_2$ for domain $\mathbf{Op}_1$.

If $\mathbf{PL}_1$ is an extension of $\mathbf{PL}_2$, then $\mathbf{PL}_1$ is a *fragment* of $\mathbf{PL}_2$, and *vice versa.*

The propositional language **La** will underly all our considerations in this chapter. Languages considered in subsequent chapters will be extensions of **La**. **La** has as its set of primitive connectives $\{\sim, \&, v, \rightarrow\}$. The formation rules are as expected, i.e. all atoms are well-formed formulae of **La**, and if A and B are well-formed formulae, so are $\sim A$, $\&AB$, vAB, and $\rightarrow AB$ (write $\sim A$, $A\&B$, AvB and $A \rightarrow B$ respectively). As informal readings of these connectives the following are recommended.

$\sim A$ – not: A,
$A\&B$ – A and B,
AvB – A or B,
$A \rightarrow B$ – A implies B.

The force of these connectives to bind propositional variables – their "valence" – decreases in the following order: $\sim, \&, v, \rightarrow$. The natural valence of a connective may be overridden by punctuation and bracketing according to the usual conventions.

2. R*-frames and -models

In this section we shall define the basic set-theoretic structure which will serve us throughout this thesis as a basis for interpreting the propositional languages we shall be dealing with. In subsequent chapters we shall restrict or extend the notion of an R*-model in various ways. The results of this and the following section will provide a point of reference for all subsequent developments.

An *R*-frame* is a structure $<\mathbf{0},\mathbf{K},R,*>$, where $\mathbf{K}$ is a nonempty set of indices (points, worlds, situations, set-ups, theories, etc.), $\mathbf{0}$ is a distinguished subset of $\mathbf{K}$ (including The Real World), R is a ternary relation on $\mathbf{K}$, i.e. $R \subseteq \mathbf{K}^3$, and $*$ is a unary operation on members of $\mathbf{K}$, i.e. $*:\mathbf{K}\to\mathbf{K}$.

We shall write $0x$ for $x \in \mathbf{0}$ and define

d1. $$a \rhd b \text{ iff } (\exists x)(0x \text{ and } Rxab).$$

By "default" R*-frames will be equipped with a number of conditions on R and $*$.[1] For any points a,b,c,d in the universe $\mathbf{K}$ of an R*-frame, we require the following conditions to hold.

r1. (Identity)	$a \rhd a$
r2. (Monotonicity)	if $a \rhd b$ and $Rbcd$ then $Racd$
*1. ($\rhd$-Inversion)	if $a \rhd b$ then $b^* \rhd a^*$

An *R*-model M* is represented by a pair $<F,V>$, where F is an R*-frame and V is a valuation function distributing propositional atoms over members of $\mathbf{K}$, i.e. $V:\mathbf{At}\to 2^{\mathbf{K}}$. The valuation is subject to the (atomic) heredity constraint, i.e.

[1] Gabbay (1976) investigates consequence relations determined by frames without our default conditions. These consequence relations, however, cannot be empty on the left-hand-side, i.e., the logics determined are "theorem-less".

(h) if $a \rhd b$ and $a \in V(p)$, then $b \in V(p)$.

Given a valuation V on propositional atoms, we interpret the non-atomic formulae of **La** by the following inductive definition of the forcing relation $\models$ (read '$a \models A$': V forces A to hold at a).

For every $a \in \mathbf{K}$:

(p) $a \models p$ iff $a \in V(p)$

(~) $a \models {\sim}A$ iff $a^* \not\models A$

(&) $a \models A \& B$ iff $a \models A$ and $a \models B$

(v) $a \models A \vee B$ iff $a \models A$ or $a \models B$

(→) $a \models A \to B$ iff $(\forall bc \in \mathbf{K})$(if $Rabc$ and $b \models A$ then $c \models B$)

The definitions of truth and validity are as follows. (It will be convenient to reserve o for an arbitrary representative of the set $\mathbf{0}$; thus, $o \models A$ says that A holds throughout $\mathbf{0}$, i.e. $(\forall x)(0x \supset x \models A)$.)

(T) A is *true at a point a in a model M* if and only if $a \models A$ in M; A is *true in M* if and only if $x \models A$ for all $x \in \mathbf{0}$ in M (or, simpler: A is true in M iff $o \models A$ in M).

(V) A is *valid in the class of all* R**-frames* if and only if A is true in every model M on an arbitrary R*-frame F.

Furthermore, we shall say that

(E) A *entails* B in M if and only if for all points a in M, if $a \models A$ then $a \models B$.

For the task of verifying formulae in models, two facts will be useful.

LEMMA 2.1. Heredity

For any formula A and any points a and b in an R*-model: if $a \models A$ and $a \rhd b$, then $b \models A$.

Proof. Induction on the complexity of A; the base is given by the atomic heredity condition (h).

For $A = {\sim}B$ we need *1: $a \rhd b \supset b^* \rhd a^*$. Suppose

(1) $a \vDash {\sim}B$ and (2) $a \rhd b$.

By (~) from (1), we have

(3) $a^* \nvDash B$

and, by *1 from (2), we obtain

(4) $b^* \rhd a^*$.

Thus, contraposing and instantiating the inductive hypothesis,

$a^* \nvDash B \ \& \ b^* \rhd a^* \ \supset b^* \nvDash B$,

it follows from (3) and (4) that $b^* \nvDash B$ whence, by (~), $b \vDash {\sim}B$.

For $A = B \rightarrow C$ we need r2: $a \rhd b \ \& \ Rbcd \supset Racd$. Suppose

(1) $a \vDash B \rightarrow C$ and (2) $a \rhd b$.

To show: $b \vDash B \rightarrow C$, i.e. (by ($\rightarrow$)) $\forall cd(Rbcd \ \& \ c \vDash B \supset d \vDash C)$. So suppose further that

(3) $Rbcd$ and (4) $c \vDash B$.

Spelling out (1) according to ($\rightarrow$) we have

(5) $\forall cd(Racd \ \& \ c \vDash B \ \supset d \vDash C)$.

By r2 we obtain from (2) and (3), $Racd$ which we can use together with (4) to detach the required $d \vDash C$ from (5). ■

THEOREM 2.2. Verification

For any R*-model M, A entails B in M if and only if $M \vDash A \rightarrow B$.

Proof.

($\Rightarrow$). Suppose that A entails B in M, i.e.

(1) $\forall a(a \vDash A \ \supset a \vDash B)$.

To show: $o \vDash A \rightarrow B$, that is, $\forall ab(a \rhd b \ \& \ a \vDash A \ \supset b \vDash B)$. So suppose further that

(2) $a \rhd b$ and (3) $a \vDash A$.

From (1) and (3) we derive

(4) $a \vDash B$.

Then the required $b \vDash B$ follows from (2) and (4) by the heredity lemma.

($\Leftarrow$). For this direction we need r1: $a \rhd a$. Suppose $o \vDash A \rightarrow B$, i.e.

(1) $\forall ab(a \rhd b \ \& \ a \vDash A \ \supset b \vDash B)$.

To show: $\forall a(a \vDash A \ \supset a \vDash B)$. So suppose further that

(2) $a \vDash A$.

Then it follows by r1 from (1) and (2) that $a \vDash B$, as required. ■

Note that for the proof of the verification theorem we have made exhaustive use of all conditions on R and * (and we had to appeal to the atomic heredity condition (h)). These conditions on models will stay with us throughout this thesis. When extending our language **La** by new connectives, we shall "match" these connectives semantically by adding a new frame-operation and extending the definition of an R*-model accordingly. Thus, we shall be in a position to take over the results of this section (the heredity lemma and the verification theorem) *provided* that we can complete the induction required in the proof of the heredity lemma for the extended language.

3. The system BM and the basic completeness result

The logic **BM** is the smallest set in **La** such that each sentence in **BM** is either an instance of the axiom schemas listed below or can be derived from such instances by successive applications of the rules listed below.

A1. (L-Simplification, l&E)	$A\&B \rightarrow A$
A2. (R-Simplification, r&E)	$A\&B \rightarrow B$
A3. (&-Composition, &C)	$(A \rightarrow B)\&(A \rightarrow C) \rightarrow .A \rightarrow B\&C$
A4. (L-Addition, lvI)	$A \rightarrow AvB$
A5. (R-Addition, rvI)	$B \rightarrow AvB$
A6. (v-Composition, vC)	$(A \rightarrow C)\&(B \rightarrow C) \rightarrow .AvB \rightarrow C$
A7. (Distribution, Dist)	$A\&(BvC) \rightarrow (A\&B)v(A\&C)$
A8. (DeMorgan k, DMk)	${\sim}(A\&B) \rightarrow {\sim}Av{\sim}B$
A9. (DeMorgan j, DMj)	${\sim}A\&{\sim}B \rightarrow {\sim}(AvB)$
A10. (Identity, I)	$A \rightarrow A$
MP. (Modus Ponens)	$\dfrac{A\ ,\ A \rightarrow B}{B}$
ADJ. (Adjunction)	$\dfrac{A\ ,\ B}{A\&B}$

PREF. (Prefixing) $$\frac{B \to C}{A \to B \to A \to C}$$

SUFF. (Suffixing) $$\frac{A \to B}{B \to C \to A \to C}$$

CP. (Contraposition) $$\frac{A \to B}{\sim B \to \sim A}$$

The rules PREF and SUFF may equivalently be replaced by the single rule

AFF. (Affixing) $$\frac{A \to B \ ,\ C \to D}{B \to C \to A \to D}.$$

As the completeness argument for **BM** will reveal, **BM** is the smallest logic determined by the class of all R*-models. For that argument we shall need a few facts about **BM** listed in the following theorem.

THEOREM 3.1.

The following formulae are theorems of **BM**.

(i) $(A \to C)\&(B \to D) \to A \vee B \to C \vee D$

(ii) $(A \to C)\&(B \to D) \to A\&B \to C\&D$

(iii) $(A \to C)\&(B \to D) \to A\&B \to C \vee D$

Moreover, the rule

CUT. $$\frac{A \to B \ ,\ A\&B \to C}{A \to C}$$

is derivable in **BM**.

Proof.

The derivations of (i), (ii) and (iii) are similar. We illustrate the method of derivation by giving the proof of (ii). From A1 and A2 we obtain by SUFF

(1) $A \to C \to A\&B \to C$ and (2) $B \to D \to A\&B \to D$

whence, using PREF and MP,

(3) $(A \to C)\&(B \to D) \to A\&B \to C$ and

(4) $(A \to C)\&(B \to D) \to A\&B \to D$.

From $(3)\&(4)$ it follows by A3 that

(5) $(A \to C)\&(B \to D) \to .(A\&B \to C)\&(A\&B \to D)$.

But by A3 again,

(6) $(A\&B \to C)\&(A\&B \to D) \to .A\&B \to C\&D$.

Thus, the theorem follows from (5) and (6) by transitivity, i.e. SUFF and MP.

To derive CUT, assume (1) $A \to B$. It follows from A10 and (1) by ADJ that

(2) $(A \to B)\&(A \to A)$

whence, by A3 and MP,

(3) $A \to A\&B$.

From (3) by SUFF:

(4) $A\&B \to C \to .A \to C$.

Now we make use of the second premiss, $A\&B \to C$, to detach $A \to C$ from (4) by MP. ■

Before turning to the completeness proof, we shall first show that all theorems of **BM** are valid in the class of all R*-models (*soundness*).

THEOREM 3.2. Soundness

If A is a theorem of **BM**, then A is valid in the class of all R*-frames.

Proof. It will suffice to pick an arbitrary R*-model M and prove that all axioms of **BM** are true in that model (i.e. that they hold at an arbitrary point $o \in \mathbf{0}$) and that the rules preserve truth-in-M. In view of the verification theorem, an axiom of the form $A \to B$ can be shown to hold throughout $\mathbf{0}$ by proving for an arbitrary point $a \in \mathrm{K}$ that if $a \models A$, then $a \models B$. The details are routine and hence omitted. (Note that we needed the special conditions on R and * only for the verification theorem. For the soundness argument, these conditions need not be invoked again.) ■

We shall now prove that formulae true in all R*-models are theorems of **BM**. Together with the soundness theorem, these results will ensure that the notions of provability in **BM** and validity on R*-frames are extensionally equivalent, or, as we shall say, that the logic **BM** is *determined* by the class of all R*-frames. Thus, in the remainder of this section we shall prove the following proposition.

THEOREM 3.3. Completeness

If A is valid in the class of all R*-frames, then A is a theorem of **BM.**

The argument follows a pattern familiar from Henkin-style completeness arguments for modal logics.[2] We shall define a *canonical model* $M_{\mathbf{BM}} = \langle 0, \mathrm{K}, R, *, V \rangle$ which refutes some arbitrarily chosen non-theorem D of **BM.** We shall then prove that the canonical model $M_{\mathbf{BM}}$ is indeed an R*-model. Since the non-theorem D was chosen arbitrarily, we may construct such a canonical R*-model for every non-theorem of **BM.** Thus, for any D such that $\nvdash_{\mathbf{BM}} D$, there is some R*-model M such that $M \nvDash D$ – which is the contrapositive of T3.3.

The details of this argument require some work. We start with defining the notion of a canonical model (and, prior to that, the notion of an **L**-theory of some kind or another) after which we shall pause for a moment to give a brief overview of the argument as a whole.

DEFINITION 3.4. **L**-theory

Let **L** be any subset of the set of well-formed formulae of some extension of the language **La.** A set of sentences T is an **L**-theory, if and only if both T is *closed under adjunction*, i.e. for any sentences A, B

(a) if $A \in T$ and $B \in T$, then $A\&B \in T$,

and T is *closed under* **L**-*implication*, i.e.

(b) if $A \in T$ and $A \rightarrow B \in \mathbf{L}$, then $B \in T$.

A set of sentences is *regular with respect to* **L** just in case

(c) $\mathbf{L} \subseteq T$.

And T is said to be *prime* if and only if

(d) if $A \vee B \in T$, then either $A \in T$ or $B \in T$.

A set of sentences satisfying the conditions (a) to (d) will be called a *saturated* **L**-*theory.*

[2] Such completeness arguments were first given in Makinson (1966).

DEFINITION 3.5. Canonical model of type [R*] for **L**

A canonical model of type [R*] for **L** is a quintuple

$$M_{\mathbf{L}} = \langle \mathbf{0}_{\mathbf{L}}, \mathbf{K}_{\mathbf{L}}, R_{\mathbf{L}}, {}^*{}_{\mathbf{L}}, V_{\mathbf{L}} \rangle,$$

(subscripts '**L**' from now on omitted) such that

(a) **K** is a nonempty set of prime **L**-theories;

(b) $\mathbf{0} \subseteq \mathbf{K}$ is a set of saturated **L**-theories such that for each non-theorem D of **L** there is a theory $x \in \mathbf{0}$ with $D \notin x$;

(c) $R \subseteq \mathbf{K}^3$ such that
$Rabc$ iff $(\forall A, B \in \mathbf{Wff})(A \rightarrow B \in a \ \&\ A \in b \supset B \in c)$;

(d) $^*: \mathbf{K} \rightarrow \mathbf{K}$ such that
$a^* = \{A : {\sim}A \notin a\}$.

(e) $V: \mathbf{At} \rightarrow 2^{\mathbf{K}}$ such that
$a \in V(p)$ iff $p \in a$;
V is extended to a relation $\models \subseteq \mathbf{K} \times \mathbf{Wff}$ such that
$a \models A$ iff $A \in a$.

Substitute **BM** for **L** and the above definition turns into a definition of a canonical model of type [R*] for **BM**.

The completeness argument will be completed after we have shown that the definition of a canonical model is "good" (i.e. that such models exist) and that such canonical models belong indeed to the class of R*-models. More specifically, we shall have to discharge five claims:

(i) **K** is nonempty.

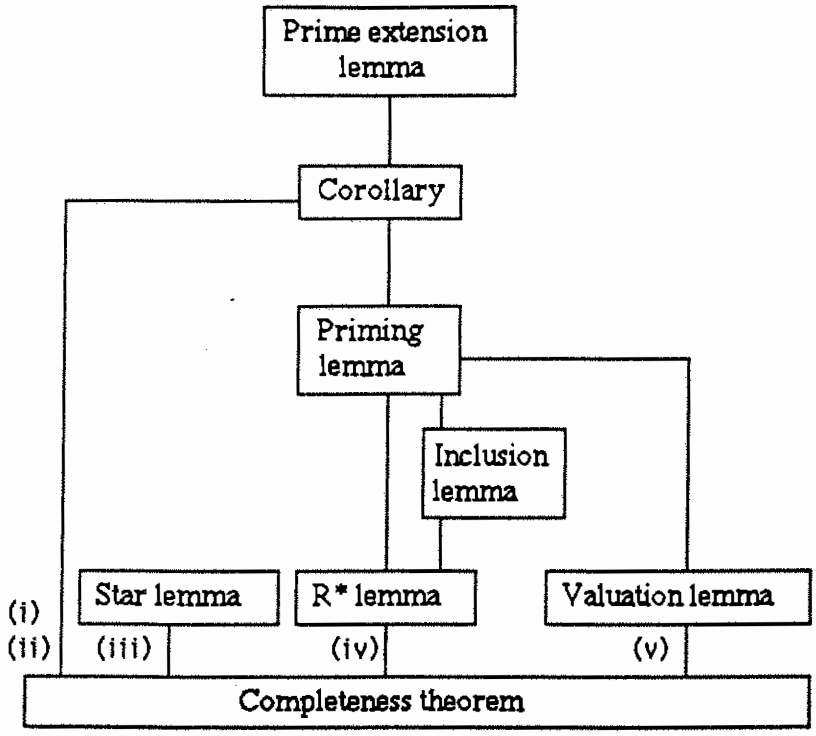
(ii) **0** is nonempty.

(iii) The function * maps members of **K** into members of **K**.

(iv) R and * satisfy the conditions r1, r2 and *1.

(v) The canonical valuation V satisfies (h) and the canonical $\models$ satisfies the truth-conditions for complex formulae.

We shall settle (i) and (ii) by constructing – in the manner of Lindenbaum – a saturated **BM**-theory, keeping an arbitrarily chosen non-theorem of **BM** out of the constructed theory (the *prime extension lemma* and its corollary). Claim (iii) will be discharged by means of the *star lemma.* We shall prove (iv) in the *R* lemma.* But for the R* lemma we shall need the *inclusion lemma* and the *priming lemma.* The latter will also be needed for the *valuation lemma* verifying (v). The dependency relations among these lemmata are charted out below. (Roman numerals along arrows indicate at which point which claims will have been discharged.)



Where T and Δ are sets of sentences, we write

$$T \vdash \Delta$$

to express the fact that for some finite collections of sentences $\{A_1, \ldots, A_m\} \subseteq T$ and $\{D_1, \ldots, D_n\} \subseteq \Delta$,

$$A_1 \& \cdots \& A_m \to D_1 v \cdots v D_m \in \mathbf{L}$$

for some contextually fixed logic **L**.

For the remainder of this section we shall mean by a *logic* any extension of the basic system **BM**. It is easily verified that the relation $\vdash_{\mathbf{L}}$ (for any logic **L**) has the following properties.

I. $\quad \Gamma, \Delta \vdash \Delta$

II. $\quad \dfrac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}$

III. $\quad \dfrac{\Gamma, A \vdash \Delta\ ,\ \Gamma, B \vdash \Delta}{\Gamma, A v B \vdash \Delta}$

IV. $\quad \dfrac{\Gamma \vdash A\ \ ,\ \Gamma, A \vdash \Delta}{\Gamma \vdash \Delta}$

V. $$\frac{\Gamma\vdash\Lambda\ ,\ \Lambda\vdash\Delta}{\Gamma\vdash\Delta}.$$

(*Proof.* I follows immediately from A2; II is derivable from A1 and SUFF; to derive III use ADJ, A6, MP, A7, and transitivity (i.e. SUFF and MP); and IV is essentially the rule CUT already derived in T3.1. Transitivity, V, is derivable from II and IV.)

LEMMA 3.6. Prime extensions (Lindenbaum)

Let T be an L-theory and let Δ be a set of formulae such that $T\nVdash_{L}\Delta$. Then there exists a prime L-theory T' such that $T\subseteq T'$ and $T'\nVdash_{L}\Delta$.

Proof. Enumerate all formulae of La. Then construct T' as follows.

$$T_0 = T$$

$$T_{n+1} = \begin{cases} T_n\cup\{A_{n+1}\}, \text{ if } T_n, A_{n+1}\nVdash\Delta \\ T_n \text{ otherwise} \end{cases}$$

$$\bigcup_{n=0}^{\omega} T_n = T'$$

Claim (a): $T\subseteq T'$. Obvious from the cumulativity of the construction.

Claim (b): $T'\nVdash\Delta$. Simple inductive argument. The base case holds by definition of T_0. For each further step use property I of $\vdash$.

Claim (c): T' is closed under adjunction. Suppose $A, B\in T'$ and yet $A\&B\notin T'$. Then (1) $T'\vdash A\&B$ and at some stage $i+1$ in the construction $A\&B$ could not have been added to T_i because $T_i, A\&B\vdash\Delta$, hence, (2) $T', A\&B\vdash\Delta$. But now we can apply IV to premisses (1) and (2) to obtain $T'\vdash\Delta$ – contradicting (b).

Claim (d): T' is closed under L-implication. Suppose (1) $A\in T'$, (2) $A\vdash B$ and yet (3) $B\notin T'$. It follows from (1) by I that $T'\vdash A$ whence from (2) by V, (4) $T'\vdash B$. But, as explained under (c), from (3) we may infer (5) $T', B\vdash\Delta$. Hence, IV applied to premisses (4) and (5) gives $T'\vdash\Delta$, contradicting (b).

Claim (e): T' is prime. Suppose (1) $AvB\in T'$ and

(2) $A\notin T'$ and $B\notin T'$.

Thus,

(3) $T', A\vdash\Delta$ and $T', B\vdash\Delta$

whence, by III,

(4) $T', A \vee B \vdash \Delta$.

But it follows from (1) that (5) $T' \vdash A \vee B$ which, together with (4), entails in virtue of IV that $T' \vdash \Delta$, again contradicting (b). ■

COROLLARY 3.7.

For any logic **L**:

(i) Let T be an **L**-theory and Δ be a set of formulae closed under disjunction (i.e. whenever $A, B \in \Delta$, $A \vee B \in \Delta$) such that $T \cap \Delta = \varnothing$. Then there exists a prime **L**-theory T' such that $T \subseteq T'$ and $T' \cap \Delta = \varnothing$.

(ii) The set **0**, and hence **K**, in a canonical model of type [R*] for **L** is nonempty.

Proof.

Ad (i). Notice that $T \nvdash \Delta$ (by closure of Δ under disjunction and disjointness of T and Δ). Thus we can infer the existence of the required theory using the lemma.

Ad (ii). Let T in the lemma be **L** and let Δ be $\{D\}$ where D is not a theorem of **L**. Then there exists a prime **L**-theory x with $D \notin x$. Moreover, since $\mathbf{L} \subseteq x$, x is a saturated **L**-theory excluding D, i.e. $x \in \mathbf{0}$. ■

The next lemma consists of four propositions. We shall use only the first two ((i) and (ii)) this section. The other two propositions will be needed in the completeness argument for certain extensions of **BM**. We shall explicitly appeal to propositions (iii) and (iv) in the proof of theorem 8.2.

see errata for new proof of 3.8

LEMMA 3.8. Priming

Let $\mathbf{K}'$ be the set of all $\mathbf{L}$-theories (where $\mathbf{L}$ is some logic). Let R' be the extension of the canonical relation $R \subseteq K^3$ to domain $\mathbf{K}'^3$, i.e., for any $a',b',c' \in \mathbf{K}'$: if $A \to B \in a'$ and $A \in b'$, then $B \in c'$ [$A,B \in \mathbf{Wff}$].

(i) For $a',b' \in \mathbf{K}'$ and $c \in \mathbf{K}$: if $R'a'b'c$, then there is a theory $a \in \mathbf{K}$ such that $a' \subseteq a$ and $Rab'c$.

(ii) For $a \in \mathbf{K}$ and $b',c' \in \mathbf{K}'$ such that $R'ab'c'$ and $B \notin c'$, there are theories $b,c \in \mathbf{K}$ such that $Rabc$, $b' \subseteq b$ and $B \notin c$.

(iii) For $a',b' \in \mathbf{K}'$ and $c \in \mathbf{K}$: if $R'a'b'c$, then there is a theory $b \in \mathbf{K}$ such that $b' \subseteq b$ and $R'a'bc$.

(iv) For $a',b' \in \mathbf{K}'$ and $c \in \mathbf{K}$: if $R'a'b'c$, then there are theories $a,b \in \mathbf{K}$ such that $a' \subseteq a$ and $b' \subseteq b$ and $Rabc$.

Proof.

Ad (i). Define a set

$$\Delta := \{D : (\exists BC)(D \to .B \to C \in \mathbf{L} \ \& \ B \in b' \ \& \ C \notin c)\}.$$

Claim (a): Δ is closed under disjunction. Suppose $D_1, D_2 \in \Delta$. Then there are B_1, B_2 and C_1, C_2 such that ($i \in \{1,2\}$)

(1) $D_i \to .B_i \to C_i$

and

(2) $B_i \in b'$, hence, $B_1 \& B_2 \in b'$

(since b' is adjunctive) and

(3) $C_i \notin c$, hence, $C_1 \vee C_2 \notin c$

(since c is prime). From (1) it follows that

(4) $(D_1 \vee D_2) \to (B_1 \to C_1) \& (B_2 \to C_2) \in \mathbf{L}$

whence by

(5) $(B_1 \to C_1) \& (B_2 \to C_2) \to .B_1 \& B_2 \to C_1 \vee C_2 \in \mathbf{L}$,

(6) $D_1 \vee D_2 \to .B_1 \& B_2 \to C_1 \vee C_2 \in \mathbf{L}$.

Thus it follows from (2), (3) and (6) by the definition of Δ that $D_1 \vee D_2 \in \Delta$ as required.

Claim (b): $a' \cap \Delta = \varnothing$. Suppose for some D that $D \in a'$ and $D \in \Delta$. Then there must be B,C such that $D \to .B \to C \in \mathbf{L}$ and $B \in b'$ and $C \notin c$. Since $D \in a'$ and a' is an $\mathbf{L}$-theory, $B \to C \in a'$. But since $R'a'b'c$ by hypothesis, $C \in c$ – contradiction.

From (a) and (b) we may now infer, using C3.7.(i), that there is a theory $a \in \mathbf{K}$ such that $a' \subseteq a$ and $a \cap \Delta = \varnothing$.

Claim (c): $R'ab'c$. Suppose $B \to C \in a$ and $B \in b'$. Since $a \cap \Delta = \varnothing$, $B \to C \notin \Delta$. It thus follows by the definition of Δ that for all B', C', if

$$B \to C \to B' \to C' \in \mathbf{L} \text{ and } B' \in b',$$

then $C' \in c$. So using Identity, $B \to C \to B \to C$, we have $C \in c$ as required.

Ad (ii). We first extend c' to a prime **L**-theory c. Let

$$\Delta_c := \{B\}.$$

Clearly, since c' is an **L**-theory and $B \notin c'$, $c' \nvdash \Delta_c$. It follows by L3.6 that there is a theory $c \in \mathbf{K}$ such that $c' \subseteq c$ and $B \notin c$.

Next we extend b' to a prime **L**-theory b. Let

$$\Delta_b := \{D : (\exists C)(D \to C \in a \ \& \ C \notin c)\}.$$

Claim (a): Δ_b is closed under disjunction. Suppose $D_1, D_2 \in \Delta_b$. Then there are C_1, C_2 such that

(1) $(D_1 \to C_1) \& (D_2 \to C_2) \in a$ and

(2) $C_1 v C_2 \notin c$.

Since

(3) $(D_1 \to C_1) \& (D_2 \to C_2) \to D_1 v D_2 \to C_1 v C_2 \in \mathbf{L}$,

it follows from (1) that

(4) $D_1 v D_2 \to C_1 v C_2 \in a$.

Hence, from (2) and (4) in virtue of the definition of Δ_b, $D_1 v D_2 \in \Delta_b$.

Claim (b): $b' \cap \Delta_b = \varnothing$. Suppose for reductio that some $D \in b'$ and $D \in \Delta_b$. Then for some $C \notin c$, $D \to C \in a$. But we have $R'ab'c'$ whence $C \in c' \subseteq c$ – contradiction.

With premisses (a) and (b) at hand we may now apply C3.7.(i) to infer the existence of a superset $b \in \mathbf{K}$ of b' such that $b \cap \Delta_b = \varnothing$.

Finally we check whether $Rabc$. Suppose $A \to C \in a$ and $A \in b$. Since $b \cap \Delta_b = \varnothing$, $A \notin \Delta_b$. So, by the definition of Δ_b, if $A \to C \in a$, as it is the case, then $C \in c$ as required.

Ad (iii). Define a set

$$\Delta := \{D : (\exists B)(A \to B \in a' \ \& \ B \notin c)\}.$$

Claim (a): Δ is closed under disjunction. Suppose $D_1, D_2 \in \Delta$. Then there are B_1, B_2 such that ($i \in \{1,2\}$)

(1) $D_i \to B_i \in a'$ and (2) $B_i \notin c$.

Thus, from (1) and (2) respectively, we may infer that

(3) $D_1 v D_2 \to B_1 v B_2 \in a'$,

since a' is an L-theory, and

(4) $B_1 \vee B_2 \notin c$,

since c is prime. Thus, there exists a B, viz. $B_1 \vee B_2$, such that $D_1 \vee D_2 \to B \in a'$ (by (3)) and $B \notin c$ (by (4)), which is to say that $D_1 \vee D_2 \in \Delta$.

Claim (b): $b' \cap \Delta = \emptyset$. Suppose for reductio that (for some A) $A \in \Delta$ and $A \in b'$. Then, in virtue of the definition of Δ, there must be some $B \notin c$ such that $A \to B \in a'$. But by hypothesis, $R'a'b'c$ whence $B \in c$ – contradiction.

We can now use C3.7(i) as before to infer from (a) and (b) the existence of a theory $b \in \mathbf{K}$ such that $b' \subseteq b$ and $b \cap \Delta = \emptyset$.

Claim (c): $R'a'bc$. Assume $A \to B \in a'$ and $A \in b$. Since $b \cap \Delta = \emptyset$, $A \notin \Delta$. Hence, $(\forall B)(A \to B \in a' \supset B \in c)$ and so $B \in c$ as required.

Ad (iv). Make the assumption and use (i) to obtain the antecedent part of (iii). Applying then (iii) gives the desired result. ■

LEMMA 3.9. Inclusion

For any sets a, b in the universe $\mathbf{K}$ of a canonical model of type [R*] for a logic **L**:

$$a \rhd b \text{ if and only if } a \subseteq b.$$

Proof. The left-to-right direction is trivial, using the fact that $A \to A \in \mathbf{L}$. For the converse suppose $a \subseteq b$. We need to show that there is some $x \in \mathbf{0}$ such that $Rxab$, i.e.

$$(\forall B \in \mathrm{Wff})(A \to B \in x \ \& \ A \in a \supset B \in b).$$

Let x be the logic **L**. Then clearly $R'\mathbf{L}ab$ since $a \in \mathbf{K}$ and $a \subseteq b$. So we have $R'\mathbf{L}ab$ for L-theory **L** and $a, b \in \mathbf{K}$. Thus, by L3.8.(i), there is a prime L-theory x such that $Rxab$ and $\mathbf{L} \subseteq x$, and in virtue of the latter conjunct, x is indeed a saturated L-theory, i.e. $x \in \mathbf{0}$. ■

Corollary 3.7.(ii) gives us already the result that the canonical **0** and **K** are well defined. Before turning to the question as to whether a canonical model satisfies the conditions on R, $*$ and V, we check now whether the canonical star function is well defined.

LEMMA 3.10. Star

For any set a in the universe $\mathbf{K}$ of a canonical model of type [R*] for a logic **L**: $a^* \in \mathbf{K}$.

Proof. We need to show that on the assumption that $a \in \mathbf{K}$, a^* is closed under (a) L-implication and (b) adjunction, and that (c) a^* is prime.

Ad (a). Suppose $A \in a^*$, whence $\sim A \notin a$, and $A \to B \in \mathrm{L}$, whence $\sim B \to \sim A \in \mathrm{L}$. Then $\sim B \notin a$ and so $B \in a^*$.

Ad (b). Suppose $A, B \in a^*$. Then $\sim A \notin a$ and $\sim B \notin a$ whence $\sim A \vee \sim B \notin a$. But $\sim(A\&B) \to \sim A \vee \sim B \in \mathrm{L}$. Hence, $\sim(A\&B) \notin a$, i.e. $A\&B \in a^*$.

Ad (c). Suppose $A \vee B \in a^*$ and $A \notin a^*$ and $B \notin a^*$. Then $\sim A, \sim B \in a$, so $\sim A \& \sim B \in a$. But $\sim A \& \sim B \to \sim(A \vee B) \in \mathrm{L}$. Hence, $\sim(A \vee B) \in a$, i.e. $A \vee B \notin a^*$ – contradiction. ■

LEMMA 3.11. R*

The relation R and the function * in a canonical model of type [R*] for a logic L satisfy the following conditions (for any $a, b \in \mathbf{K}$).

r1. $a \rhd a$

r2. if $a \rhd b$ and $Rbcd$, then $Racd$

*1. if $a \rhd b$ then $b^* \rhd a^*$.

Proof. In view of the inclusion lemma L3.9, r1 is trivial.

For r2 assume

(1) $a \rhd b$

and $Rbcd$, i.e. for arbitrary formulae A, B,

(2) if $A \to B \in b$ and $A \in c$ then $B \in d$.

Suppose further for some formulae C, D that

(3) $C \to D \in a$ and (4) $C \in c$.

We need to show that $D \in d$. It follows from (1) by L3.9 that $a \subseteq b$. Hence (from (3)), $C \to D \in b$, and so, from (2) and (4), $D \in d$.

For *1 suppose that $a \rhd b$ whence (L3.9) $a \subseteq b$. By L3.9 it suffices to show that $b^* \subseteq a^*$. So suppose $B \in b^*$. Then $\sim B \notin b$ whence $\sim B \notin a$ and so $B \in a^*$. ■

LEMMA 3.12. Valuation

The forcing relation $\models$ in a canonical model of type [R*] for a logic L satisfies the conditions (h), (p), $(\sim)$, $(\&)$, $(\vee)$, and $(\to)$.

Proof. (p) and $(\sim)$ hold by definition. (h) follows immediately from the inclusion lemma L3.9. The right-to-left direction of $(\&)$ follows from the adjunctiveness of members of **K**; for the converse use A1 and A2 respectively. Dually, the left-to-right direction of $(\vee)$ follows from the primeness of sets in **K**; for the converse use A4 and A5 respectively. The left-to-right direction of ($\rightarrow$) is also easy to prove, using the definition of R.

For the right-to-left direction of ($\rightarrow$) we contrapose and assume that $A \rightarrow D \notin a$. We need to find two sets $b, c \in \mathbf{K}$ such that $Rabc$, $A \in b$ and $D \notin c$. Define

$$b' := \{B : A \rightarrow B \in \mathbf{L}\} \text{ and } c' := \{C : (\exists B)(B \rightarrow C \in a \ \& \ B \in b')\}.$$

Claim (a): $b' \in \mathbf{K}'$. For closure under adjunction suppose $B_1, B_2 \in b'$. Then $(A \rightarrow B_1) \& (A \rightarrow B_2) \in \mathbf{L}$ whence $A \rightarrow B_1 \& B_2 \in \mathbf{L}$ and so $B_1 \& B_2 \in b'$. For closure under **L**-implication, suppose $B_1 \in b'$ and $B_1 \rightarrow B_2 \in \mathbf{L}$. Then $A \rightarrow B_1 \in \mathbf{L}$, hence, by transitivity, $A \rightarrow B_2 \in \mathbf{L}$, i.e. $B_2 \in b'$.

Claim (b): $c' \in \mathbf{K}'$. For closure under adjunction suppose $C_1, C_2 \in c'$. Then there are $B_1, B_2 \in b'$ such that

(1) $B_1 \rightarrow C_1 \in a$ and $B_2 \rightarrow C_2 \in a$.

Since b' is adjunctive,

(2) $B_1 \& B_2 \in b'$.

Since a is an **L**-theory, it follows from (1) that

(3) $B_1 \& B_2 \rightarrow C_1 \& C_2 \in a$.

Hence, from (2) and (3) by the definition of c', $C_1 \& C_2 \in c'$. For closure under **L**-implication, suppose (1) $C_1 \in c'$ and (2) $C_1 \rightarrow C_2 \in \mathbf{L}$. It follows from (1) that there is some $B \in b'$ such that $B \rightarrow C_1 \in a$. It follows from (2) by PREF that $B \rightarrow C_1 \rightarrow . B \rightarrow C_2 \in \mathbf{L}$, hence, $B \rightarrow C_2 \in a$ for $B \in b'$, i.e. (by the definition of c') $C_2 \in c'$.

Claim (c): $R'ab'c'$. Suppose $B \rightarrow C \in a$ and $B \in b'$ and yet $C \notin c'$. It follows from the latter assumption by the definition of c' that if $B \rightarrow C \in a$ then $B \notin b'$ – contradiction.

Claim (d): $A \in b'$ and $D \notin c'$. That $A \in b'$ follows trivially from the definition of b'. Suppose then that $D \in c'$. Then for some $B \in b'$, $B \rightarrow D \in a$, i.e. (1) $A \rightarrow B \in \mathbf{L}$ and (2) $B \rightarrow D \in a$. It follows from (1) by SUFF that $B \rightarrow D \rightarrow . A \rightarrow D \in \mathbf{L}$ whence, in virtue of (2), $A \rightarrow D \in a$, contradicting our hypothesis.

Putting (a) to (d) together, the condition for part (ii) of the priming lemma L3.8 obtains. Hence, there are $b, c \in \mathbf{K}$ ($b' \subseteq b$, $c' \subseteq c$) such that $Rabc$, $A \in b$, and $B \notin c$. ∎

As explained above, the proof of T3.3 (Completeness) is now completed.

4. Some further relevant logics

In this section we shall consider various extensions of the system **M**. We shall indicate how these extensions can be modelled by means of suitably constrained R*-models. The systems to be presented in this section will be defined by adding to the above axiomatisation of **BM** axioms and rules from the following list of key postulates.

A11. (DN-Elimination, DNE)	$\sim\sim A \to A$
A12. (DN-Introduction, DNI)	$A \to \sim\sim A$
A13. (Conjunctive Syllogism, WB)	$(A \to B)\&(B \to C) \to .A \to C$
A14. (Excluded Middle, X)	$A \vee \sim A$
A15. (Reductio, Rd)	$A \to \sim A \to \sim A$.
A16. (Contraposition, Cp)	$A \to B \to . \sim B \to \sim A$
A17. (Prefixing, B)	$A \to B \to .C \to A \to .C \to B$
A18. (Suffixing, CB)	$A \to B \to .B \to C \to .A \to C$
A19. (Contraction, W)	$(A \to .A \to B) \to .A \to B$
A20. (Permutation, C)	$(A \to .B \to C) \to (B \to .A \to C)$
A21. (Mingle, M)	$A \to .A \to A$
A22. (M3)	$A \vee (A \to B)$
A23. (Weakening, K)	$A \to .B \to A$
ER. (E-rule, ER)	$\dfrac{A}{A \to B \to B}$

Some of the logics produced by adding selected postulates from this list to **BM** are:

B	**BM** + DNE + DNI
G	**B** + X
DW	**B** + Cp
DJ	**DW** + WB
DK	**DJ** + X
DL	**DK** + Rd
TW	**DW** + B + CB
TJ	**TW** + WB
TK	**TJ** + X
TL	**TK** + Rd
T	**TL** + W

EW	**TW + ER**
EWX	**EW + X**
EWR	**EW + Rd**
E	**EWX + Rd**
RW	**EW + C**
RWX	**RW + X**
RWK	**RW + K**
R	**RWX + Rd**
RM	**R + M**
RM3	**RM + M3**
K	**R + K** [3]

It is, of course not claimed that these axiomatisations are irredundant. Thus, e.g., as soon as **BM** is extended by the schemas B, CB, or Cp, the corresponding rule forms, primitive in **BM**, become redundant. So do the two DeMorgan axioms A8 and A9 in the presence of both double negation axioms A11 and A12. There are also many alternative axiomatisations of the systems just defined. For example, the set {CP,A11,A12} may be replaced by A11 together with the DNI-suppressing contraposition rule $A \rightarrow {\sim} B \;/\; B \rightarrow {\sim} A$. And addition of any of the following principles to **RW** produces **R**:

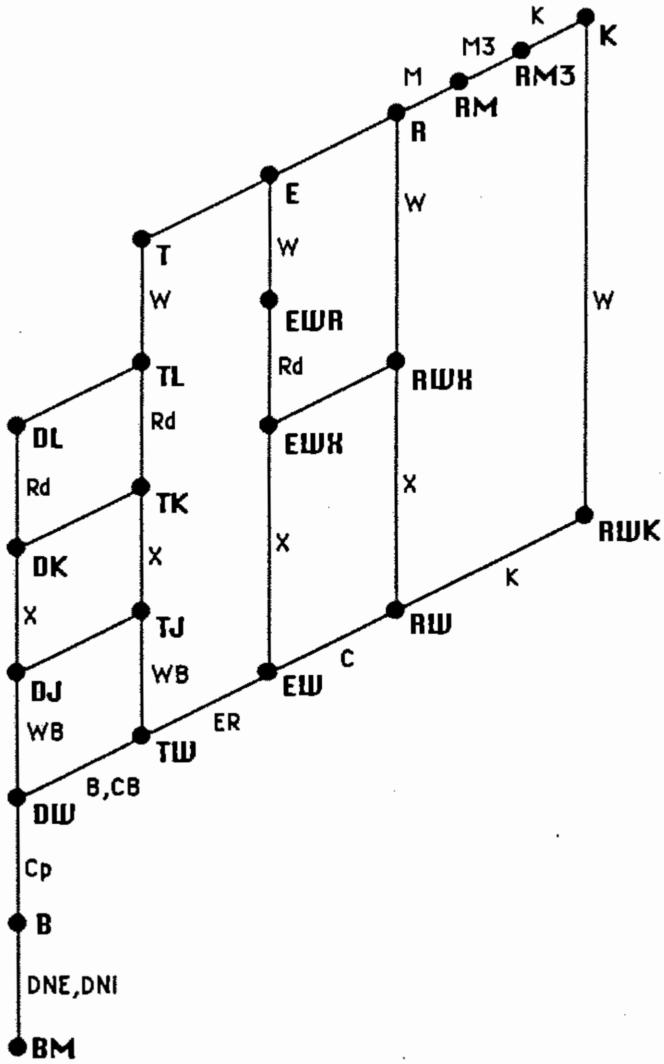
Rd.	$A \rightarrow {\sim} A \rightarrow {\sim} A$
W.	$(A \rightarrow .A \rightarrow B) \rightarrow (A \rightarrow B)$
S.	$A \rightarrow (B \rightarrow C) \rightarrow .(A \rightarrow B) \rightarrow .(A \rightarrow C)$
WI.	$A \& (A \rightarrow B) \rightarrow B$
WB.	$(A \rightarrow B) \& (B \rightarrow C) \rightarrow .A \rightarrow C$
WC.	$(A \rightarrow .B \rightarrow C) \rightarrow .A \& B \rightarrow C$,

In fact, many of these principles are equivalent in **R**, as are permutation of premisses, C, and assertion,

CI. $\quad A \rightarrow .A \rightarrow B \rightarrow B$.[4]

[3] All of these logics have made their appearance in the relevant literature; hence, we shall not engage here in a detailed discussion of the motivations behind them or their relative merits and demerits. Most of the systems are discussed in Routley, Meyer et al. (1982). For the D-systems see also the articles by Brady, and for contraction-free logics see Slaney (1980).

The (proper) inclusion relations among these systems are summarised in the diagram below (if $\mathbf{L}_1$ is connected with $\mathbf{L}_2$ by an upward path, then $\mathbf{L}_1$ is theoremwise included in $\mathbf{L}_2$).



A few remarks on the significance of some of these systems may be appropriate. The logic **K**, the strongest system in our list, is the classical propositional calculus. Just short of **K** are the semi-relevant logics logics **RM3** and **RM**. Though neither of the outright paradoxes

[4] These results are well-known and proofs will therefore be omitted. Standard references for these and further results about "axiom chopping" and the inclusion relations among most of the systems just defined, are Anderson and Belnap (1975), Routley, Meyer et al. (1982). For the D-systems see Brady (1985) and for W-free systems see Slaney (1980).

$$A \to .B \to A \qquad\qquad A \to .{\sim}A \to B$$
$$A \to B \vee {\sim}B \qquad\qquad A \& {\sim}A \to B$$

are theses of either of these logics, they do allow the derivation of somewhat milder irrelevancies, like ${\sim}(A \to A) \to .B \to B$ and $A \& {\sim}A \to B \vee {\sim}B$. As decent theories of implication, these logics are thus ruled out. However, **RM** and **RM3** are *paraconsistent* in the sense that the closure under any of these logics of an inconsistent set of sentences is non-trivial, i.e. a proper subset of the set of all well-formed formulae of the language under consideration. If one's main concern is with paraconsistency, then in particular **RM3** is highly recommendable. For, **RM3** is determined by a set of three-valued matrices:

$\to$	0	1	2	$\sim$
0	0	2	2	2
1	0	1	2	1
2	0	0	0	0

Designated values: 0 and 1; $x\&y = min(x,y)$, $x \vee y = max(x,y)$.

Thus, deciding theoremhood for **RM3** involves only marginally more work than deciding whether a formula is a two-valued tautology. This makes **RM3** just about the simplest paraconsistent logic on the market. Lest one should think that **RM3** is "*ad hoc*", we point out that the matrix values may be given interpretations which make **RM3** look good as a paraconsistent logic. Think of 0, 1, 2 as the values *true, false,* and *true and false* respectively. We leave it to the reader to decide whether the values assigned to compound formulae according to the above matrices make sense on this interpretation.

The logics **T**, **E**, and **R** are the principal systems investigated in Anderson and Belnap (1975). While **R** is their favoured theory of *relevant implication,* **E** was put forward as a theory of *entailment,* combining relevant implication with necessity. Anderson and Belnap (in [Ent]) have motivated their ideas about implication and entailment by analysing proofs in a Fitch-style natural deduction system. The system **T** of *ticket entailment* is one result of such an analysis: it is the result of constraining the rule for reiterating premisses in a subordinate proof so as to conform intuitively to an idea of Gilbert Ryle's that a strict distinction has to be drawn between tickets for inferences (implicational formulae) and facts used to cash in such tickets.[5]

[5] See Anderson and Belnap (1975), pp. 44ff. A less elusive motivation, apparently also going back to Anderson, is given in Dunn (1986), p.127.

Contraction free relevant logics are of much interest because they hold the promise for a non-trivial naive set theory (i.e. with unrestricted comprehension schema) and a non-trivial truth semantics for languages closed under the truth schema

'*A*' is true-in-L if and only if '*A is true*' is true-in-L.

Both theories can be trivialised in either one of two ways. One may either form the Russell class by means of the unrestricted comprehension schema (respectively, create the liar paradox by means of the naive truth schema) and then produce a contradiction (using Rd and MP); closure under a logic containing *ex falso quodlibet* will result in a trivial theory. Alternatively, we may use the contraction axiom W together with MP (or just the rule form of contraction) to produce triviality by an argument due to Curry.[6]

Triviality arguments of the first kind are blocked by adopting a logic in which EFQ and its easily recognizable cognates cease to be theorems. However, it turns out that Curry type paradoxes survive in even severly cut back relevant logics like **RWX**.[7] The strongest result to date, due to Brady (198+), is that naive set theory based on (the quantificational extension of) **DK** is absolutely consistent.

We turn now to our basic logic **BM**. This system, though hitherto not unknown, has been somewhat neglected. In building up a lattice of relevant logics, the start is usually made with **B**.[8]

But, first, from a semantic point of view, **BM** is a more natural choice of a basic system than **B**. Waiving any of our default conditions r1, r2, or *1 in the definition of an R*-frame, or the heredity condition (h) on R*-models, we obtain frames and models that validate only the empty set of formulae. In this sense **BM** – determined by the class of all R*-models, is the smallest (non-trivial) logic that can be modelled by means of the semantic methods used in this thesis.

Secondly, it will be recalled that **B** results from **BM** by adding both double negation axioms, DNI and DNE. The basic system **BM** does not include much that is objectionable from an intuitionist point of view: none

[6] See Curry (1942).

[7] See Slaney (1980). The question as to whether Curry-type paradoxes can be produced for naive set theory or semantics based on any of the systems **EW**, **EWX**, **EWR**, **RW** or **RWK** is still open.

[8] As in Routley, Meyer et al. (1982). In the projected volume two of *Relevant Logics and Their Rivals*, however, will be based on **BM** as a minimal relevant logic.

[8] Gabbay (1976), chapter 15. However, the consequence relation determined by such frames is non-trivial.

of the key offenders, excluded middle, double negation elimination, classical reductio, or classical contraposition are theorems. Thus, although full **BM** includes the intuitionistically unacceptable DeMorgan principle DMk, $\sim(A\&B)\rightarrow\sim A\vee\sim B$, the $\{\rightarrow,\sim\}$-fragment of **BM** is a subintuitionistic system. The basic system **BM** is therefore not only a point of departure for subintuitionist relevant logics in the $\{\rightarrow,\sim\}$-fragment of **La** but gives also rise to a branch of "quasi-intuitionist" relevant logics.

We shall now show how to extend the determination result of the last section for **BM** to cover logics defined by adding to **BM** any combination of postulates taken from the above list. Let **L** be such a logic. For the soundness result we shall need to impose additional constraints p_1, p_2, ... on R*-frames, for each new postulate P_1, P_2, ... added to **BM**. Thus, we prove the soundness of **L** with respect to that subclass of R*-frames that satisfies the new conditions. This calls for an extension of the R*-lemma and/or the star lemma in the completeness argument for **L**: we now have to ascertain that a canonical model of type [R*] for **L** does indeed belong to the newly defined subclass of R*-models; naturally, the new axioms (or rules) will be used to verify that the canonical models satisfy the new constraints on models. If proof-theoretic postulates and modelling conditions thus "fit", we shall say that they correspond to each other. We make this notion of correspondence precise as follows.

DEFINITION 4.1. Correspondence

Let **L** be a logic, let P be a proof-theoretic postulate (i.e. either an axiom or a rule), let p be a condition on R*-frames, and let M_{L} be a canonical model of type [R*] for **L**. Then P corresponds to p if and only if

(a) if an R*-model satisfies p, then P is true in M (soundness), and

(b) if P is an axiom of **L** (a rule of **L**), then M_L satisfies p (completeness).

Thus suppose $\mathbf{L} = \mathbf{BM}+P_1+...+P_n$. If we can show that P_1 corresponds to p_1 and ... and p_n corresponds to P_n, then we have in effect extended the determination result for **BM** to the result that **L** is determined by the class of all R*-frames satisfying the conditions p_1 to p_n. Conditions on R*-frames corresponding to axiom schemas A11 to A23 and the rule ER are displayed in the next theorem.

THEOREM 4.2. Correspondence

(We define:

d2.	$R^2abcd := (\exists x)(Rabx \,\&\, Rxcd)$ and
d3.	$R^2a(bc)d := (\exists x)(Raxd \,\&\, Rbcx)$.)
(i)	A11 corresponds to $0x \supset x^{**} \rhd x$
(ii)	A12 corresponds to $0x \supset x \rhd x^{**}$
(iii)	A13 corresponds to $Rabc \supset R^2a(ab)c$
(iv)	A14 corresponds to $0x \supset x^* \rhd x$
(v)	A15 corresponds to Raa^*a
(vi)	A16 corresponds to $Rabc \supset Rac^*b^*$
(vii)	A17 corresponds to $R^2abcd \supset R^2a(bc)d$
(viii)	A18 corresponds to $R^2abcd \supset R^2b(ac)d$
(ix)	A19 corresponds to $Rabc \supset R^2abbc$
(x)	A20 corresponds to $R^2abcd \supset R^2acbd$
(xi)	A21 corresponds to $Rabc \supset a \rhd c \vee b \rhd c$
(xii)	A22 corresponds to $0x \,\&\, a \rhd b \supset a \rhd x$
(xiii)	A23 corresponds to $Rabc \supset a \rhd c$
(xiv)	ER corresponds to $(\exists x)(0x \,\&\, Raxa)$.

Proof. Except for the first two correspondences, a full proof of this theorem (and proofs of further correspondences) can be found in Routley, Meyer et al. (1982), section 4.4. The first two correspondences are easy to prove: for completeness use the definition of the canonical star function together with DNE and DNI respectively and then the already established identity condition $a \rhd a$. To give a flavour of the verification of correspondences, we illustrate the proof of (vi).

Soundness. Suppose (1) $a \vDash A \rightarrow B$. We need to show that $a \vDash {\sim}B \rightarrow {\sim}A$. So assume

(2) $Rabc$ and (3) $b \vDash {\sim}B$

and for reductio

(4) $c \nvDash {\sim}A$.

It follows from (1) that

(5) $(\forall xy)(Raxy \,\&\, x \vDash A \supset y \vDash B)$.

From (2) we have (6) Rac^*b^* and from (3) and (4),

(7) $b^* \nvDash B$ and (8) $c^* \vDash A$.

Hence, from (5), (6) and (8):

(8) $b^* \vDash B$

in contradiction to (7).

Completeness. Assume $Rabc$. In virtue of the definition of the canonical relation R, this means that

(1) $C \rightarrow D \in a$ & $C \in b \supset D \in c$ for any formulae C, D.

We need to show that Rac^*b^*. So assume

(2) $A \rightarrow B \in a$ and (3) $A \in c^*$

and for reductio $B \notin b^*$, i.e.

(4) $\sim B \in b$.

Since a is a theory closed under a logic **L** which *ex hypothesi* includes all instances of the contraposition schema, it follows from (2) that

(5) $\sim B \rightarrow \sim A \in a$

whence, in virtue of (1), (5) and (4), $\sim A \in c$, i.e., by the definition of the canonical *-function, $A \notin c^*$, contradicting (3).

Correspondences where the consequent part of the modelling condition postulates the existence of certain points require more work in the completeness half of the argument: we need to construct prime theories with the required relational properties. Since we have already illustrated the method of constructing such theories in L3.8, we refer the reader to Routley, Meyer et al. (1982), L4.4, for the particular constructions required for the continuation of this proof. ■

Let C be a class of R*-frames. If a logic **L** is determined by C, then we shall say that a frame F in C is an R*(L)-frame (or, where there is no danger of ambiguity, an L-frame). Similarly, if M is a model induced on a frame in C, then M will be said to be an R*(L)-model (L-model). Just as there are equivalent axiomatisations of logics, so there are equivalent characterisations of classes of frames. In particular in view of completeness arguments, it would be cumbersome to carry a heavy baggage of conditions on frames, gathered by building up logics and sets of modelling conditions from **BM** in a step-by-step fashion. Fortunately, both economical axiomatisations and characterisations of classes of frames are available for the logics defined in this section. As an example of how much pruning can be done, we provide a reaxiomatisation of the system

R together with a simple definition of an R-frame.

Postulates for **R**

L-Simplification (&I).	$A \& B \to A$
R-Simplification (&I).	$A \& B \to B$
&-Composition (&C).	$(A \to B) \& (A \to C) \to A \to B \& C$
L-Addition (vI).	$A \to A v B$
R-Addition (vI).	$B \to A v B$
v-Composition (vC).	$(A \to C) \& (B \to C) \to A v B \to C$
Distribution (Dist).	$A \& (B v C) \to (A \& B) v (A \& C)$
DN-Elimination (DNE).	$\sim\sim A \to A$
Contraposition (Cp).	$A \to \sim B \to . B \to \sim A$
[Identity (I).	$A \to A$
[Suffixing (B).	$A \to B \to . B \to C \to A \to C$
[Prefixing (CB).	$A \to B \to . C \to A \to . C \to B$
[Contraction (W)	$(A \to . A \to B) \to . A \to B$
[Importation (W*).	$(A \to . B \to C) \to . A \& B \to C$
[Permutation (C).	$(A \to . B \to C) \to (B \to . A \to C)$
[Assertion (CI).	$A \to . (A \to B) \to B$
Modus Ponens (MP).	$\dfrac{A\ ,\ A \to B}{B}$
Adjunction (ADJ).	$\dfrac{A\ ,\ B}{A \& B}$

Deleting Identity and choosing any one axiom from each of the angle-bracketed pairs (B,CB), (W,W*), (C,CI) will result in a set of independent postulates for R. (Of course, many more such alternative pairs are conceivable.)

Following Dunn (1986), an R*(R)-frame may be defined as an R*-frame satisfying the following conditions on R and *.

r1. (Identity)	$a \rhd a$
r2. (Monotonicity)	$Rbcd\ \&\ a \rhd b \supset Racd$
r3. (Associativity)	$R^2abcd \supset R^2a(bc)d$

r4. (Idempotence) $Raaa$

r5. (Commutativity) $Rabc \supset Rbac$

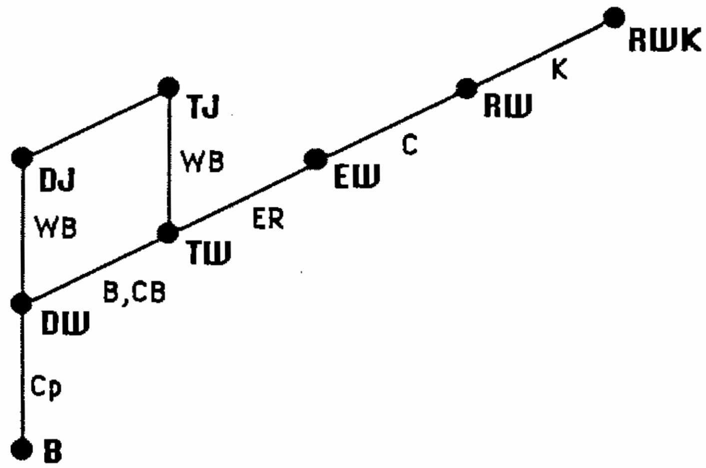
*2. (Inversion) $Rabc \supset Rac^*b^*$

*3. (Involution) $a^{**} = a$

Remark on reduced frames. A reduced R*-frame is an R*-frame satisfying the condition

0. (Reduction) **0** is a singleton set.

Accordingly, when specifying a reduced R*-frame, the set $\mathbf{0} \subseteq \mathbf{K}$ may be replaced by a distinguished point $0 \in \mathbf{K}$, whereupon the reduction condition 0 becomes redundant. Reduced frames were the kind of structures used to provide semantics for **R** and weaker positive relevant logics in Routley and Meyer's "The semantics of entailment" ((1973), (1972a), (1972b)). However, the kind of completeness argument adopted in these papers, does not extend to logics that are not a supersystem of **TW**+WI (called **C** in Routley, Meyer et al. (1982)) in the full connective set of **La**. (The hangup is located in the right-to-left direction of the inclusion lemma: in the case of unreduced models we needed to show that members of **K** are closed under **L**-implication. For reduced models we need to show that members of **K** are closed under 0-implication. But in order to ascertain this property of members of the canonical model set **K**, the logic **L** must contain the schemas Cp, B, CB and WI.) In particular, the argument does not extend – at least not in any straightforward way – to **RWK** or any of the systems in the above diagram that are weaker than **R**. Slaney (1987) has since shown that the original original completeness argument of Routley and Meyer argument can be amended so as to make reduced modelling available for logics without WI. With respect to the logics introduced in this chapter, Slaney has proved the completeness with respect to reduced R*-frames of all logics in the following fragment of the diagram displayed earlier:



Addition of Excluded Middle to these logics forfeits the reduced modelling property. (Slaney does not consider the system **BM**. But we conjecture that a reduced modelling theorem holds for **BM** too.)

Our reasons for preferring unreduced R*-frames will not emerge in detail until section 8. There it will turn out that one is faced with a difficulty in principle when attempting to enrich in a *straightforward* way R*-frames with a binary relation S to interpret a unary modal operator. The difficulty can be bypassed by sticking to unreduced frames. Moreover, the unreduced frames of section 8 are quite powerful: they allow a relational modelling of modal logics whose classical counterparts fall outside the scope of Kripke style frames (and, instead, are usually provided with neighbourhood semantics).

Chapter II

Modal extensions of relevant logics

5. The language $\mathbf{La}^{\Box}$

The background to all our considerations in this chapter will be provided by a language $\mathbf{La}^{\Box}$ which extends **La** by a unary (necessity) operator $\Box$. Thus, the set of primitive connectives of **La** is $\{\sim,\Box,\&,\vee,\rightarrow\}$. As an intuitive reading of formulae of the form $\Box A$ we suggest: *necessarily A*. A further (possibility) operator $\Diamond$ may be defined as usual, i.e.

D$\Diamond$. $\quad\quad \Diamond A := \sim\Box\sim A$.

An expression of the form $\Diamond A$ may be read: *possibly A*.

Nothing in this part hinges on these suggested readings. In particular, we leave it open, whether the necessity involved is of a logical, physical, deontic, or temporal kind. Claims to the effect that the box-operator as characterised in a particular modal logic captures the inferential hallmarks of a modality about which we may have pre-theoretical intuitions, will not be made until part two of this thesis.

By a *modal logic* we shall mean any formal system in the language $\mathbf{La}^{\Box}$ that extends our basic relevant logic **BM**. Thus, as a special case – of not much interest though – **BM**, when based on $\mathbf{La}^{\Box}$, is itself a modal logic.

The modal logics in this thesis are thus specified two-dimensionally, as it were. Any particular modal logic is, first, an extension of some (non-modal) relevant logic **L** and, secondly, characterised by a set of specifically modal postulates. Most of the results in this chapter will be schematic: they will be valid not for a particular modal logic, that is, an extension of a particular non-modal logic by a particular set of modal postulates, but for a class of modal logics, that is, for any extension of some non-modal logic satisfying certain conditions by some set of modal postulates satisfying certain other conditions. Unless indicated otherwise, in this chapter we shall mean by a (non-modal) logic any one of the formal systems presented in chapter one.

6. Relational semantics: R*S-frames and models

As in the case of **La**, the modal language $\mathbf{La}^{\Box}$ will be interpreted by means of certain set-theoretical structures. The basis of the determination results for the logics in the language **La**, discussed in chapter one, with respect to certain classes of R*-frames, were two facts: (a) to each connective in **La** there corresponded a certain operation on the universe set **K** of an R*-frame, and (b) these operations had properties "corresponding" to the inferential properties of the connectives in **La**, as laid down in the postulates for the various logics considered. Thus, the simple set-theoretic operations $\cup$ and $\cap$ corresponded to the Boolean connectives $\vee$ and $\&$, and the relation R and the function $*$ corresponded to the intensional connectives $\rightarrow$ and $\sim$ respectively. In order to interpret the language $\mathbf{La}^{\Box}$, we shall match the connective $\Box$ with a binary relation S in the semantics. Determination results will be forthcoming by giving S just the "right" properties. Since modal logics are, by definition, mere extensions of non-modal logics in a richer language, one should hope and expect that the semantics for modal logics are, in as straightforward a sense as possible, mere *extensions* of the semantics for non-modal logics. That is to say, the R*-part of the semantics should remain intact and the definition of a model should just be extended by a valuation clause for the new connective $\Box$.

In order to interpret the modal language $\mathbf{La}^{\Box}$, let us then extend R*-frames by adding a relation $S \subseteq \mathbf{K}^2$. Where F is an R*-frame, we shall say that $\langle F, S\rangle$ is an R*S-frame, i.e. an R*S-frame is represented by a quintuple $\langle 0, \mathbf{K}, R, S, *\rangle$, where – as for R*-frames – **K** is a non-empty set, $0 \subseteq \mathbf{K}$, $R \subseteq \mathbf{K}^3$, $*: \mathbf{K} \rightarrow \mathbf{K}$; and S is a binary relation between members of **K** (the modal accessibility relation).

An R*S-frame satisfies the conditions

for all $a \in \mathbf{K}$,

r1. (Identity)	$a \rhd a$,
r2. (Monotonicity)	$Rbcd \;\&\; a \rhd b \supset Racd$ and
*2. ($\rhd$-Inversion)	$a \rhd b \supset b^* \rhd a^*$

as for R*-frames and, in addition, the following condition on S:

d4.	$Sa := \{x : Sax\}$
s1. (S-Monotonicity)	$a \rhd b \supset Sb \subseteq Sa$

(The condition s1 will be needed in order to complete the inductive proof of the heredity lemma.)

A valuation function V mapping atomic sentences into sets of points induces a model <F,V> on an R*S-frame F. Again, the distribution of atomic sentences over members of **K** is subject to the atomic heredity condition,

(h) $\qquad a \triangleright b \ \& \ a \in V(p) \supset b \in V(p).$

The truth conditions for formulae with their principal connectives chosen from $\{\sim, \&, \vee, \rightarrow\}$ are as for R*-models. We define

DI |. $\qquad |A| := \{a : a \models A\}$

and add to the clauses (p), ($\sim$), ($\&$), ($\vee$), and ($\rightarrow$) a truth condition for formulae of the form $\Box A$:

($\Box$) $\qquad a \models \Box A$ iff $Sa \subseteq |A|$ [for all $a \in \mathbf{K}$, $A \in$ **Wff**].

Finally, the definitions of *truth* (at a point), *validity* (in a class of frames), and *entailment* (according to a model) are as for R*-models. That is, where a is a point in the universe set **K** of an R*S-model M, a formula A is true at a in M just in case $a \models A$ in M; A is true in M iff $x \models A$ for all $x \in \mathbf{0}$ in M; A is valid (in a class **C** of R*S-frames) iff A is true in all models induced on frames in **C**; and A entails B according to M iff $x \models B$ whenever $x \models A$, for all $x \in \mathbf{K}$ in M.

At the end of section 2 it has been noted that the heredity condition for arbitrary formulae is satisfied in an R*S-model, provided that we can extend the inductive proof of the heredity lemma L2.1, by the case for formulae of the form $\Box A$. We can:

LEMMA 6.1. Heredity

For any formula A and any points a and b in an R*S-model: if $a \models A$ and $a \triangleright b$, then $b \models A$.

Proof. Induction on the complexity of A. For sentential atoms and the connectives of **La**, the argument is the same as in L2.1. We need to consider one additional case, namely $A = \Box B$; for this case we shall use the inclusion condition s1. So assume

(1) $a \vDash \Box B$ and (2) $a \rhd b$.

It follows from (1) by the clause ($\Box$) that

(3) $Sa \subseteq |B|$.

Premiss (2) yields by s1 that

(4) $Sb \subseteq Sa$

whence in virtue of (3) we have $Sb \subseteq |B|$ which is to say (by ($\Box$)) that $b \vDash \Box B$, completing the inductive proof. ■

As for R*-models, the verification theorem follows from the heredity lemma in conjunction with the condition r1:

THEOREM 6.2. Verification

For any R*S-model M, A entails B according to M if and only if $M \vDash A \to B$ (i.e. for **0** in M: $(\forall x)(0x \supset x \vDash A \to B)$).

Thus armed we can now proceed to the soundness and completeness theorems of the next section.

7. C-modal logics and the basic completeness result

Let **L** a logic in the sense of chapter one: an extension of the basic system **BM** by any combination of the axioms A11 to A23 or the rule ER. For any such logic **L**, we define a system **L.C** as the smallest set of sentences in the language $\mathbf{La}^{\Box}$ that is closed under the following rules.

RL. if $\vdash_{\mathbf{L}} A$ then $\vdash_{\mathbf{L.C}} A$

MP. $$\frac{A\ ,\ A \to B}{B}$$

ADJ. $$\frac{A\ ,\ B}{A \& B}$$

RC. $$\frac{A_1 \& \cdots \& A_n \to A}{\Box A_1 \& \cdots \& \Box A_n \to \Box A} \quad (1 \leq n)$$

Where **L** is a logic, we shall say that **L.C** is the smallest *C-modal logic* based on **L**. In the sequel we shall frequently refer to "the" logic **L.C**. It will be a convenient fiction to treat '**L.C**' as a singular term. But, of course, whenever we make assertions about the system **L.C**, these assertions must be understood as universal generalisations with the (tacit) universal quantifier binding the *variable* 'L' in 'L.C'. That is, such assertions must be treated as assertions about any logic **L** extended by the rule RC. And *mutatis mutandis* for terms like '**L.E**', '**L.M**', '**L.KT4**', etc which will be used later.

Given the class of all R*(L)-frames, that is, the class of R*-frames determining a logic **L**, we now want to define a class of all R*(L)S-frames. Let F be an R*(L)-frame $<\mathbf{0},\mathrm{K},R,*>$ and let F' be a frame $<\mathbf{0}',\mathrm{K}',R',*',S>$. F' is an *S-extension of* F just in case $\mathbf{0}=\mathbf{0}'$, $\mathrm{K}=\mathrm{K}'$, $R=R'$, $*=*'$, and S is a relation on $\mathbf{K}^2$ satisfying the condition s1. The class of all R*(L)S-frames is the class of all S-extensions of all R*(L)-frames. Thus, *in nuce*, an R*(L)S-frame adds to an R*(L)-frame just a binary relation S constrained by the condition s1, leaving everything else as it is. Consequently, for any non-modal formula A, if A is true in an R*(L)-model M, then A is true in all S-extensions of M.

THEOREM 7.1. Soundness

For any logic **L**: if A is a theorem of **L.C**, then A is valid in the class of all R*(L)S-frames. In particular, if $\vdash_{\mathbf{BM.C}}$ then $\models_{R*S}A$.

Proof. Given the verification theorem T6.2, the verification of the non-modal postulates of **L** is as for R*(L)-models. (To verify postulates that are not theorems of **BM**, use the soundness direction of the relevant correspondence schema listed in T4.2.) It remains to show that the rule RC preserves truth at all points in **0**. (According to the definition D| |, we may rewrite $a\models A_1$ & $\cdots$ & $a\models A_n$ as $a\in|A_1|\cap\cdots\cap|A_n|$.) So assume that the premiss of RC holds throughout **0**, i.e.

(1) $|A_1|\cap\cdots\cap|A_n|\subseteq|A|$.

Assume secondly that $a\models\Box A_1\& \cdots \&\Box A_n$ for an arbitrary point $a\in\mathbf{K}$, i.e.

(2) $a\in|\Box A_1|\cap\cdots\cap|\Box A_n|$,

which, by ($\Box$), is to say that

(3) $Sa\subseteq|A_1|\cap\cdots\cap|A_n|$.

It follows from (1) and (3) that

(4) $Sa\subseteq|A|$, i.e. $a\models\Box A$

as required. ■

The completeness argument for a C-modal logic **L.C** simply extends the argument for non-modal logics in section 3. We shall define the notion of a canonical model which refutes some non-theorem of **L.C**. We shall then observe that such a canonical model is an R*(L)S-model. Hence, we conclude that we may construct for any non-theorem D of **L.C** an R*(L)S-model such that $M \nVdash D$ – or, contrapositively:

THEOREM 7.2. Completeness

If A is valid in the class of all R*(L)S-frames, then A is a theorem of **L.C**.

Now for the details.

DEFINITION 7.3. Canonical model of type [R*S] for **L.C**

A canonical model of type [R*S] for a C-modal logic **L.C**, is a sextuple

$$M_{\mathbf{L.C}} = \langle \mathbf{0}_{\mathbf{L.C}}, \mathbf{K}_{\mathbf{L.C}}, R_{\mathbf{L.C}}, S_{\mathbf{L.C}}, {}^*_{\mathbf{L.C}}, V_{\mathbf{L.C}} \rangle,$$

(subscripts **L.C** from now on omitted) such that

(a) **K** is a nonempty set of prime **L.C**-theories;

(b) $\mathbf{0} \subseteq \mathbf{K}$ is a set of saturated **L.C**-theories such that for each non-theorem D of **L.C**, there is a theory $x \in \mathbf{0}$ with $D \notin x$;

(c) $R \subseteq \mathbf{K}^3$ such that
$Rabc$ iff $(\forall A, B \in \mathbf{Wff})(A \rightarrow B \in a \ \& \ A \in c \supset B \in c)$;

(d) $*: \mathbf{K} \rightarrow \mathbf{K}$ such that
$a^* = \{A : {\sim}A \notin a\}$;

(e) $S \subseteq \mathbf{K}^2$ such that
Sab iff $(\forall A \in \mathbf{Wff})(\Box A \in a \supset A \in b)$;

(f) $V: \mathbf{At} \rightarrow 2^{\mathbf{K}}$ such that
$a \in V(p)$ iff $p \in a$;
V is extended to a relation $\models \subseteq \mathbf{K} \times \mathbf{Wff}$ such that
$a \models A$ iff $A \in a$.

As before, in order to show that the just defined canonical models exist and that they fall into the class of R*(L)S-models, we need to ascertain six facts:

(i) **K** is nonempty.

(ii) **0** is nonempty.

(iii) The function * is a mapping from **K** into **K**.

(iv) R and * satisfy the conditions required for R*(L)-models.

(v) S satisfies the condition s1.

(vi) The canonical valuation satisfies the heredity condition (h) and $\models$ satisfies the truth-conditions for complex formulae.

We can be brief: As **L.C** is a modal logic, **L.C** extends, by definition of 'modal logic' the system **BM**. Hence, the arguments in section 3 apply without modification to discharge claims (i) to (iii), and, for the logic **BM.C**, also claim (iv). For C-modal logics based on logics stronger than **BM**, we use the correspondence theorem T4.2 in order to verify (iv). New work needs to be done in verifying (v) and (vi).

LEMMA 7.4.

The relation S in a canonical model of type [R*S] for **L.C** satisfies the condition

s1. $\qquad a \rhd b \supset Sb \subseteq Sa.$

Proof. Assume that (1) $a \rhd b$ and that (2) Sbc. We need to show that Sac, i.e. that if $\Box A \in a$ then $A \in c$, for some formula A. So suppose further that (3) $\Box A \in a$. Now we use the inclusion lemma

$$a \rhd b \text{ iff } a \subseteq b$$

for canonical models of type [R*S] for C-modal logics (proof is as for L3.9) to infer from (1) and (3) that $\Box A \in b$ whence, by (2), $A \in c$ as required. ■

Finally, we show that the valuation function V satisfies the valuation clauses defining an R*S-model. The argument for the heredity condition and the valuation clauses for the connectives of **La** is given in L3.13. In order to extend the argument to the box-operator, we need to prove another priming lemma.

LEMMA 7.5. Priming

Let **L** be a C-modal logic, let a be a prime **L**-theory, and define $b' := \{B \in \mathbf{Wff}: \Box B \in a\}$. Then

(i) b' is an **L**-theory;

(ii) there exists a superset b of b' such that b is a prime **L**-theory, and for any formula A such that $A \notin b'$, $A \notin b$.

Proof. Ad (i). For closure under adjunction suppose that $B_1, B_2 \in b'$. Then $\Box B_1, \Box B_2 \in a$ and since a is an **L**-theory, $\Box B_1 \& \Box B_2 \in a$. Now, in virtue of the rule RC,

$$\Box B_1 \& \Box B_2 \rightarrow \Box(B_1 \& B_2)$$

is a theorem of any C-modal logic **L**. Hence, $\Box(B_1 \& B_2) \in a$ and so, by the definition of b', $B_1 \& B_2 \in b'$. For closure under **L**-implication, suppose $B_1 \in b'$ and $B_1 \rightarrow B_2 \in \mathbf{L}$. From the latter conjunct we derive by a special case of RC that $\Box B_1 \rightarrow \Box B_2 \in \mathbf{L}$. From the former conjunct it follows (by the definition of b') that $\Box B_1 \in a$ and so, since a is an **L**-theory, $\Box B_2 \in a$, i.e. $B_2 \in b'$. Thus, b' is an **L**-theory.

Ad (ii). Given that b' is an **L**-theory, the prime extension lemmma L3.6 guarantees the existence of a prime **L**-theory b such that $b' \subseteq b$. Furthermore, if $A \notin b'$, then b' may be extended to a prime **L**-theory b such that $A \notin b$ (by the corollary C3.7.(i) to L3.6). ■

LEMMA 7.6.

The relation $\models$ in a canonical model of type [R*S] for **L.C** satisfies the condition

($\Box$) $\qquad a \models \Box A$ iff $Sa \subseteq |A|$

Proof. ($\Rightarrow$) The proof of this direction of ($\Box$) is trivial.

($\Leftarrow$) Assume

(1) $\forall x \in \mathbf{K}(Sax \supset A \in x)$

and, for reductio,

(2) $\Box A \notin a$.

Define $b' := \{B: \Box B \in a\}$. It follows from the definition that $A \notin b'$. So, by the preceding priming lemma, there exists a prime **L.C**-theory b such that $b' \subseteq b$ and $A \notin b$. Note also that for all $\Box B \in a$, $B \in b' \subseteq b$; hence, Sab. But then (by (1)) $A \in b$ – contradiction. ■

The proof of the last lemma concludes the completeness argument for **L.C**.

8. C-modal logics in context and semantics for some extensions of L.C

In the first part of this section, the class of C-modal logics will be situated in a wider context of classes of modal logics. Various extensions of **L.C** will be considered in the second part; we shall provide the essential prerequisites for proving the completeness of such extensions in the form of a correspondence theorem, relating additional modal postulates to conditions on the accessibility relation S.

8.1. A classification of modal logics

By a (single-conclusion) deducibility relation $\vdash_{\mathbf{L}}$ determined by a logic **L** we mean a set of ordered pairs, $2^{\mathbf{Wff}}\times\mathbf{Wff}$, such that $\langle\Gamma,A\rangle\in\vdash_{\mathbf{L}}$ just in case A is deducible from Γ according to **L** (and in that case we write '$\Gamma\vdash_{\mathbf{L}}A$'). The collection to the left of the turnstile is called 'the premisses' and the formula on the right-hand-side of the turnstile is called 'the conclusion'. Just now we shall refrain from enquiring into the exact meaning of the relation of deducibility; in a moment, however, two possible candidate explications will emerge. For the time being, it will suffice to notice that a deducibility relation $\vdash_{\mathbf{L}}$ is a *metalogical* relation, specifying which transitions from premisses to conclusion are sanctioned by **L** as good inferences.

When, for example, proving facts about $\vdash$ by induction, three cases concerning the cardinality of the set of premisses stick out: (1) the premise set is empty; (2a) there is exactly one premise; (2b) the premise set contains one or more premisses. Where the logic under consideration is modal, the distinction between these three cases issues in three basic properties which consequence relations determined by modal logics may or may not have.[1]

(nec) if $\vdash A$ then $\vdash\Box A$;

(mon) if $A\vdash B$ then $\Box A\vdash\Box B$;

(reg) if $A_1,\ldots,A_n\vdash B$ then $\Box A_1,\ldots,\Box A_n\vdash\Box B$ $(n\geq 1)$.

In addition, we list a fourth property:

[1] The following conditions are due to Scott (1971). The congruence condition below is taken from Bull and Segerberg (1984).

(cgr) if $A \dashv\vdash B$ then $\Box A \dashv\vdash \Box B$.

We now return to the question as to what we should mean by 'the sentence A is deducible (according to **L**) from the set of sentences Γ'. (To make the point of our *purely heuristic* considerations, we assume that sets of premisses are *finite*.) By way of an answer, we suggest the following: B is deducible from $A_1, \ldots, A_n$ according to **L** just in case **L** "says so", that is to say, just in case a representation C of the fact that $A_1, \ldots, A_n \vdash_{\mathbf{L}} B$ is a theorem of **L**, i.e. $\vdash_{\mathbf{L}} C$. But how do such deducibility facts about **L** get represented as theorems in **L**? Here are two candidates.

Ded.1 $A_1, \ldots, A_n \vdash^1{}_{\mathbf{L}} B$ iff $A_1 \& \cdots \& A_n \rightarrow B \in \mathbf{L}$;

Ded.2 $A_1, \ldots, A_n \vdash^2{}_{\mathbf{L}} B$ iff $A_1 \rightarrow . \cdots \rightarrow . A_n \rightarrow B \in \mathbf{L}$;
$(n \geq 0)$.

If the logic **L** is the classical propositional calculus **K**, then there is not much to choose between Ded.1 and Ded.2. In virtue of the classical equivalence

Exportation & Importation

$$(A \& B \rightarrow C) \leftrightarrow (A \rightarrow . B \rightarrow C),$$

$\vdash^1$ and $\vdash^2$ are equivalent. However, for logics which lack either one of Exportation or Importation, such as relevant logics, $\vdash^1$ and $\vdash^2$ are quite distinct. (But notice that Importation, $(A \rightarrow . B \rightarrow C) \rightarrow (A \& B \rightarrow C)$, is a theorem of the relevant logic **R**; hence, $\vdash^2{}_{\mathbf{R}}$ is a (proper) subset of $\vdash^1{}_{\mathbf{R}}$.) Where **L** is a relevant logic, this distinction can be brought out in Churchian terms (Church (1951)) as follows. While $A_1, \ldots, A_n \vdash^2{}_{\mathbf{L}} B$ expresses the fact that B may be deduced from $A_1, \ldots, A_n$, *using* all of the premisses $A_1, \ldots, A_n$ – that there is a *relevant deduction* of B *from* $A_1, \ldots, A_n$ – $\vdash^1$ need not satisfy a use-of-premisses condition. As a consequence, $\vdash^1$ is *monotonic*: if $\Gamma \vdash^1 A$ holds, so does $\Delta \vdash^1 A$, for any $\Delta \supseteq \Gamma$; by contrast, $\vdash^2$ is non-monotonic.

The distinction between $\vdash^1$ and $\vdash^2$ necessitates a differentiation in the clause (reg) above. We need to distinguish between (reg1) if $A_1, \ldots, A_n \vdash^1 B$ then $\Box A_1, \ldots, \Box A_n \vdash^1 \Box B$ $(n \geq 1)$.

(reg2) if $A_1, \ldots, A_n \vdash^2 B$ then $\Box A_1, \ldots, \Box A_n \vdash^2 \Box B$ $(n \geq 1)$.

(Clearly, for (nec), (mon), and (cgr), such a differentiation would be otiose.)

We now represent the conditions (nec), (mon), (reg1), (reg2), and (cgr) as possible rules of a modal logic, generating theorems from theorems.

RN. (Necessitation) $\dfrac{A}{\Box A}$.

RM. (Monotonicity) $\dfrac{A \to B}{\Box A \to \Box B}$

RC. (Conjunctive Regularity) $\dfrac{A_1 \& \cdots \& A_n \to A}{\Box A_1 \& \cdots \& \Box A_n \to \Box A}$ $(n \geq 1)$

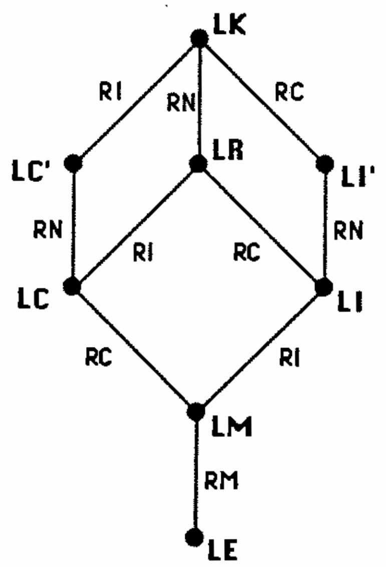
RI. (Implicative Regularity) $\dfrac{A_1 \to . \cdots \to A_n \to A}{\Box A_1 \to . \cdots \to \Box A_n \to \Box A}$ $(n \geq 1)$

RE. (Congruence) $\dfrac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$

A modal logic is

congruential, if closed under RE,
monotonic, if closed under RM,
C-regular, if closed under RC,
I-regular, if closed under RI,
regular, if both C-regular and I-regular,
necessitative, if closed under RN,
C-normal, if both C-regular and necessitative,
I-normal, if both I-regular and necessitative,
normal, if both regular and necessitative.

Where **L** is a logic (in the sense of chapter one), **L.E** denotes the smallest extension of **L** that is closed under the rule RE (**L.E** is the smallest congruential modal logic based on **L**). Similarly, **L.M, L.C, L.I, L.R, L.C′, L.I′, L.K,** and **L.N** denotes the smallest monotonic, C-regular, I-regular, regular, C-normal, I-normal, normal, and necessitative modal logic based on **L** respectively. The inclusion relations among these systems are depicted below. (The smallest necessitative system **L.N** is omitted from the diagram as the system is of little importance and would "disturb" the picture.)



These inclusion relations are immediate reflections of the definitions of the systems and need no further justification. The split at the **L.M**-node is justified in view of the preceding discussion concerning the rules RC and RI. However, where **L** is a logic including all instances of both Exportation and Importation, like classical logic, **L.C=L.I=L.R** and **L.C′=L.I′=L.K** whence the (classically) familiar linear order of congruential, monotonic, regular and normal systems of modal logic emerges.

The pairs of rules RC and RN, and RI and RN can be merged into the two rules of C-normality and I-normality respectively:

RC′. (C-Normality) $$\frac{A_1 \& \cdots \& A_n \rightarrow A}{\Box A_1 \& \cdots \& \Box A_n \rightarrow \Box A} \quad (n \geq 0)$$

RI′. (I-Normality) $$\frac{A_1 \rightarrow. \cdots \rightarrow. A_n \rightarrow A}{\Box A_1 \rightarrow. \cdots \rightarrow \Box A_n \rightarrow \Box A} \quad (n \geq 0)$$

Though proof-theoretically quite elegant, the rules RC′ and RI′ are very cumbersome baggage in model-theoretic investigations. It is therefore good to know that these rules can be replaced by more manageable sets of postulates including the axiom schemas

$\Box$C. (C-regularity)	$\Box A \& \Box B \rightarrow \Box(A \& B)$ and
$\Box$I. (I-regularity)	$\Box(A \rightarrow B) \rightarrow . \Box A \rightarrow \Box B$.

THEOREM 8.1.

A modal logic is

(i) C-regular iff it is closed under RM and contains $\Box$C,

(ii) I-regular iff it is closed under RM and contains $\Box$I,

(iii) C-normal iff it is closed under RM and RN and contains $\Box$C,

(iv) I-normal iff it is closed under RN and contains $\Box$I,

(v) regular iff it is closed under RM and contains both $\Box$C and $\Box$I,

(vi) normal iff it is closed under RN and contains both $\Box$C and $\Box$I.

Proof. Propositions (v) and (vi) are immediate corollaries to (i) and (ii) and (iii) and (iv) respectively. The left-to-right directions of (i) to (iv) are proved by applying the rules RC (RI) to premisses $A \rightarrow B$ and $A \& B \rightarrow A \& B$ ($A \rightarrow B \rightarrow A \rightarrow B$). For the converse directions note first that we have

$$\Box A \& \Box B \leftrightarrow \Box(A \& B)$$

in all C-regular modal logics. (One half is of course $\Box$C; the other half follows by &E, RM, and ADJ.) Thus, by the associativity of conjunctions and replacement, we have

(*) $\quad \Box A_1 \& \cdots \& \Box A_n \rightarrow \Box(A_1 \& \cdots \& A_n)$.

To prove the right-to-left directions of (i) to (iv) it will suffice to show (a) that $\Box$C together with closure under RM implies closure under RC, (b) that $\Box$I together with closure under RM implies closure under RI, and (c) that $\Box$I and closure under RN implies closure under RM.

Ad (a). Assume

(1) $A_1 \& \cdots \& A_n \rightarrow B$.

It follows by RM that

(2) $\Box(A_1 \& \cdots \& A_n) \rightarrow \Box B$

whence by transitivity of provable implication from (*) and (2),

(3) $\Box A_1 \& \cdots \& \Box A_n \rightarrow \Box B$.

Ad (b). Assume

(4) $A_1 \rightarrow . \cdots \rightarrow A_n \rightarrow B$.

Thus, by RM,

(5) $\Box A_1 \rightarrow \Box(A_2 \rightarrow . \cdots \rightarrow A_n \rightarrow B)$.

The following formula is an instance of $\Box$I:

(6) $\Box(A_2 \rightarrow . \cdots \rightarrow A_n \rightarrow B) \rightarrow . \Box A_2 \rightarrow \Box(A_3 \rightarrow . \cdots \rightarrow A_n \rightarrow B)$.

Hence, by transitivity from (5) and (6), we obtain

(7) $\Box A_1 \rightarrow . \Box A_2 \rightarrow \Box(A_3 \rightarrow . \cdots \rightarrow A_n \rightarrow B)$.

After repeating the last two steps $n-3$ times, we shall eventually arrive at the required

(8) $\Box A_1 \rightarrow . \cdots \rightarrow \Box A_n \rightarrow \Box B$. ∎

8.2. Extensions of C-modal logics and their semantics

In sections 6 and 7 we have shown that L.C is determined by the class of all R*(L)S-frames. This result sets the lower bound for the scope of modelling afforded by the semantics in terms of R*S-frames to the class of C-modal logics. In section 9 we shall consider a more powerful modelling technique whose scope extends to congruential model logics, that is, extensions of the system L.E. For the remainder of this section we shall contend ourselves with extending the present modelling technique to C-modal systems enriched by modal postulates chosen from the list below.

RN. (Necessitation)	$\dfrac{A}{\Box A}$
$\Box$I. (Impl. Regularity)	$\Box(A \rightarrow B) \rightarrow . \Box A \rightarrow \Box B$
$\Box$T. ($\Box$-Elimination)	$\Box A \rightarrow A$
$\Box$D. (Consistency)	$\Box {\sim} A \rightarrow {\sim}\Box A$
$\Box$4. (LL-Expansion)	$\Box A \rightarrow \Box\Box A$
$\Box$B. (Brouwer)	$A \rightarrow \Box\Diamond A$
$\Box$5. (LM-Expansion)	$\Diamond A \rightarrow \Box\Diamond A$

Under the principal – necessity- – interpretation of the box-operator, $\Box$T expresses the triviality that what is necessarily true, is true *simpliciter*. But under other interpretations, such as the deontic "it is obligatory that ...", $\Box$T is far from being a truism and should indeed be rejected. When $\Box$ is interpreted as an obligation operator, $\Box$T is usually – and arguably erroneously – weakened to the schema □D which is derivable from □T (using CP and TRANS). On this deontic interpretation, □D requires that obligations be consistent. □B – or rather the equivalent $A \rightarrow \sim\Diamond\sim\Diamond A$ – is the so-called Brouwer'sche axiom. Contrary to what the name suggests, it was not introduced by Brouwer but by Becker (1930) as a principle about the iteration of the impossibility modality: if A is true, then it is impossible that A is impossible. Together with □T, □B allows to reduce modalities in a way reminiscent of "Brouwer's Rule": an uneven (even) iteration of absurdity (= $\sim\Diamond$) is equivalent to simple (double) absurdity.[2] □4, the distinctive postulate of C.I. Lewis's system **S4**, leads – in the presence of □T – to a similar reduction of modalities: iterated necessity is equivalent to simple necessity. □5 is the distinctive postulate of Lewis's system S5; it is perhaps best be viewed as a weakening of □B: for it to be true that it is impossible for A to be impossible, it suffices, according to □5, that A be possible, while, according to Brouwer, it suffices that A be true.

Let $\mathbf{L.CP_1 \cdots P_n}$ be a C-modal logic including some selection $P_1, \ldots, P_n$ from the above list of postulates. In order to prove that $\mathbf{L.CP_1 \cdots P_n}$ is sound and complete with respect to some class of R*S-models, we need to extend the argument of section 8 by proofs of two facts for each postulate P added to **L.C**. We need to find a modelling condition p for P and verify (a) that R*S-models that satisfy p, validate P, and (b) that a canonical model of type [R*S] for $\mathbf{L.CP_1 \cdots P_n}$ satisfies the condition p if $\mathbf{L.CP_1 \cdots P_n}$ contains the schema P or, if P is a rule, is closed under P respectively. In the terminology introduced earlier: we need to prove correspondences for each additional postulate.

THEOREM 8.2. Correspondence

(i) □I corresponds to
si. $\exists x(Rabx \,\&\, Sxc) \supset \exists yz(Say \,\&\, Sbz \,\&\, Ryzc)$.

(ii) RN corresponds to sn. $0a \,\&\, Sab \supset a \triangleright b$.

(iii) □T corresponds to st. Saa.

(iv) □D corresponds to sd. $\exists x(Sax^* \,\&\, Sa^*x)$.

(v) □4 corresponds to s4. $Sab \,\&\, Sbc \supset Sac$.

[2] See Becker (1930), footnote 2.

(vi) $\Box$B corresponds to sb. $Sab \supset Sb^*a^*$.

(vii) $\Box$5 corresponds to s5. $Sa^*c \;\&\; Sab \supset Sb^*c$.

Proof. As it will be expected, one uses the semantic condition to validate the corresponding proof-theoretic postulate (soundness) and, conversely, one uses the postulate in order to show that the canonical model under consideration satisfies the corresponding condition. In the cases of (ii), (iii), (v), (vi) and (vii), this task is easily accomplished and, hence, the proofs will be omitted. As the reader will expect, the soundness parts of (i) and (iv) are also easily established. Thus, we shall supply only the completeness-halves of (i) and (iv).

Ad (i). Assume that for some $x \in \mathbf{K}$,

(1) $Rabx$ and (2) Sxc.

Define

$$y' := \{A : \Box A \in a\} \text{ and } z' := \{B : \Box B \in b\}.$$

Claim (a): y' and z' are **L.CI**-theories. This follow by L7.5(i) immediately from the facts that **L.CI** is a C-modal logic and that a and b are prime **L.CI**-theories.

Claim (b): For any formula A: if $\Box A \in a$ then $A \in y'$, and if $\Box A \in b$ then $A \in z'$. Both implications hold in virtue of the definitions of y' and z' respectively.

Claim (c): For any formulae A, B: if $A \rightarrow B \in y'$ and $A \in z'$ then $B \in c$. Make the assumptions. Then

(3) $\Box(A \rightarrow B) \in a$ and (4) $\Box A \in b$.

Since a is *ex hypothesi* an **L.CI**-theory, it follows from (1) that

(5) $\Box A \rightarrow \Box B \in a$.

Using (1), we may infer from (4) and (5) that there is some theory x such that

(6) $\Box B \in x$

whence, by (2),

(7) $B \in c$.

We now apply L3.8(iv): in virtue of (a), (c) and the hypothesis that $c \in \mathbf{K}$, the antecedent condition of L3.8.(iv) is satisfied. Hence, there are $y, z \in \mathbf{K}$ such that $Ryzc$ and $y' \subseteq y$ and $z' \subseteq z$ whence, by (b), Say and Sbz, as required.

Ad (iv). Define $x' := \{A : \Box A \in a^*\}$. Since $a \in \mathbf{K}$, so is a^*. Thus, since a is *ex hypothesi* a prime **L.CD**-theory, so is a^*. It follows by L7.5.(i) that x' is an **L.CD**-theory. Moreover, it is immediate from the definition of x' that (for any formula A)

(1) $\Box A \in a^* \supset A \in x'$.

Next we show that

(2) $\Box A \in a \supset {\sim}A \notin x'$.

Assume that $\Box A \in a$. Then ${\sim}\Box A {\sim}\in a$ since a is closed under $\Box$D. Thus $\Box{\sim}A \notin a^*$ whence, by the definition of x', ${\sim}A \notin x'$.

We shall now construct a prime extension x of x' such that Sa^*x and Sax^*. Define $\Delta := \{A : \Box{\sim}A \in a\}$.

Claim (a): Δ is closed under disjunction. Suppose $A, B \in \Delta$. Then $\Box{\sim}A \in a$ and $\Box{\sim}B \in a$ whence $\Box{\sim}A \& \Box{\sim}B \in a$. Since a is an **L.CD**-theory, it follows that $\Box({\sim}A \& {\sim}B) \in a$ whence $\Box{\sim}(A \vee B) \in a$. So, by the definition of Δ, $A \vee B \in \Delta$.

Claim (b): $\Delta \cap x' = \varnothing$. Assume for some sentence A that $A \in \Delta$ and, for contradiction, that $A \in x'$. Then $\Box{\sim}A \in a$. Thus, it follows from (2) that $A \notin x'$ – contradiction.

By C3.7.(i), it follows from (a) and (b) that there is some theory x such that

(3) $x \in \mathbf{K}$,

(4) $x' \subseteq x$, and

(5) $\Delta \cap x = \varnothing$.

By the definition of S in a canonical model, it follows immediately from (1), (3) and (4) that Sa^*x. To show that Sax^*, suppose that $\Box A \in a$, so $\Box{\sim}{\sim}A \in a$. Then by the definition of Δ: ${\sim}A \in \Delta$. Hence, by (5), ${\sim}A \notin x$, i.e. $A \in x^*$ as required. ■

With the correspondence theorem T8.2 at hand, we can now extend the soundness and completeness result for **L.C** to C-modal logics **L.C**+P, where P denotes any set of postulates chosen from the list displayed at the beginning of this section. We need some convention for naming modal logics and adapt the conventions of Lemmon (1977) to our present needs. Given some basic modal system **L.U** (**U**$\in$ {**E,M,C,C′,K,I,I′,R**}) and axiom schemas $\Box\mathrm{P}_1$,..., $\Box\mathrm{P}_n$, $\mathbf{L.UP}_1 \cdots \mathbf{P}_n$ names the smallest extension of **L.U** by all instances of the schemas $\Box\mathrm{P}_1$,..., $\Box\mathrm{P}_n$. A thus name for a modal system thus generated is the *Lemmon code* for that system.[3] For example, **BM.K4** denotes the smallest normal modal logic

[3] The term 'Lemmon code' is borrowed from Bull and Segerberg (1984), p.20.

based on **BM** that contains all instances of the schema □4; and **L.K4** is a generic name for any smallest normal modal logic containing □4 based on some (non-modal) logic **L**. Sometimes synonyms for Lemmon codes are well entrenched in the literature and in such cases we reserve the right to use these synonyms, especially when the corresponding Lemmon code is rather long.

Only one of the *relevant* modal systems thus definable has, to my knowledge, so far appeared in print. This is the system **R.KT4** which may be axiomatised by adding to the postulates for **R** the pair of normality rules RC′ and RI′ and the schemas □T and □4. Equivalently (by T8.1) **R.KT4** results by adding to postulates for **R** the rule RN[4] together with the schemas □I, □C, □T, and □4. The latter formulation is found in Meyer (1966) and Routley and Meyer (1972a) where the system goes under the name of **NR**.[5]

There is a sense in which **R.KT4** should have been Anderson and Belnap's Official Theory of Entailment. For, according to Anderson and Belnap, entailment is necessary relevant implication, and while the authors favour **S4** as the correct theory of logical necessity, their champion theory of relevant implication is the system **R**. However, entailment according to **R.KT4** does not coincide with entailment according to **E**, Anderson and Belnap's favourite theory of entailment: the Minč-formula

$$(A \to (B \to C))\& (B \to A \vee C) \to B \to C$$

fails to be a theorem of **E**, while its appropriately translated counterpart $((A \Rightarrow B) := \Box(A \to B))$

$$(A \Rightarrow (B \Rightarrow C))\& (B \Rightarrow A \vee C) \Rightarrow B \Rightarrow C\text{, i.e.}$$
$$\Box(\Box(A \to \Box(B \to C))\& \Box(B \to A \vee C) \to \Box(B \to C))$$

is a theorem of **R.KT4**.[6] The authors of *Entailment* decided to "postpone, at least to Volume II, the decision as to whether to laugh or cry" (p.352). The present author is inclined to bid farewell to entailment in the sense of **E**. While the idea of entailment as necessary good (relevant) implication is simple and natural, the Fitch-style natural deduction system for **E** – Anderson and Belnap's principal vehicle for motivating **E** – is decidedly not. Therefore, *if* **R** provides the theory of implication and *if* the S4-

---

[4] Or the closure of all axiom schemas under □.

[5] To record yet another name: R.KT4 = NR is the system $\mathbf{R}^{\Box}$ appearing in Anderson and Belnap (1975) and Read (1988).

[6] Chidgey found a eight elements matrix for **E** rejecting the Minč-formula; see Anderson and Belnap (1975), p.352. A three-worlds model for **E** in which the Minč-formula fails is displayed in Read (1988).

modal axioms capture the hallmarks of logical necessity, then **R.KT4**, and not **E**, should be thought of as the logic of entailment.

The system **K.KT** is Fey's system **T** (von Wright's **M**); **K.KTB** is the "Brouwer'sche system" **B**; and **K.KT5** is better known under the name **S5**. All these systems have deontic cousins which result by weakening the schema □T to □D. Thus for example, **K.KD** is the minimal (or standard) deontic logic **D** in Chellas (1980), earlier introduced in Hansson (1971) as Standard Deontic Logic, **SDL**. Clearly, relevant versions of these better known modal logics may be obtained by basing the respective modal superstructure on a relevant logic **L** rather than on the classical propositional calculus **K**.

8.3. Reduced frames for R.KT4 ?

At the end of section 4 we have briefly mentioned that stronger relevant logics, that is, from **TW+WI** "upwards", can be proved sound and complete with respect to reduced R*-frames. Reduced R*-frames, it will be remembered, are R*-frames in which the set **0** of distinguished points is reduced to a single point. Similarly, an R*S-frame is said to be reduced just in case the set **0** in such a frame is a singleton set. We know, for example, that the system **R** can be modelled by means of reduced frames.[7] A natural question to ask, therefore, is whether the modal extensions of **R** considered in this section are determined by reduced R*-frames extended by a relation S (with appropriate constraints, including s1) in just the same way in which we earlier extended unreduced frames by a relation S. One would expect the answer 'Yes'. For, just as we have conservatively extended **R** by a set of modal postulates, so we should be able to extend the frames for **R** by whatever is required to interpret the extended language, leaving the R*-component of such frames completely undisturbed. Or, as we have put it earlier: just as we have grafted syntactically a theory of necessity onto **R**, so we should graft its semantic counterpart onto R*-frames and models. It turns out, however, that reduced R*-frames – unlike unreduced frames – are not stable with respect to modal extensions. If reduced R*-frames are to be transformed into a semantics for modal extensions of **R**, then that transformation cannot be a mere extension; changes in the R*-component of such frames will be called for, and, accordingly, the answer to our question at the beginning of this paragraph is 'No'. We shall demonstrate

[7] See for example Routley and Meyer (1973).

this by showing that the comparatively strong system **R.KT4** is incomplete with respect to the class of all reduced R*S-frames.

By a reduced R*S-frame, we mean an R*S-frame satisfying the condition

0. (Reduction) **0** is a singleton set.

We denote that single element of **0** by '0'.

THEOREM 8.3.

For any reduced R*S-model M, if $0 \vDash A \rightarrow B$ then $0 \vDash \Box A \rightarrow \Box B$.

Proof. Since both the heredity lemma L6.1 and the verification theorem T6.2 are true for all R*S-models, these results hold in particular for R*S-models satisfying the reduction condition 0. By the verification theorem we need to show that if $|A| \subseteq |B|$ then $|\Box A| \subseteq |\Box B|$,
i.e. $\forall x (x \vDash \Box A \supset x \vDash \Box B)$. Thus, suppose that

(1) $|A| \subseteq |B|$ and

(2) $a \vDash \Box A$.

Then $\forall y (Say \supset y \vDash A)$ from (2) and, hence (from (1)),

(3) $\forall y (Say \supset y \vDash B)$.

It remains to show that $a \vDash \Box B$, i.e. $\forall z (Saz \supset z \vDash B)$. Thus assume Sab. Then (from (3)) $b \vDash B$ as required. ■

A reduced R*(R)S-frame is an R*(R)S-frame (i.e. an R*S-frame satisfying the characteristic modelling conditions on R and $*$ for the system **R**) satisfying the reduction condition 0. The proof that **R** is sound with respect to the class of all reduced R*(R)S-frames is routine and, hence, omitted.

The formula

(T) $(p \rightarrow p \rightarrow .p) \vee (p \rightarrow {\sim}(p \rightarrow p))$.

is a theorem of **R**. (T) may be viewed as an instance of excluded middle in the form $(t \rightarrow A) \vee (A \rightarrow f)$ where the sentential constants t and f are defined as $A \rightarrow A$ and ${\sim}(A \rightarrow A)$ respectively. By the soundness theorem for **R** with respect to reduced R*(R)S-frames, we have

(1) $0 \vDash (p \rightarrow p \rightarrow .p) \vee (p \rightarrow {\sim}(p \rightarrow p))$

for the 0-point of an arbitrary reduced R*(R)S-model M. Hence, by the forcing condition ($\vee$) for disjunctive formulae,

(2) $0\models p\to p\to .p$ or $0\models p\to{\sim}(p\to p)$.

It follows now from the closure of 0 under the rule RM (T8.3) that

(3) $0\models\Box(p\to p)\to\Box p$ or $0\models\Box p\to\Box{\sim}(p\to p)$

whence, by (v) again,

(4) $0\models(\Box(p\to p)\to\Box p)v(\Box p\to\Box{\sim}(p\to p))$.

Thus, the formula

(F) $(\Box(p\to p)\to\Box p)v(\Box p\to\Box{\sim}(p\to p))$,

is valid in the class of all R*(R)S-frames and so in particular in the class of all R*(R)S-frames that satisfy the conditions si, sn, st, and s4 (R*(R)S(kt4)-frames).

When unreduced models are under consideration, the argument breaks down at step (3). Although **0** is "collectively" closed under $\Box$-monotonicity, i.e.

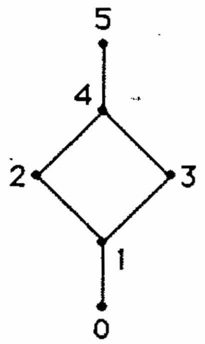
$$(\forall x\in\mathbf{0})(x\models A\to B)\supset(\forall x\in\mathbf{0})(x\models\Box A\to\Box B),$$

it is not generally the case that each member of **0** is thus closed, i.e.

$$\text{not: }(\forall x\in\mathbf{0})(x\models A\to B\ \supset x\models\Box A\to\Box B),$$

as it would be required for the transition from (2) to (3).

The matrix set M6 below shows that (F) is *not* a theorem of **R.KT4**.



$\to$	0	1	2	3	4	5	$\sim$	$\Box$
0	5	5	5	5	5	5	5	0
1	0	4	4	4	4	5	4	0
2	0	3	4	3	4	5	3	0
3	0	2	2	4	4	5	2	3
4	0	1	2	3	4	5	1	4
5	0	0	0	0	0	5	0	5

Designated values: 4 and 5; $x\&y=min(x,y)$, $xvy=max(x,y)$.

R.KT4 is sound with respect to M6. But for A=1, (F) gets evaluated to the undesignated value 3.[8] Thus, we have found a formula, viz (F), which

[8] Both claims have been verified using Belnap and Chidgey's matrices-versus-formulae testing program TESTER (Version PGH-1980A). For the pair (T) and (F) and the lead to M6 I am indebted to Dr John Slaney.

is valid according to the definition of a reduced R*(R)S(kt4)-frame but fails to be provable according to **R.KT4**. So **R.KT4** is incomplete with respect to the class of all R*(R)S(kt4)-frames.

Routley and Meyer (1972a) have obtained a completeness result for **R.KT4** with respect to certain kinds of reduced frames. Their frames, however, differ significantly from our R*S-frames in that the heredity relation $\rhd$ is now "necessitated": $a \rhd b$ iff for some $x \in$ K, $S0x$ and $Rxab$. While this move may be seen as giving some semantic substance to the claim that entailment combines relevant implication with necessity, its success depends crucially on the fact that **R.KT4** is closed under RN and contains □T and □I; the semantics in Routley and Meyer (1972a) are ingeniously tailor-made for **R.KT4** (though they presumably extend to other normal modal logics containing the schema □T) and consequently lack the versatility of the notion of an R*S-frame as defined in this thesis.

9. Neighbourhood semantics: R*N-frames and models

In the preceding sections we have provided semantics for conjunctively regular modal logics (i.e. modal logics extending the system **L.C**). This leaves the question as to how to model modal logics that fall short of being closed under the rule RC, like the systems **L.E**, **L.M** and their implicatively regular extensions **L.I** and **L.I′**. An answer will be provided in this and the following two sections in terms of so-called *neighbourhood frames*.[9] After having defined the concept of a neighbourhood (R*N-) frame, we shall engage in a brief *excursus* about the relationship between R*N- and R*S-models.

An R*N-frame is a structure $<$**0**,**K**,R,N,$*>$, where **K**$\neq\varnothing$, 0$\subseteq$**K**, $R\subseteq$**K**^{3}, and $*$:**K**$\rightarrow$**K**, just as for R*-frames. The new element N (the "neighbourhood"-function) maps members of **K** into collections of subsets of **K** ("neighbourhoods"), i.e. N:**K**$\rightarrow 2^{2^{\mathbf{K}}}$.

Again we assume that R and $*$ are subject to the conditions familiar from the definition of an R*-frame. As a default condition for the function N, we add:

[9] Neighbourhood semantics for unary intensional operators were first publicised in Montague (1968) and Scott (1970). The classical study of neighbourhood semantics for modal logics based on classical logic **K** is Segerberg (1971).

n1. (Inclusion) $a \triangleright b \supset Na \subseteq Nb$.

An R*N-model is an R*N-frame F together with a valuation function $V\text{:At}\rightarrow 2^{\mathbf{K}}$ induced on F, such that the atomic heredity condition (h) is satisfied. The forcing clauses for the nonmodal connectives are as for R*-models and for modal formulae we define:

($\Box$) $a \models \Box A$ iff $|A| \in Na$.

The definitions of truth at a point and in a model, entailment according to a model, and validity on a class of frames are the usual ones.

Excursus. What is the relationship between relational and neighbourhood models? The concept of a neighbourhood model is not, in any obvious sense, a generalisation of that of a relational model. However, in an extensional sense, the class of all relational models can be identified with a certain subclass of neighbourhood models. This extensional sense in which relational models are special kinds of neighbourhood models is supplied by the notion of *pointwise equivalence*: two models M_1 and M_2 are pointwise equivalent just in case there is a one-to-one mapping between the two model sets $\mathbf{K}_1$ and $\mathbf{K}_2$ such that for every formula A, A holds at the point $a_1 \in \mathbf{K}_1$ iff A holds at the corresponding point $a_2 \in \mathbf{K}_2$. The subclass of neighbourhood models which are in this sense equivalent to the class of relational models is the class of *augmented R*N-models.* An R*N-model is augmented just in case it satisfies the condition (for every point a and set of points X)

na. $X \in Na$ iff $\cap Na \subseteq X$.

THEOREM 9.1.

(i) Every R*S-model M_S is pointwise equivalent to some augmented R*N-model M_N and (ii) every augmented R*N-model M_N is pointwise equivalent to some R*S-model M_S.

Proof. (The theorem is a standard result in the literature on modal logics; full proofs may be found in Chellas (1980) (T7.9) and Segerberg (1971) (T2.8 and corollary). Thus we shall provide only a sketch of the argument which, we hope, will nevertheless be instructive.) To show (i) we define the model M_N induced by M_S as follows. M_N is just like M_S

except that the relation S is now replaced by an S-induced neighbourhood function:

(N) $\qquad X \in Na$ (in M_N) iff $Sa \subseteq X$ (in M_S).

It is easily verified that M_N is an R*N-model. In particular, since $\cap Na = Sa$ (by the definition), M_N is augmented. By induction on the complexity of formulae, we show that M_S and M_N are pointwise equivalent. The only case of interest in the induction arises when a formula is of the form $\Box A$, and that case is quickly dealt with using the definition (N).

For (ii) assume that we are given an augmented R*N-model M_N. The R*S-model M_S induced by M_N is just like M_S except that the function N is now replaced by an N-induced relation S:

(S) $\qquad Sab$ (in M_S) iff $b \in \cap Na$ (in M_N).

Again, we can show that M_S is an R*S-model. And by induction on the complexity of formulae we show that M_N and M_S are pointwise equivalent, the only non-trivial case being again formulae of the from $\Box A$:

$(M_N,a)\models\Box A$	iff $\lvert A \rvert \in Na$ in M_N	by ($\Box$) for R*N-models
	iff $\cap Na \subseteq \lvert A \rvert$ in M_N	by augmentation
	iff $(\forall x)(Sax \supset (M_S,b)\models A)$	by (S) and ind. hyp.
	iff $(M_S,a)\models\Box A$	by ($\Box$) for R*S-models. ■

Note that as a corollary to T9.1 and the soundness and completeness theorems for **L.C** with respect to the class of all R*S-models, we have thus obtained a determination result for **L.C** with respect to the class of augmented R*N-models. (*End of excursus*)

The proof of the heredity lemma for R*N-models builds as straightforwardly on the basic heredity lemma for R*-models (L2.1) as does the corresponding argument concerning R*S-models (L6.1).

LEMMA 9.2. Heredity

For any formula A and points a and b in an R*N-model: if $a\models A$ and $a \rhd b$, then $b \models A$.

Proof.

In addition to what we have shown in lemma 2.1, we need to consider one more case in order to complete the inductive proof: $A = \Box B$. As to be expected, we shall use the condition n1. Assume

(1) $a \vDash \Box B$ and (2) $a \rhd b$.

It follows by ($\Box$) from (1) that

(3) $|B| \in Na$

and by n1 from (2) that

(4) $Na \subseteq Nb$

whence we infer from (3) and (4),

(5) $|B| \subseteq Nb$

which, by ($\Box$) again, is to say that $b \vDash \Box B$. ■

Given the heredity lemma, we are entitled to assert the verification theorem:

THEOREM 9.3. Verification

For any R*N-model M, A entails B according to M if and only if $M \vDash A \to B$ (i.e. $(\forall x \in 0)(x \vDash A \to B)$).

10. The basic completeness result for congruential modal logics

A congruential modal logic is a modal logic closed under the rule

RE. (Congruence) $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$$

For any logic **L**, **L.E** is the smallest congruential modal logic based on **L**. In this section we shall prove that **L.E** is determined by the class of all R*(L)N-frames.

THEOREM 10.1. Soundness

If A is a theorem of **L.E**, then A is valid in the class of all R*(L)N-frames.

Proof. We extend the soundness result for **L** with respect to R*(L)-frames by showing that the rule RE preserves truth in an arbitrarily chosen R*(L)N-model M. Thus, assume that $M \vDash A \leftrightarrow B$. Then, by the verification theorem T9.3, $|A| = |B|$ in M. Hence, for any point $a \in M$, $|A| \in Na$ iff $|B| \in Na$, i.e. (by ($\Box$)) $a \vDash \Box A$ iff $a \vDash \Box B$ whence (again by T9.3) $M \vDash \Box A \leftrightarrow \Box B$. ■

In order to prove completeness, we start again by defining the notion of a canonical model. The canonical models to be defined presently, are simple extensions of the canonical R*-models of section 3. Thus, in verifying that our canonical models are indeed R*N-models, we can build again on the results already proved about canonical R*-models (and R*(L)-models).

Definition 10.2. Canonical model of type [R*N] for **L.E**

A canonical model of type [R*N] for a congruential modal logic **L.E** is a sextuple

$$M_{\mathbf{L.E}} = \langle \mathbf{0}_{\mathbf{L.E}}, \mathbf{K}_{\mathbf{L.E}}, R_{\mathbf{L.E}}, N_{\mathbf{L.E}}, {}^*{}_{\mathbf{L.E}}, V_{\mathbf{L.E}} \rangle$$

(subscripts **L.E** henceforth omitted) such that

(a) **K** is a nonempty set of prime **L.E**-theories;

(b) $\mathbf{0} \subseteq \mathbf{K}$ is a set of saturated **L.E**-theories such that if D is not a theorem of **L.E**, then there is a theory $x \in \mathbf{0}$ with $D \notin x$;

(c) $R \subseteq \mathbf{K}^3$ such that
$Rabc$ iff $(\forall A, B \in \mathbf{Wff})(A \rightarrow B \in a \;\&\; A \in b \supset B \in c)$;

(d) $^* : \mathbf{K} \rightarrow \mathbf{K}$ such that
$a^* = \{A : {\sim}A \notin a\}$;

(e) $N : \mathbf{K} \rightarrow 2^{2^{\mathbf{K}}}$ such that
$Na = \{X : (\exists A)(\Box A \in a \;\&\; X = |A|)\}$;

(f) $V : \mathbf{At} \rightarrow 2^{\mathbf{K}}$ such that
$a \in V(p)$ iff $p \in a$; $a \models A$ iff $A \in a$.

Given the results of sections 3 and 4 about canonical R*(L)-models, it will suffice to prove the following.

(i) The neighbourhood function N satisfies the condition n1.

(ii) Canonical models (as just defined) satisfy the valuation condition for formulae of the form $\Box A$.

(iii) The function N is well-defined.

Observe first that since **L.E** is an extension of **L**, the following inclusion lemma holds by the same argument as given for L3.9.

LEMMA 10.3. Inclusion

For any sets a, b in the universe K of a canonical model of type [R*N] for L.E:

$$a \rhd b \text{ iff } a \subseteq b.$$

Using this lemma, we now verify (i).

LEMMA 10.4.

The function N in a canonical model of type [R*N] for L.E satisfies the condition

n1. $$a \rhd b \supset Na \subseteq Nb.$$

Proof. Assume that

(1) $a \rhd b$ and (2) $X \in Na$.

It follows from (2) that there is some formula A such that

(3) $|A| = X$ and (4) $\Box A \in a$.

From (1) we may infer by L10.3 that

(5) $a \subseteq b$

whence, by (4),

(6) $\Box A \in b$.

Thus, combining (3) and (6), it follows by the definition of N that $X \in Nb$ as required. ■

Proposition (ii) is trivially true in virtue of the definition of the canonical neighbourhood function N. It remains to show that N is well-defined. We need to show that whenever two formulae A and B determine the same valuation set, i.e. $|A| = |B|$, and $|A| \in Na$, then $|B| \in Na$ (for arbitrary $a \in$ K). The result follows quite easily, once we have proved the following lemma.

LEMMA 10.5.

For any canonical model of type [R*N] for L.E:

if $|A| \subseteq |B|$, then $A \rightarrow B$ is a theorem of L.E.

We proceed on the assumption that the lemma has been proved (the proof will follow presently). To show that the definition D10.2(e) of N is correct, suppose that $|A| = |B|$. It follows by L10.5 that $\vdash A \rightarrow B$ and $\vdash B \rightarrow A$, i.e. $\vdash A \leftrightarrow B$ whence by RE, $\vdash \Box A \leftrightarrow \Box B$. Now make the second assumption, viz $|A| \in Na$. Then $\Box A \in a$ and hence, $\Box B \in a$, i.e. $|B| \in Na$ as required.

Proof of Lemma 10.5. Assume that

(1) $(\forall x \in \mathbf{K})(A \in x \supset B \in x)$

and, for reductio, that

(2) $\nvdash A \rightarrow B$.

It follows from (2) by the definition of $\mathbf{0}$ that there is some theory $a \in \mathbf{0}$ such that

(3) $A \rightarrow B \notin a$.

At this point we may interpolate the main part of the proof of L3.12: from the assumption (3) we can show that there exist sets $a, b \in \mathbf{K}$ such that

(4) $Rabc$

and

(5) $A \in b$ and (6) $B \notin c$.

Since $a \in \mathbf{0}$, it follows from (4) that

(7) $b \rhd c$

whence by L10.3,

(8) $b \subseteq c$.

So, from (5) and (8),

(9) $A \in c$.

But then, by (1),

(10) $B \in c$,

contradicting (6) as required. ■

The last lemma completes the argument that canonical models for **L.E**, as defined in D10.2, exist and that such models are R*(L)N-models. Since non-theorems of **L.E** fail to be true in canonical models (for each non-theorem there is always a point in $\mathbf{0}$ where it fails), it follows that **L.E** is complete with respect to the class of all R*(L)N-frames.

THEOREM 10.6. Completeness

If A is valid in the class of all R*(L)N-frames, then A is a theorem of **L.E**.

11. Extensions of LE

In the left column of table 11.1 we list a number of axioms and rules which may be added to a system L.E to produce, among others, the modal logics explicitly mentioned in section 8. Corresponding modelling conditions are exhibited in the right column. To state these modelling conditions in a reasonably concise way, we define

$$X^* := \{x : x^* \in X\}$$

and

$\hat{R}(Na)(Nb)(Nc) :=$
$[\forall YZ][\{x : (\forall yz)(Rxyz \ \& \ y \in Y \supset z \in Z)\} \in Na \ \& \ Y \in Nb \supset Z \in Nc]$.

Rule / Axiom			*Modelling condition*
RM.	$\dfrac{A \to B}{\Box A \to \Box B}$	nm.	$X \subseteq Y \ \& \ X \in Na \supset Y \in Na$
RN.	$\dfrac{A}{\Box A}$	nn.	$0 \subseteq X \supset (\forall x \in 0)(X \in Nx)$
$\Box$C.	$\Box A \& \Box B \to \Box(A \& B)$	nc.	$X \in Na \ \& \ Y \in Na \supset X \cap Y \in Na$
$\Box$I.	$\Box(A \to B) \to. \Box A \to \Box B$	ni.	$Rabc \supset \hat{R}(Na)(Nb)(Nc)$
$\Box$T.	$\Box A \to A$	nt.	$X \in Na \supset a \in X$
$\Box$D.	$\Box{\sim}A \to {\sim}\Box A$	nd.	$X \in Na^* \supset X^* \in Na$
$\Box$4.	$\Box A \to \Box\Box A$	n4.	$X \in Na \supset \{x : X \in Nx\} \in Na$
$\Box$B.	$A \to \Box\Diamond A$	nb.	$a \in X \supset \{x : X^* \in Nx^*\} \in Na$
$\Box$5.	$\Diamond A \to \Box\Diamond A$	n5.	$X^* \in Na^* \supset \{x : X^* \in Nx^*\} \in Na$

TABLE 11.1

The correspondence claims implicit in the above table consist again of two parts. Let P be one of the proof-theoretic postulates in the table and let p be the modelling condition for which it is claimed that it corresponds to P. For the soundness part, we have to show that P is true in every model that satisfies the condition p. For the completeness part we have to show that a canonical model for a congruential modal logic containing (or being closed under) P satisfies the condition p.

THEOREM 11.2. Correspondence

The proof-theoretic postulates and the semantic conditions as set out in table 11.1 stand in the correspondence relation to each other.

Proof. The soundness part is routine and, hence, omitted. Given the definition of the neighbourhood function in canonical models (being essentially the valuation clause for $\Box$ in R*N-models), and given the way the semantic conditions are "read off" the proof-theoretic postulates, the completeness part of the theorem is also easily proved. We illustrate the method of proof by providing the completeness halfs of the correspondences between RN and rn (a), and between $\Box$I and ri (b).

(a). Assume that $\mathbf{0} \subseteq |A|$. By the definition of the canonical set $\mathbf{0}$, it follows that $\vdash A$, hence, by RN, $\vdash \Box A$. Thus, since $\mathbf{0}$ is a set of regular theories, $\Box A \in x$ for all $x \in 0$. So, by the definition of the canonical neighbourhood function N, $|A| \in Nx$ for all $x \in \mathbf{0}$, as required.

(b). Assume that

(1) $Rabc$.

Spelling out $\hat{R}(Na)(Nb)(Nc)$, we may assume further that

(2) $\{x : (\forall yz)(Rxyz \;\&\; A \in y \supset B \in z)\} \in Na$

and that

(3) $|A| \in Nb$.

We need to show that $|B| \in Nc$. The canonical models considered here satisfy the forcing condition for $\rightarrow$-formulae. Thus we may replace (2) by

(4) $\{x : A \rightarrow B \in x\} \in Na$, i.e. $|A \rightarrow B| \in Na$.

Hence,

(5) $\Box(A \rightarrow B) \in a$.

By hypothesis, a is closed under a modal logic containing each instance of the schema $\Box$I, whence from (5):

(6) $\Box A \rightarrow \Box B \in a$.

It follows from (1) and (6) by the definition of the canonical relation R that

(7) $\Box A \rightarrow \Box B \in a \;\&\; \Box A \in b \supset \Box B \in c$.

The antecedent part of (7) is true in virtue of (6) and (3). Thus we can detach

(8) $\Box B \in c$

which is to say that $|B| \in Nc$. ■

The basic completeness result for **L.E** of section 10 combines with the above correspondence theorem to yield determination results for any extension of **L.E** by sets of postulates drawn from the list in table 11.1. In particular we thus obtain alternative semantics for C-modal logics, previously modelled by means of relational frames. For example, the smallest C-modal system (based on L), **L.C** = **L.EMC** = **L.MC** is determined by the class of all R*(L)N-frames in which the function N is closed under supersets (nm) and closed under (finite) intersections (nc).

Part Two
Applications

Chapter III

Modal logics for reasoning about theories

12. Formal properties of theories

We document the results of inquiry in the form of theories. Not all inquiry is directed towards nature or touchable artefacts. Philosophy, for one, is a discipline which, to a large extent, turns to theories themselves as objects of inquiry. In this, and the following chapter we shall investigate certain formal aspects of theories. In the present chapter, some of the modal systems presented in chapter two will be interpreted as formal metatheories of theories of various kinds as they present themselves at a particular point of time. In chapter four we shall focus on certain formal properties of theories along a time axis: how theories ought to change in the light of new evidence.

The principal aim in this chapter will be to argue for a move from classical to relevant modal logics when the modal operators are interpreted in certain ways. We shall consider in some detail two such interpretations: epistemic and deontic interpretations. It is well known that classical epistemic and deontic logics yield what appear to be paradoxical consequences. Some of these paradoxes, so I shall argue, can be resolved by adopting a relevant rather than classical basis. Others persist even when the move to relevant modal systems is made. Thus, for example, the schema

$$oA \rightarrow o\,(A \vee B\,)$$

(where o reads 'it ought to be that ...') will be derivable in the relevant logics of obligation proposed in section 15. But (*) may be instantiated to the *prima facie* paradoxical

> if it ought to be that you post the letter, then it ought to be that you post the letter or burn it.

This and similar "paradoxes" have been extensively discussed in the literature and I have nothing to add to that discussion. I share the view of the proponents of "Standard Deontic Logic" that Ross-type paradoxes and certain paradoxes of commitment are due to inadequate formalisation and that they can be satisfactorily dealt with by moving to a richer language in which distinctions can be drawn that cannot be drawn in the rather simple deontic language of the standard systems.[1] This view, however,

[1] A richer language in which Ross-type paradoxes disappear is the language of dynamic logic; see Segerberg (1980). Paradoxes of commitment, like the one displayed

will not be defended here.

In part one we have already used the term 'theory' as a generic term for certain linguistic objects: sets of sentences in a specified formal language satisfying certain closure conditions (see the definition of an **L**-theory, D3.4). This is the sense in which we shall go on to employ the term 'theory'; as it will emerge, it is a sense suitable for the purpose of the present investigation.

All of the kinds of theories considered here will be closed under logical equivalence: if the sentences A and B are logically equivalent, then A is a member of a theory T if and only if B is a member of T. This condition is by no means a trivial one. The set of sentences believed to be true by an agent at a particular time are arguable not thus closed, even if we choose our weakest logic **BM** to provide the notion of logical equivalence. For, in **BM**, A and $(A\&B)vA$ are logically equivalent. Let A be 'it is 38°C in Canberra'; let B be 'it is -12°C in Dnepropetrovsk'. John, living in Canberra, may very well believe that A. But it seems just wrong to take John's belief in A as a ground for ascribing to him the belief that $(A\&B)vA$. John may never have heard about Dnepropetrovsk in which case there is no reason to suppose that John is opinionated in any way about the temperature in Dnepropetrovsk. In general, we should not ascribe beliefs to an agent which involve concepts not to be found in the conceptual resources from which the agent in question forms his beliefs.

All of the theories considered here in any detail will be closed under logical consequence: if $A \rightarrow B$ is a theorem of our logic, then if A is in the theory, so is B. Again, sets of beliefs are not thus closed – even when we restrict the vocabulary of our language to expressions of whose meanings the believer in question does have a grasp. The reason for this failure is that actual agents are just imperfect reasoners: they frequently fail to recognise not only remote but also immediate consequences of their beliefs. But when belief is combined with a commitment-notion, the set of sentences an agent is committed to believe as true is closed under logical consequence. Such a set will be called an *acceptance set*. Acceptance sets – sets of sentences that are true according to an agent at a time – will be the topic of section 14. Sets of sentences that ought to be true according to an agent at a time will be studied in section 15. Such sets will be referred to as *norm sets* and these too will be closed

in Hansson (1971), pp. 132f., can be resolved by introducing a dyadic operator of conditional obligation; see e.g. Hansson (1971) and van Fraassen (1972). A relevant version of dyadic modal logic is presented in Fuhrmann (198+).

under logical consequence.

Closure under logical consequence is an important property of theories in the less general sense of 'theory' in which the term is used ordinarily. It is an epistemological fact that the acceptance of a theory proceeds via the acceptance of some surveyable exposition of that theory; as finite beings we cannot have an overt (as opposed to tacit) attitude to infinite sets of sentences. But once I accept a theory via such an expository base, what I thereby accept is, in a certain sense beyond my control. I am not at liberty to claim acceptance of, say, the philosophical doctrine of anti-realism – as set out, e.g., in the writings of Dummett – while rejecting one of its consequences, no matter how unexpected it may be, when presented with a proof. If I accept the exposition of a theory, then I thereby accept the theory as a whole.[2] Thus, there is a natural and important sense of 'acceptance' in which a set of sentences accepted as true is closed under logical consequence. 'Acceptance' in this sense is closely related to the notion of rational belief as it occurs in the literature on doxastic logics.[3]

We shall suppose – until further notice – that the theories we are dealing with can be formulated in the language **La**, having the connective set $\{\sim, \&, \vee, \rightarrow\}$. In terms of these connectives and their English counterparts we formulate now a number of general conditions theories may or may not satisfy.

(cmp) Completeness	if $A \notin T$ then $\sim A \in T$;
(cns) Consistency	if $\sim A \in T$, then $A \notin T$;
(adj) Adjunctivity	if $A \in T$ and $B \in T$, then $A \& B \in T$;
(dej) Dejunctivity	if $A \& B \in T$, then $A \in T$ and $B \in T$;
(add) Additivity	if $A \in T$ or $B \in T$, then $A \vee B \in T$;
(prm) Primeness	if $A \vee B \in T$, then $A \in T$ or $B \in T$;
(det) Detachedness	if $A \rightarrow B$, then, if $A \in T$ then $B \in T$.

[2] In ordinary language the distinction between the exposition of a theory and the closure of such an exposition under logical consequence is blurred. There is not much to choose, except for considerations of style, between locutions like "Special Relativity Theory says that ... " and "It follows from Special Relativity Theory that ... ".

[3] There are, however, differences which will be emphasised later. One difference is that our concept of acceptance can be introduced without a detour via the figment of an "ideally rational believer".

Further conditions on theories emerge when setting theories in relation to the world. Thus, theories may reflect the world truly or even comprehensively:

(ver) Veracity if $A \in T$, then A,

(cpr) Comprehensiveness if A, then $A \in T$.

Another external factor that may have an impact on a theory is a canon of inference, a logic. Indeed, as pointed out earlier, theories – in a more colloquial sense of the term – are expected to show some respect to logic. We list three ways in which theories may fulfill such an expectation:

(ccl) Congruential Closure from $A \leftrightarrow B \in \mathbf{L}$ infer $A \in T$ iff $B \in T$,

(mcl) Monotonic Closure from $A \rightarrow B \in \mathbf{L}$ infer if $A \in T$ then $B \in T$,

(reg) Regularity from $A \in \mathbf{L}$ infer $A \in T$.

Finally, consider theories which contain sentences expressing that sentences are members or not members of the theory in question. We extend the language **La** by a unary operator $\Box$ such that if $\Box$ is applied to a sentence A, then the sentence $\Box A$ is meant to express that A is a thesis of the theory under consideration. There are four basic ways in which a theory may systematically contain such reflections on the concept of membership in the theory:

(rve) Reflective Veracity if $\Box A \in T$, then $A \in T$,

(rcp) Reflective Comprehensiveness
if $A \in T$, then $\Box A \in T$,

(rve*) Reflective Veracity* if $\sim\Box A \in T$, then $A \notin T$,

(rcp*) Reflective Comprehensiveness*
if $A \notin T$, then $\sim\Box A \in T$.

These conditions are reflective analogues to (ver) and (cpr) (hence, the mnemonic labels). According to (rve) ((rve*)), the theory T reflects correctly on what sentences are (not) members of T. According to (rcp) ((rcp*)), the theory's reflection of membership (non-membership) in the theory is complete: if A is (is not) a member of T, then T "says so".

13. Modal logics for reasoning about theories

We shall translate the conditions on theories introduced in section 12 into postulates in the modal language **La$^\Box$** (sec. 13.1). Such a translation may be viewed as a purely formal mapping from one set of expressions into another. But it may also be understood as implicitly making a substantial claim about the semantics of certain English expressions. For the developments in sections 14 and 15 it will be important to make the substantial claim. In section 13.2 we turn to the particular modal systems that result by adding **La$^\Box$**-translates of the conditions mentioned in section 12 to one of the base logics between **BM** and **K** of chapter one. A sample of significant differences between choosing **K** rather than a relevant logic as the basis for modal systems will be displayed; these differences will be of consequence for the evaluation of various *interpreted* modal systems considered in subsequent sections.

13.1. An extension of La$^\Box$ by Basic Logical English

In section 12 we have formulated a number of conditions in English with respect to theories formulated in the formal language **La$^\Box$**. We shall now extend **La$^\Box$** by a small fragment of English, namely the expressions *not, and, or, if-then,* and *is in T* ($\in T$). The extended language **ELa$^\Box$** is based on a denumerable set of *atomic sentences* from which new sentences may be formed by applying any of the connectives of **ELa$^\Box$**. The set of connectives of **ELa$^\Box$** consists of the connectives of **La$^\Box$**, $\{\sim, \&, \vee, \rightarrow\}$, together with the unary connectives *not* and $\in T$, and the binary connectives *and, or*, and *if-then.* The unary connective $\notin T$ is defined: $A \notin T$:= *not* $(A \in T)$. Just as we have favoured infix notation over prefix notation in naming well-formed sentences of **La$^\Box$**, so we shall stipulate that we write *A and B, A or B, if A then B* for the results of applying *and*, respectively *or, if-then,* to the sentences A and B.

All of the sentences in section 12 expressing conditions on theories, except (ccl), (mcl), and (reg), are well formed formulae of the language **ELa$^\Box$**. We now collapse back **ELa$^\Box$** into **La$^\Box$** by adding to the definition of **ELa$^\Box$** the following identities:

I. *not* $= \sim$,

II. *and* $= \&$,

III. *or* $= \vee$,

IV. *if-then* $= \rightarrow$,

V. $\in T = \Box$.

Adding I to V to the definition of **ELa**$^{\Box}$ has the effect that applying the operators on the left hand-side to formulae, results in the very same object as when the corresponding operation on the right-hand-side had been applied. Thus e.g., *if* A *then* B and $A \rightarrow B$ are only different names for the same formula. It does not matter then, whether we state the conditions in section 12 in the extended vocabulary of **ELa**$^{\Box}$ or whether we restrict ourselves to the linguistic resources of **La**$^{\Box}$. We may thus reformulate these conditions in **La**$^{\Box}$ as follows.

(cmp)	$\sim\Box A \rightarrow \Box \sim A$
(cns)	$\Box \sim A \rightarrow \sim\Box A$
(adj)	$\Box A \& \Box B \rightarrow \Box (A \& B)$
(dej)	$\Box (A \& B) \rightarrow \Box A \& \Box B$
(add)	$\Box A v \Box B \rightarrow \Box (A v B)$
(prm)	$\Box (A v B) \rightarrow \Box A v \Box B$
(det)	$\Box (A \rightarrow B) \rightarrow . \Box A \rightarrow \Box B$
(ver)	$\Box A \rightarrow A$
(cpr)	$A \rightarrow \Box A$
(rve)	$\Box\Box A \rightarrow \Box A$
(rcp)	$\Box A \rightarrow \Box\Box A$
(rve*)	$\Box \sim\Box A \rightarrow \sim\Box A$
(rcp*)	$\sim\Box A \rightarrow \Box \sim\Box A$.

As one would expect, the closure conditions (ccl) and (mcl), and the regularity condition (reg) are treated as logical rules:

(ccl) $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$$

(mcl) $$\frac{A \rightarrow B}{\Box A \rightarrow \Box B}$$

(reg) $$\frac{A}{\Box A} .$$

Collapsing **ELa**$^{\Box}$ back into **La**$^{\Box}$ by means of the identities I to V, may be viewed as a purely formal reduction of one set of (uninterpreted) objects to another. However, in chapter two we have offered a set of interpretations of **La**$^{\Box}$ using the notion of an R*S-model or – frequently equivalently – that of an R*N-model. And when we listed conditions on theories in section 12, we did not just write down well-formed formulae of **ELa**$^{\Box}$ but we used English sentences to communicate some basic properties of theories to the reader, thus assuming that the reader would share with the author a grasp of how the English expressions occurring in the conditions are to be interpreted . When **La**$^{\Box}$ and the fragment of English used in stating the conditions (Basic Logical English) are thus viewed as *interpreted* languages, then the identities I to V carry a substantive semantical claim: that our favourite class of R*S-models used to interpret the language **La**$^{\Box}$ may be used to model our semantic intuitions concerning **BLE**. Indeed, if formal logic is to lay any claim on reflecting reasoning in the medium of natural language, then such correspondences between certain logical expressions in natural languages and the connectives of formal languages, as expressed in I to V, must be maintained and defended. In the following sections, the substantial interpretation of I to V will be essential for a judicious choice among the many modal systems that have been or may be presented as epistemic or deontic logics.

Excursus on the adequacy of the Routley-Meyer semantics. When favouring a particular logic **L** governing certain English expressions that correspond to connectives of **La**$^{\Box}$, we *ipso facto* favour semantic theories which are demonstrably adequate with respect to **L**. If **L** is determined by a certain class C of R*S-models, then an interpretation of **La**$^{\Box}$ in terms of that class of R*S-models is such an adequate semantic theory. And as far as **L** is an adequate theory of *not* $(=\sim)$, *and* $(=\&)$, *or* $(=\vee)$, *if-then* $(=\rightarrow)$ and $\in T$ $(=\Box)$, an interpretation of **BLE** in terms of the class C of R*S-models must be an adequate semantic theory of the **BLE**-fragment of English.

Although this line of thought can hardly be rejected, Routley-Meyer style semantics for propositional logics have been attacked on the grounds that they fail to be "intuitively adequate", that they are "merely mathematical", "purely formal", "not applied". For a semantics to be

[3] See e.g. Copeland's articles "On when a semantics is not a semantics" (1979), "Pure semantics and applied semantics" (1983), "What is a semantics for classical negation?" (1986).

fully adequate, so it is claimed, we need not only establish its adequacy with respect to an acceptable logic, but also an intuitively satisfactory interpretation of its terms that is relevant to the purpose of assigning truth-conditions to the sentences of the language to be interpreted. In all these objections to the Routley-Meyer semantics – particularly levelled at the ternary relation and the star function – it has never been made sufficiently clear how intuitively satisfactory an interpretation must be in order to vindicate a semantics. Does the sort of picturesque heuristics accompanying possible worlds semantics suffice, or do we need more serious interpretations of model structures such as the temporal interpretations of Kripke-style models for tense logics?

Whatever the intended requirements are, we reject the demand for adequacy in this emphatic sense. When producing theories of *any* kind, including semantic theories, we must decide on an appropriate level abstraction. We frequently find that the explanatory power and conceptual economy of a theory can be enhanced by introducing theoretical terms which do not answer to any pre-theoretic concepts. Some theories, however, are simple and immediately accessible to "the man on the street"; but they may be false. Some theories are both simple and true; but why should all semantic theories be of that kind? To be sure, we have no objection to grading semantic theories according to their "naturalness". But just as the Theory of Relativity should not be discounted as a serious theory of the physical world on the grounds that it locates objects in an unfamiliar type of space, so Kripke's or Routley and Meyer's semantic theories should not be discounted as serious theories about the truth-conditions of certain English sentences just because they introduce relations between "worlds", or "situations", which resist interpretations in terms of everyday notions. The crucial test for a semantic theory is whether it fits the data – in our case: whether the logic determined by a certain class of R*S-models is an acceptable formalisation of reasoning in the medium of **BLE**. It would be "nice" if such a fit could be adjudicated, as it were, *a priori*, by inspecting the semantic theory alone. But there is no reason to suppose that this can always be accomplished; and since 'ought' ought to imply 'can' (though it frequently does not, as we shall see), there is accordingly no reason to *require* of semantic theories that they always allow such *a priori* adjudication. (*End of excursus*)[4]

[4] For further elaborations of this theme cf. van Benthem (1986), ch.9.

13.2. Some benchmark differences between relevant and classical bases for modal logics

Consider the system **KEC!**, i.e. the extension of classical logic **K** by the congruence rule

RE. $$\frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B}$$

and the schema

$\Box$C!. $\Box(A \& B) \leftrightarrow \Box A \& \Box B$.

In **KEC!** we can derive all instances of the schema

*I. $\Box A \& \Box \sim A \rightarrow \Box B$.

Proof. Assume

(1) $\Box A \& \Box \sim A$.

By $\Box$C! (from right to left) and MP from (1):

(2) $\Box(A \& \sim A)$.

In **K**, and so in **KEC!** we have

(3) $A \& \sim A \leftrightarrow (A \& \sim A) \& B$.

Thus, from (3) by RE:

(4) $\Box(A \& \sim A) \leftrightarrow \Box((A \& \sim A) \& B)$

whence from (2) by MP:

(5) $\Box((A \& \sim A) \& B)$.

So by $\Box$C! (from left to right) and MP from (5):

(6) $\Box(A \& \sim A) \& \Box B$,

from which we may infer $\Box B$ by &E and MP. ∎

In monotonic extensions of **KEC!**, the derivation of *I is even simpler. Consider e.g. the proof of *I in the system **KC**, the smallest conjunctively regular modal logic based on **K**:

Assume:

(1) $\Box A \& \Box \sim A$.

From (1) by $\Box$C and MP:

(2) $\Box(A \& \sim A)$.

From the classical tautology

*i. $A \& {\sim}A \rightarrow B$,

we obtain by RM:

(3) $\Box(A \& {\sim}A) \rightarrow \Box B$

whence $\Box B$ from (2) and (3) by MP. ■

Consider any monotonic modal logic based on **K**. All instances of the following schemas are theorems of **K**:

*ii.	${\sim}A \rightarrow .A \rightarrow B$,
*iii.	$A \rightarrow .B \rightarrow A$,
*iv.	$A \& B \rightarrow .A \rightarrow B$.

Hence, the following schemas are derivable in one step, using the rule RM:

*II.	$\Box{\sim}A \rightarrow .\Box(A \rightarrow B)$,
*III.	$\Box A \rightarrow \Box(B \rightarrow A)$,
*IV.	$\Box(A \& B) \rightarrow \Box(A \rightarrow B)$.

For modal systems based on the relevant or even semi-relevant systems charted out in the diagram of section 4, all these derivations are blocked where appeal is made to the classical theses *i - *iv. This is easily verified by considering the characteristic matrix for the strongest of our relevant and semi-relevant systems **RM3** which we have displayed in section 4. For *i and *ii let $A=1$ and $B=2$. Then $(1 \& {\sim}1 \rightarrow 2)=2$ and $({\sim}1 \rightarrow .1 \rightarrow 2)=2$, thus refuting *i and *ii respectively. For *iii let $A=1$ and $B=2$. Then $(1 \rightarrow .0 \rightarrow 1)=2$, thus refuting *iii. For *iv let $A=0$ and $B=1$. Then $(0 \& 1 \rightarrow .0 \rightarrow 1)=2$, thus refuting *iv.

14. Interpreted modal logics I. Epistemic modalities

14.1. Acceptance sets. LR as a basic logic of acceptance.

As mentioned earlier, by the *acceptance set* of an agent X at a given time t, I mean the set of all sentences X is committed to believe at t. By 'commitment' I mean commitment in virtue of logical consequence: if X accepts that A and B is a logical consequence of A, then X accepts that

B.

It is this notion of acceptance (commitment to belief) that allows to criticise agents – for example in debates – when they have explicitly confessed to certain beliefs, say $p_1, \ldots, p_n$, from which we can derive an implausible belief q. Such a situation may be described as follows. X at t believes each of $p_1, \ldots, p_n$ and is aware of these beliefs. In virtue of his beliefs in each of $p_1, \ldots, p_n$ and the fact that q is a logical consequence of these beliefs, X is committed at t to believe that q. But (a) X may not be be aware of his commitment at t to q, and, moreover, (b) he may even explicitly believe that *not*$\neg q$ in which case his acceptance set is inconsistent (though, so we may assume, he is not aware of it). At some later time t', X's commitment to q is pointed out to him. Then X has two options: either add q to his set of explicit beliefs or reject q and change his beliefs accordingly so as to avoid further commitment to q. If the first option is taken and (b) is the case, then X holds explicitly at t' an inconsistent pair of beliefs, which, in the dominant intellectual climate of the Western World, needs very good arguments to defend. If X submits to the criticism, and thus takes the second option, then it will not suffice to simply declare disbelief in q. X will have to change his acceptance set such that q is no longer in it. Such changes of belief are no trivial manoeuvres, as we shall explain in chapter four. Assuming that X favours the first option, it is false at both t and t' that X *believes* (in any straightforward sense) that q, although it is true at t that X *accepts* that q. Moreover, X may continue at t', and for some time thereafter, to accept that q despite explicitly declaring disbelief in q. This may easily happen when X's efforts to change his acceptance set so as to avoid commitment to q are unsuccessful; perhaps, because he has overlooked certain alternative derivations of q (that is, derivations not from $p_1, \ldots, p_n$) from what he accepts.

Another important area of application for the notion of acceptance is reasoning about databases. A database is a set of data structures (the datafile) representing states of affairs, together with an inference engine (an implementation of a logic). A datafile is like the axiomatic base for a theory. The set of all data structures retrievable at a time t from the database by means of the inference engine represents what, according to the database, is accepted as true at t. Thus, the set of retrievable data is an acceptance set in our sense. It is sometimes just as important to know what can be retrieved as what can not be retrieved from a database and to draw inferences from such knowledge. What we need then is a logic of acceptance (according to database D) or retrievability (from D).

The importance of such a logic may be illustrated by considering certain forms of non-monotonic reasoning. Generally speaking, a non-monotonic inference to q from a set of sentences Γ,

(1) $\Gamma \vdash q$,

is such that adding a premise p to Γ may block the inference from $\Gamma \cup \{p\}$ to q. The truth of (1) does not guarantee the truth of

(2) $\Gamma' \vdash q$

for any superset Γ' of Γ. To take an example, the following is – unfortunately – a highly feasible rule of inference:

(3) *Jones is a tenured academic* $\vdash$ *Jones is male and white.*

But (3) is sensitive to the addition of premisses. If we add to the premiss that *Jones is a tenured academic* the further premiss that *Jones is an aboriginal activist employed in the Women Studies Department*, then we ought to withdraw the inference to *Jones is male and white* from the premiss that *Jones is a tenured academic* – that inference has been *defeated.*

The example shows that the acceptability of a non-monotonic inference from a given set of data D does not only depend on what is in D but also on what is not in D. This suggests that non-monotonic reasoning requires the combination of two levels of monotonic reasoning: reasoning by means of the inference engine coupled with the database in question (the object-level) and reasoning *about* the set of data retrievable (and not retrievable) from the database by means of the inference engine (the meta-level).[5]

Having made clear the notion of acceptance that will be in focus here, we may continue by asking what properties acceptance sets have. This is a question different from the one as to what properties it would be *desirable* for acceptance sets to have. Quite plainly, it would be desirable if the set of sentences an agent accepts as true would coincide with the set of true sentences, that is, if it were both comprehensive and veridical (and, perhaps, also reflectively so). Such acceptance sets, however, would be of little interest: they are not of the kind we can expect to encounter in reality. Neither do we expect acceptance sets always to be consistent, although we usually prefer consistent acceptance sets to inconsistent ones. This marks an important difference between an acceptance set in our

[5] This combination of levels of reasoning underlies Doyle and McDermott's approach to non-monotonic reasoning; see McDermott and Doyle (1980) and McDermott (1982). An integration of both levels of reasoning has been attempted in Moore's (1983) Autoepistemic Logic.

sense and a *rational belief set*, as the term is used in the literature on doxastic logics and related topics: rational belief sets are, by definition, consistent.[6] The subject of an acceptance set is an ordinary agent. By contrast, the subject of a rational belief set is an ideal agent, an ideally rational agent. For, whatever an (ordinary) agent accepts explicitly, determines an acceptance set. But the set of sentences an agent accepts explicitly may be inconsistent. Hence, the acceptance set of such an agent does not qualify as a rational belief set. A necessary condition for an agent to have a rational set of beliefs is that he completely shuns inconsistent beliefs – at least in this sense, the agent must be ideal.

We shall use the notation $a_{X,t}A$ to express that the agent X accepts at t that A is true. In the sequel the subscripts to the operator a will be omitted on the understanding that the references to an agent and a point of time are contextually fixed. We state postulates for a logic of acceptance in the language $\mathbf{La}^a$ which results from $\mathbf{La}^a$ by replacing the unary operator $\Box$ by the unary (acceptance-) operator a.

By definition, acceptance sets are closed under logical consequence. Hence, the envisaged logic of acceptance will be a monotonic modal system:

RM. $$\frac{A \rightarrow B}{aA \rightarrow aB}$$

Like RM, further postulates result from the commitment ingredient in the notion of acceptance. Thus we require that acceptance sets be adjunctive and detached, hence:

aC. $aA \& aB \rightarrow a(A \& B)$,

aI. $a(A \rightarrow B) \rightarrow .aA \rightarrow aB$.

As a basic logic of acceptance we therefore propose a logic out of the family of smallest regular modal logics, **L.R**.

6 See e.g. Ellis (1979) and Lenzen (1980). A precursor to the notion of a rational belief set is the notion of a defensible set (of sentences), or model set, as it occurs in Hintikka's classic (1962). I am using the term '*rational* belief set' opaquely, as it were. In so using the term I do not endorse the view that the properties conventionally associated with it, are essential features of rationality.

14.2. Considerations about the choice of a base logic

The next question to ask is which of the logics **L** considered in section 4 should we extend by RM, aC, and aI in order to obtain an intuitively adequate basic logic of acceptance? In view of the wide range of choice it would be preposterous to even attempt an answer that singles out one particular system. A unique choice would have to depend on a variety of factors, many of which cannot be assessed without reference to the particular context in which the logic is to be applied.[7] (Thus, for example, choice of a logic for AI applications depends significantly on how its implementations fare – in particular languages and on particular machines – under a set of efficiency parameters.) But we can at least draw a demarcation line between certain classes of logics.

Consider for example the classical modal system **K.R** under the acceptance interpretation. **K.R** contains each of the following schemas (see the proofs in sec. 13):

*I.	$aA \& a \sim A \to aB$,
*II.	$a \sim A \to a(A \to B)$,
*III.	$aA \to a(B \to A)$,
*IV.	$a(A \& B) \to a(A \to B)$.

*I and *II (in the presence of aI) rule out the possibility of inconsistent but non-trivial acceptance sets. This restricts the range of applicability of **K.R** as a logic of acceptance to a (very) small subset of real life acceptance sets. In particular for reasoning about databases, **K.R** is virtually ruled out as a candidate. For, sufficiently large databases are, as a rule of thumb, locally inconsistent; it would be disastrous to render them therefore trivial.

*III and *IV introduce commitments into acceptance sets which should flatly be refused (so do, of course, *I and *II). For a counterexample to *III, consider the following. I know, and, hence, believe that **RM** is a proper supersystem of **R**. Thus, I do not believe that

(1) if A is provable in **RM**, then A is provable in **R**.

In particular I do not believe the following instance of (1):

---

[7] Cf. Routley (1979).

(2) if $p\&q \rightarrow p$ is provable in **RM**, then $p\&q \rightarrow p$ is provable in **R**.[8]

I also believe that all instances of $A\&B \rightarrow A$ are theorems of **R**. So I accept commitment to the belief that

(3) $p\&q \rightarrow p$ is provable in **R**.

Yet I feel perfectly entitled to refuse commitment to (2) as a consequence of belief in (3). These sorts of counterexamples are familiar from the literature on the paradoxes of material implication. What is perhaps less familiar is the fact that principles like *III (and *IV) do not only offend deeply entrenched intuitions about valid inference but also lead to patently absurd consequences when diachronic aspects of acceptance sets are taken into consideration. Suppose that I have in fact got the inclusion relation between **RM** and **R** wrong: I believe that **RM**$\subseteq$**R**. Thus I accept (2) as true. Of course, inspecting the axioms of **R**, I also have excellent reasons to believe in (3). Now suppose that you have convinced me that I should no longer accept the false belief (2). As pointed out at the beginning of this section, it does not suffice to simply declare that I now reject (2); I have to adjust my acceptance set so as to effectively remove commitment to (2). If acceptance is closed under implication provable in **K**, I shall therefore be required to give up (3) also. But this is absurd. For, in contrast to (2), (3) is as well-grounded a belief as beliefs can be. Thus, if I want to stick to (3), I am bound to embrace the false (2); and if I want to reject what I believe to be false, viz. (2), then I must also reject what I believe to be true, viz. (3).

The example just given tells similarly against *IV. Suppose that I do believe that $p\&q \rightarrow p$ happens to be a theorem of both **R** and **RM**. That should not commit me to the false sentence (2). And if I want to give up (2), I should not be compelled to choose between giving up the belief that $p\&q \rightarrow p$ is provable in **R** and the belief that $p\&q \rightarrow p$ is provable in **RM**. *IV is just as implausible as *III is.

To some up these considerations: classical logic **K** is unsuitable as the basis of a logic of acceptance.[9] It completely fails with respect to inconsistent acceptance sets and it provides a wrong guide to the commitments of agents even when these commitments are consistent.

[8] I am not saying that disbelief in (2) follows from disbelief in (1) – that would be a blatant *non sequitur*. I claim that disbelief in (2) is justified for similar reasons as disbelief in (1): just as there is no general argument from theoremhood in **RM** to theoremhood in **R**, so there is in particular no argument from the fact that $p\&q \rightarrow p$ is provable in **RM** to the fact that the same formula is provable in **R**.

[9] So is classical logic without contraction, **RWK**, since all instances of *I to *IV are theorems of **RWK.R**.

We are thus left with logics weaker than **K** (and **RWK**) as possible bases for a logic of acceptance. I shall argue now that we we are best served with a fully relevant logic as such a base.

Consider the semi-relevant logics **RM** and **RM3**. Both logics contain the schemas

*vi. $A \rightarrow A \rightarrow A$, and

*vii. $A \& \sim A \rightarrow B \vee \sim B$.

Thus, by RM, any basic logic of acceptance based on either **RM** or **RM3** contains the schemas

*VI. $aA \rightarrow a(A \rightarrow A)$, and

*VII. $a(A \& \sim A) \rightarrow a(B \vee \sim B)$

But *VI and *VII are just false principles about the deductive structure of acceptance sets. There is no way to *infer* from my commitment to *It is raining now* a commitment to *If (it is raining now), then (it is raining now)*. Similarly, although I may in fact believe that

(1) the equation $x^n + y^n = z^n$ either has or has not a solution for $n > 2$ in the positive integers x, y, z s.t. $xyz \neq 0$,

a commitment to (1) cannot be *inferred* from a commitment to

(2) Epimenides lies and Epimenides speaks the truth.

This kind of objection to *VI and *VII is not particularly tied to the acceptance interpretation of the operator a. The objection is primarily directed against the claim that the schemas *vi and *vii should be part of a characterisation of the notion of implication. From a relevant point of view, *vi and *vii are simply invalid principles of inference and thus give rise to an inadequate notion of commitment, if the latter notion is to be explicated in terms of provable implication. A base logic for a logic of acceptance is thus better be sought among fully relevant systems.

14.3. Does fragmentation save classical acceptance logic?

Considerations like the ones advanced in the last paragraph are unlikely to move a determined classicist towards rejecting *VI and *VII. However, some classical logicians do feel some discomfort at the thought of what material implication "does to" inconsistent acceptance sets. For, the reality of such sets and the need to treat them in a non-trivial way have become inescapable. One strategem to save classical logic in the face of inconsistent but non-trivial acceptance sets is *fragmentation*. The idea to treat inconsistent sets as composites of consistent fragments goes back to Jaškowski's work on Discussive Logics. More recently the idea has been revived by Schotch and Jennings in the context of deontic logics. Rescher and Brandom have proposed fragmentation as a universal tool to defuse the (classically) explosive potential of inconsistent theories. The fragmentation strategy, in particular with respect to acceptance sets, has also been recommended by Lewis and Stalnaker.[10] I shall now discuss this proposal in the context of a logic of acceptance.

The strategy is designed to block the derivation of *I by rejecting the adjunctivity principle

aC. $aA \& aB \rightarrow a(A \& B)$.

This appears to be a desperate move to make. If an agent has committed himself to accept A and to accept B, then, surely, he has taken upon himself a commitment to accept *A and B*? That commitments be adjunctive seems to be the very least we should expect of any notion of commitment. We are in danger of loosing any grip on the concept of commitment, if we cannot demand a defence of *A and B* from an agent who has committed himself to A and to B. Thus, very good reasons indeed need to be advanced for dispensing with the requirement that acceptance sets be closed under adjunction.

Fragmentation theories attempt to make the rejection of adjunctivity for acceptance sets more palatable by offering a redescription of what it is for an agent to apparently be committed to a belief of the form $A\& \sim A$.[11] According to Lewis and Stalnaker, the total belief system of such an agent is split into, possibly overlapping, fragments, one giving rise to a commitment to A, another committing the agent to $\sim A$. But, so Lewis and Stalnaker contend, there is no fragment that gives rise to a commitment to $A\& \sim A$. How plausible is such a view?

---

[10] See Jaškowski (1948), Schotch and Jennings (1981), Rescher and Brandom (1980), Lewis (1982), and Stalnaker (1984).

[11] Schotch and Jennings reasons for rejecting aC – which are very similar to Lewis' and Stalnaker's – will be discussed in the section on deontic interpretations of the a-operator.

There is a core to the fragmentation theory of belief systems which does seem very plausible indeed. Beliefs are conveniently bundled so as to support routines which are frequently mutually exclusive. In order to make a cup of tea I routinely invoke a bundle of beliefs as to where I can fill the jug with water, how the jug is to be switched on, where I keep the tea, how long the tea leaves have to be immersed in water, etc. Similarly, I invoke another bundle of beliefs when riding a bicycle: in which direction to push the pedals, how to shift the gears, when it is prudent to switch on the lights, etc. For all I know, it is unlikely that there will ever be a single occasion on which I shall have to invoke both bundles of beliefs; making a cup of tea while riding my bicycle is an unlikely action to perform. Thus, while there have been many occasions on which it was correct to attribute to me the belief that *the tea is in the cupboard* and also many occasions on which it was correct to attribute to me the belief that *one shifts into a lower gear by pushing the gear shift downwards*, there has never been a single occasion on which it would have been correct to attribute to me the belief that *the tea is in the cupboard and cycling in a lower gear comes about by pushing the gear shift downwards*. To quote Lewis: "Different fragments came into action in different situations, and the whole system of beliefs never manifested itself all at once." ((1982), p.436).

It seems hard to deny the observation that lies at the core of the fragmentation theory. But this core needs substantial supplementation if a case is to be made against the adjunctivity of acceptance sets. So far we have only been presented with an observation which makes it implausible that beliefs *simpliciter* are generally closed under adjunction. But that view has been implausible all along. (The core of the fragmentation theory provides an interesting explanation of why that view is so implausible.) Belief *simpliciter* is not the requisite notion of belief for which the adjunction schema is in serious dispute. What we need is an argument that commitment-to-believe (acceptance) fails to be adjunctive. Thus, the core of the fragmentation theory is extended to the view that the set of sentences an agent is committed to believe is not, as we have proposed, the closure of his explicitly held beliefs under adjunction and provable implication, but the union of all fragments thus closed.[12] But this extension of the core theory *is* implausible. If an agent is committed to A and to B, then – whether or not the agent in question has ever cared to consider the beliefs in conjunction – he is committed to A *and* B. It is

[12] See Lewis (1982) and Stalnaker (1984).

of no avail to plead schizophrenia. Just as an agent who believes A, is committed to B, if B follows logically from A – regardless of whether his mental life is such that he has never formed the belief that B in his head – so an agent is committed to the conjunction $A\&B$, if he has committed himself to both conjuncts. The arguably attractive core of the fragmentation theory lends no support at all to rejecting adjunction for acceptance sets.

But even if we accept the extension of the core theory, thus having an independent ground for rejecting the schema aC, classical logic is not yet in the clear. According to the extended theory, each fragment of a belief system is closed under adjunction and provable implication. If 'provable implication' means implication provable in **K**, then inconsistent fragments are trivial fragments and, hence, the set of sentences accepted according to a fragmented belief system (being the union of all fragments) is also trivial. In other words, the schema

*I $\qquad aA \& a \sim A \rightarrow aB$

would still be valid in a logic of acceptance-according-to-a-single-fragment. This calls for a further extension of the core theory: all fragments must be consistent. (So the antecedent of *I never obtains.) According to Stalnaker, the consistency condition on fragments

> "must hold. This is clear since the only set of propositions conforming to the first two conditions [closure under provable material implication and closure under adjunction] but violating the third [consistency] is the set of all propositions, and no belief state in which all propositions were believed could distinguish any actions as appropriate or inappropriate."[13]

For obvious reasons, there is no need in the present context to comment on this "argument".

Again, the observation on which the fragmentation theory is based, lends no support to the thesis that fragments are always consistent. Although the core theory goes some way towards explaining why inconsistencies in the belief states of agents remain frequently undetected, namely because inconsistent pairs of beliefs may be isolated from each other, there is no reason to suppose that inconsistent pairs of beliefs are always quarantined in separate fragments of one's total set of beliefs.

Indeed, the latter view clashes just as much with the data as the now discredited thesis that it is impossible for an agent to believe that A and

[13] (1984), p.83.

to believe that $\sim A$, for some sentence A. Consider Jones who has published a paper, say, about the concept of knowledge, in which he has advanced the theses $A_1, \ldots, A_n$. Smith writes a reply to Jones' paper, arguing that if one accepts $A_1, \ldots, A_n$, then one is committed to $\sim A_i$ (for some $i \in \{1, \ldots, n\}$). That is, Smith refutes Jones' theory by a *reductio ad absurdum* argument, pointing out that Jones' theory is inconsistent. Now, it may be true that Jones' belief system is fragmented; that his beliefs about how to boil a cup of tea are tucked away in one fragment, that his beliefs about how to ride a bicycle are part of another fragment, etc. And although Jones has presumably failed to realise that the theses $A_1, \ldots, A_n$ may not be consistently combined, he certainly has combined $A_1, \ldots, A_n$ to a single theory, viz. Jones' theory of knowledge. Thus, we may assume that just as Jones' beliefs about riding bicycles are bundled in one fragment, there is a fragment of Jones' total belief system which contains his beliefs about what constitutes knowledge. But that fragment is inconsistent, as Smith has shown. If, however, Stalnaker's thesis about the consistency of fragments were true, then beliefs against which a *reductio* argument can be advanced are never combined so as to be part of a single fragment; for, such a fragment would be inconsistent. That consequence of Stalnaker's thesis is hardly credible.

In Rescher and Brandom's fragmentation theory of inconsistent sets, 'fragment' is a term of art. There is no serious connection to the core theory underlying both Lewis' and Stalnaker's proposals, although the core theory is sometimes invoked for heuristic purposes. For Brandom and Rescher a fragment of an inconsistent set S of sentences is defined as a maximal consistent subset of S. Apart from the occasional masquerade of heuristics as motivation, Rescher and Brandom's central claim is that there is no need to abandon classical logic in order to make non-trivial inferences from inconsistent sets of sentences: we may replace an inconsistent theory by the family of all its maximal consistent subtheories. Where S is an inconsistent set of sentences and $\{S_1, \ldots, S_n\}$ the set of all fragments of S, a sentence A is a non-trivial consequence of S just in case A follows classically either from S_1 or from ... or from S_n.

The proposal suffers from a number of problems. First, it must appear as entirely *ad hoc*: no other motivation is seriously advanced except to save classical logic in the face of inconsistent theories.

Secondly, when "replacing" an inconsistent theory by the set of its fragments, information is bound to be lost. Thus, the adequacy of such replacements is in serious doubt. To take the simplest example: the inconsistent theory S contains some sentence of the form $p\& \sim p$. Yet that sentence is – for obvious reasons – not a member of the replacement theory. To take yet another, more concrete, example, consider naive set theory with the naive abstraction schema expressing that to every condition $F(x)$ there corresponds a set y whose elements are exactly those objects that satisfy $F(x)$:

(1) $(\exists y)(\forall x)(x \in y \leftrightarrow F(x))$.

For $F(x)$ we may fill in the condition for membership in the Russell set, i.e.

(2) $(\exists y)(\forall x)(x \in y \leftrightarrow \sim(x \in x))$.

For $x=y$ we thus obtain:

(3) $y \in y \leftrightarrow \sim(y \in y)$

which is classically equivalent to

(4) $y \in y \& \sim(y \in y)$.

According to Rescher and Brandom, naive set theory consists of essentially two fragments, one containing one half of the naive abstraction schema (1), the other containing the other half. In other words, no instance of the naive abstraction schema is a theorem of the proposed classical reconstruction of naive set theory!

Thirdly, the proposal of Rescher and Brandom would have something to recommend it, if it could be argued that it is much simpler to first split inconsistent theories in maximal consistent subtheories and then close these fragments under classical logic. But such an argument has never been provided and the prospects for that it can be provided are not good. It seems much simpler, especially from a computational point of view, to take inconsistent sets of sentences for what they are and draw inferences from them by means of a paraconsistent logic.

We have offered a number of arguments to the effect that the fragmentation theory provides no grounds for rejecting aC as a valid principle of a logic of acceptance. Lewis' and Stalnaker's core theory is based on comparatively uncontroversial observations about the occurrent beliefs of agents. These observations, however, are irrelevant to the normative concept of acceptance. The extended core theory, on the other hand, which would have consequences for a logic of acceptance, runs counter to basic intuitions concerning the concept of commitment and contains highly implausible claims about how the belief systems of agents are subdivided into parts. In contrast to Lewis and Stalnaker, Brandom and Rescher's aim is much more modest: they merely try to show that there is a sense in which classical logic *can* be used to draw inferences from inconsistent sets of sentences. Their theory falls far short from providing the very good reasons required for rejecting aC under the acceptance interpretations. Moreover, in the course of processing inconsistent theories in the way proposed by Brandom and Rescher, these theories get mutilated in rather drastic ways. Their programme is thus better described as providing a systematic method for cutting down inconsistent theories to consistent ones – a "logic of inconsistency" (the title of their book) is not really on offer. In conclusion: no version of the fragmentation theory so far proposed undermines the validity of the schema aC under the acceptance interpretation.

14.4. Extensions of the basic acceptance logic L.R

According to **L.R**, acceptance sets are both additive,

$$aA \vee aB \rightarrow a(A \vee B),$$

and dejunctive,

$$a(A \& B) \rightarrow aA \& aB.$$

(These schemas follow by RM from basic properties of the base logic **L**, like v-composition, &-composition, ADJ, and MP.) We shall now consider a number of perhaps more contentious principles that may be added to the basic logic **L.R**. Naturally, these are only a few of the many principles that *can* be considered. However, the selection is not entirely arbitrary, since the principles we shall now consider, give expression to some of the key properties – mentioned in section 13 – that may be ascribed to acceptance sets.

Should we require acceptance sets to be regular with respect to the logic that determines the commitment relation under which an acceptance set is to be closed? If so, then we should want to add to **L.R** the rule

RN. $$\frac{A}{aA}.$$

That a logic of acceptance should be closed under the rule RN is indeed so plausible, that perhaps not L.R but L.K (=L.R+RN) should be viewed as a basic system. If a logic is accepted as the proper canon for determining what one is committed to believe given one's explicit beliefs, then, surely, one has committed oneself to the logic in question – that is to say, one should accept the theses of that logic. I don't see how a convincing case for RM but against RN could be made. So I recommend

RN as an intuitively valid rule of inference for a logic of acceptance.

More controversial are the following postulates:

aD.	$a \sim A \rightarrow \sim aA$,
aW.	$aaB \rightarrow aB$,
a5′.	$a \sim aB \rightarrow \sim aB$,
a4.	$aB \rightarrow aaB$,
a5.	$\sim aB \rightarrow a \sim aB$.

The first three schemas should be rejected. Counterexamples to aD are provided by any inconsistent acceptance set. aW and a5′ require that acceptance sets be veridical with respect to what is and what is not accepted. But there is nothing in the notion of commitment-to-believe (acceptance) that could prevent an agent from forming erroneous beliefs about what he believes. I may believe that I am committed to believe that A (thus accepting that I accept that A) when in fact I am not committed to A. Certainly, the mere *belief* that I am committed to A can not commit me to A. So aW should be rejected. Similarly, I may believe that I am not committed to A (thus accepting that I do not accept A). But that belief may be just false: I may in fact be committed to A. What I am committed to is not merely a matter of what I believe to be committed to. So a5′ should be rejected too.

By contrast, both a4 and a5 are quite plausible principles for a logic of acceptance. If I am committed to A, then I cannot reject my commitment to A. So if I am committed to A, then I am commmitted to acknowledge my commitment to A, as required by a4. Similarly, if I am not committed to A, then I am committed to believe it: I cannot coherently reject commitment to the fact that I am not committed to A. Thus, a K45-modal system based on a relevant logic emerges as a good candidate for a logic of acceptance.

There are a number of modal operators, corresponding to natural epistemic attitudes, that may be defined by means of the acceptance operator a and negation. a and $\sim$ combine to four basic modes of acceptance:

$$a \;,\; \sim a \;,\; a\sim \;,\; \sim a \sim .$$

Given our interpretation of aA as 'A is accepted', the other three basic $\{\sim, a\}$-modalities may be read as follows:

$\sim aA$	–	A is rejected,
$a \sim A$	–	A is denied,
$\sim a \sim A$	–	A is consistent (with what is accepted).

It is important to distinguish rejection from denial. In traditional terminology, acceptance and rejection are contradictories whereas acceptance and denial are contraries. Denial, rejection and consistency are important enough notions to merit their representation in an epistemic logic by means of defined operators:

Dr. $rA := \sim aA$

Dn. $nA := a \sim A$

Dq. $qA := \sim a \sim A$.

Combining these defined notions, we obtain three more operators which represent natural and important epistemic attitudes.

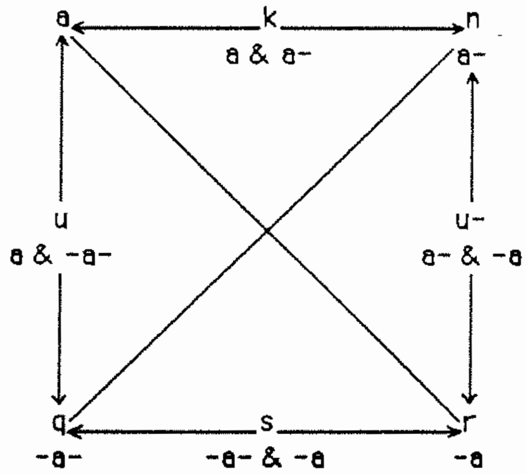
Dk. $kA := aA \& a \sim A$

Ds. $sA := \sim a \sim A \& \sim aA$

Du. $uA := aA \& \sim a \sim A$.

A sentence of the form kA represents the fact that the acceptance set under consideration is *inconsistent* ("kontradiktory") with respect to the sentence A. sA expresses that judgement as to whether A is the case or not is *suspended*. Judgements may be suspended either because the sentence in question has not yet been considered or because the information available does not yet justify either acceptance or denial. uA is a formal representation of the fact that A is *consistently accepted* ("univocally" accepted).[14] The logical relations among these acceptance modalities may be summarised in an epistemic square of oppositions:

[14] It would perhaps be preferable to name this modality 'c' instead of 'u'. However, I want to reserve 'c' as a name for a deontic modality in section 15. So I resort to the less mnemonic 'u' here.



The following theorem shows that already in the basic system $\mathbf{L.R}^a$ we can derive some interesting theorems about the logical relations between the epistemic notions just defined.

THEOREM 14.1.

Let $\mathbf{L}$ be a logic at least as strong as $\mathbf{DW}$ and let the operators r, n, q, k, s, u be as just defined. Then all instances of the following schemas are derivable in $\mathbf{L.R}^a$:

(1)	$kA \to aA$	(2)	$kA \to nA$
(3)	$uA \to aA$	(4)	$uA \to pA$
(5)	$u{\sim}A \to nA$	(6)	$u{\sim}A \to {\sim}aA$
(7)	$sA \to qA$	(8)	$sA \to {\sim}aA$
(9)	$qA \leftrightarrow {\sim}nA$	(10)	$aA \leftrightarrow n{\sim}A$
(11)	$sA \leftrightarrow {\sim}(nA \vee n{\sim}A)$	(12)	$sA \leftrightarrow qA \& q{\sim}A$
(13)	$sA \leftrightarrow s{\sim}A$	(14)	$s{\sim}A \to {\sim}aA$
(15)	$kA \leftrightarrow nA \& n{\sim}A$	(16)	$aA \to {\sim}(qA \& q{\sim}A)$
(17)	$uA \leftrightarrow aA \& qA$	(18)	${\sim}u{\sim}A \leftrightarrow {\sim}(a{\sim}A \& q{\sim}A)$
(19)	$uA \to {\sim}u{\sim}A$	(20)	$uA \& uB \to u(A \& B)$
(21)	$u(A \to B) \to .uA \to uB$	(22)	$\dfrac{A \to B}{uA \to uB}$
(23)	$\dfrac{A \to B}{qA \to qB}$	(24)	$\dfrac{A \to B}{nB \to nA}$

Proof. The schemas follow essentially from the definitions of the non-primitive operators. For most derivations, $\mathbf{L}$ may be $\mathbf{B}$. Only for the proof of (21) do we need systemic contraposition, Cp. In order to derive the rules (22), (23) and (24), use RM. We provide the proof of (21):

By Id, Cp and Replacement, we have

(1) $a(A \to B) \to a(\sim B \to \sim A)$.

By aI: (2) $a(\sim B \to \sim A) \to .a \sim B \to a \sim A$.

And by Cp and Dq:

(3) $a \sim B \to a \sim A \to .qA \to qB$.

So from (1), (2) and (3) by TRANS:

(4) $a(A \to B) \to .qA \to qB$.

From (4) and aI we obtain by ADJ, &-Composition and MP:

(5) $a(A \to B) \to (aA \to aB) \& (qA \to qB)$.

The following formula is a theorem of our weakest logic **BM** (see T3.1):

(6) $(aA \to aB) \& (qA \to qB) \to .aA \& qA \to aB \& qB$

whence from (2) and (3):

(7) $a(A \to B) \to .aA \& qA \to aB \& qB$.

But

(8) $a(A \to B) \& q(A \to B) \to a(A \to B)$.

So from (5) and (4) by transitivity,

(9) $a(A \to B) \& q(A \to B) \to .aA \& qA \to aB \& qB$

which, by Dq and Du, may be abbreviated to the required schema

(10) $u(A \to B) \to .uA \to uB$. ■

15. Interpreted modal logics II. Deontic modalities

15.1. The descriptive interpretation of deontic logic

> "Deontic logic can be defined as the study of those sentences in which only logical words and normative expressions occur *essentially*. Normative expressions include the words 'obligation', 'duty', 'permission', 'right', and related expressions. These expressions may be termed *deontic words*, and sentences involving them *deontic sentences*." (Foellesdal and Hilpinen (1971), p.1)

This quote from Foellesdal and Hilpinen's introductory essay to the collection Hilpinen (1971) delineates in broadest outline the scope of deontic logics. Within this scope there is room for a variety of systems of deontic logic, each of which may lay claim to formalising certain aspects of discourse involving normative expressions. One lesson has generally been learned from the abundance of apparent paradoxes and open problems deontic logics have been faced with since their inception: when proposing a deontic logic, it should better be made clear in advance what aspects of deontic discourse are meant to be formalised by that logic. Accordingly, we shall now explain in some detail what we shall mean by expressions like oA occurring in the deontic logic to be proposed later.

Our approach here will be based on a *descriptive* conception of deontic operators like *it ought to be that* ... or *it is permitted that* .... For the sake of simplicity, we shall consider for now a formal language, $\mathbf{La}^{o}$, which is just like $\mathbf{La}^{\Box}$ except that the unary operator $\Box$ is replaced by the operator o with the same syntactic properties as $\Box$: o maps sentences into sentences. By the descriptive conception of the obligation operator o we mean the following. At any time t, an agent X has certain views as to what ought to be the case. These views may derive from different sources: legal codes, social norms or expectations, the agent's conscience, etc. They may also pertain to agents other than the one who holds the views in question. For example, Joan may hold the view that it ought not to be that John puts his feet on the desk. Let $O_{X,t}$ be the set of all those sentences A such that

(O) A is in $O_{X,t}$ just in case X accepts at t that A ought to be true.

Such a set will be referred to as a *norm set* (for X at t). 'Acceptance', as it occurs in the right-hand-side of (O), is a convenient term in this context. For one, it suggest – rightly, as we shall see – a close relation to the notion of acceptance treated in the last section. On the other hand, it allows to remain neutral on the issue as to whether norms can be proper bearers of truth-values. All we mean here by 'X accepts at t that it ought to be that p', is that X acknowledges it as appropriate to require the truth of p at t. (If one wishes so, one may add that X at t also believes it to be true that p ought to be the case.)

We translate into $\mathbf{La}^{o}$ the fact that for some sentence A, $A \in O_{X,t}$ as oA. Strictly speaking, we should write something like $o_{X,t}A$. But as in the case of the acceptance operator a we shall omit indices for agents and times on the understanding that these points of reference are held constant in any single context. A sentence of the form oA is thus *descriptive*: it says that according to X at t, A ought to be the case; equivalently: oA is a sentence describing a property of the set $O_{X,t}$, viz. that A is a member of that set. Given this descriptive conception of the obligation operator o, the task for the deontic logician is now to formulate invariant properties of norm sets in the language $\mathbf{La}^o$, just as we have attempted in the previous section to formulate invariant properties of acceptance sets in $\mathbf{La}^a$.

Alchourron (1972) has contrasted the descriptive interpretation of the obligation operator with a *prescriptive* one. To simplify matters, he introduces an absolute monarch, Rex, who is unanimously recognised by his subjects as the only source of obligations. Alchourron then goes on to draw the contrast between the descriptive and the prescriptive interpretation as follows.

> "We saw in the descriptive interpretation that a state of affairs p is obligatory and forbidden when Rex has commanded that p and that not p. I pointed out then that even if this is not an impossible situation, it is a regrettable one. But why is it regrettable? Not because it is unfair, unjust or bad from some axiological point of view but because the authority has betrayed his prescriptive intention in prescribing too much (incompatible results). In this sense I believe that [the axiom $oA \rightarrow {\sim}o{\sim}A$] represents a conceptual criterion for deontic consistency in the field of prescriptive (normative) discourse, so I understand that it must be accepted in the prescriptive normative interpretation." (p.454f.)

It appears that, according to Alchourron, on the prescriptive interpretation of the operator o, the inference from oA to ${\sim}o{\sim}A$ is justified because it is assumed that "prescriptive intentions are not betrayed". That is, it is assumed that what ought to be the case according to the norm set under consideration, is consistent. The prescriptive interpretation thus turns out to be a special case of the descriptive interpretation: on the prescriptive interpretation sentences of the form oA are descriptive of *ideal*, and that means, according to Alchourron, at least consistent norm sets. This interpretation of the quote from Alchourron (1972) fits well with his assertion (in (1969) and (1972)) that for consistent sets of norms, the logic of obligation under the descriptive interpretation ("normative logic") coincides with the logic of obligation under the prescriptive interpretation ("logic of norms").

Most systems of deontic logic in the literature do contain the schema

oD. $\quad o{\sim}A \rightarrow {\sim}oA.$

Given the descriptive interpretation of the obligation operator o, this principle is clearly invalid. Conflicts of duty are a regrettable but common phenomenon. Norm sets not only may, but frequently are inconsistent. This is not only so because individuals are usually subject to obligations derived from many, possibly uncoordinated, sources. But one and the same source of obligations, like a legal code, may actually issue incompatible norms, a fact well-known to jurists, though strangely denied by some legal philosophers.[15] So the charitable interpretation of deontic logics containing the schema oD must be the prescriptive one: a self-imposed restriction to the invariant features of *consistent* norm sets. This is also the view of van Fraassen who writes about his logic of conditional obligation, which contains a conditional version of oD, that "it would be more apt to say that we have here a logic of obligations that remain after obligational conflicts are resolved."[16] Similarly, Hansson proposes Standard Deontic Logic (**SDL**) – which is the modal system **K.KD** in the language $\mathbf{La}^{o}$ – as descriptive of consistent ("rational") sets of norms (that is, as *pre*scriptive in Alchourron's sense):

> "I will here take the view that deontic statements (formulas of **SDL**) are descriptive, that they describe what is obligatory, forbidden and permitted respectively, according to some (undetermined) system of norms or moral or legal theory. [...] The deontic axioms which will be discussed later, then, do not have the status of logical truths, but they express properties of the norm systems used. Those who are attracted by the axioms may then, if they so want, regard them as criteria of rationality or of inner coherence of norm-systems or moral or legal theories." ((1971), p. 123)

The problem with a logic of obligations restricted to consistent norm sets, is that it is largely useless as a meta-ethical tool. As pointed out earlier, sufficiently complex norm sets are likely to be inconsistent and thus oD is likely to lead us astray. What we need arguably more than a meta-theory of ideal norm sets, is a meta-theory of norm sets as they actually are and play their role in guiding human action. This is not to reject the consistency of norm sets as a regulative ideal. But striving to actualise

---

[15] Most notably Kelsen (1960), pp. 209ff., 280, 329. See also von Wright (1963), p. 148. In fairness it must be said that many legal philosophers keep, like jurists, to the facts; cf. e.g. Mautner (1971) and (1973). For a sample of moral dilemmas, presented with a view to refute the consistency thesis, see Routley and Plumwood (1984), §3. More and very acute moral dilemmas may be extracted from the essays in Lockwood (1985).

[16] (1972), p. 436.

regulative ideals is a dynamic process: it requires a gradual change of non-ideal sets to better approximations to the ideal. The recognition of such a regulative ideal is thus better accommodated in a theory of how norm sets ought to change, a topic that will be treated in chapter IV.

We shall now inquire into the basic invariant properties of norm sets. As a result, a basic set of principles concerning the operator o will emerge. By combining these principles with a base logic **L**, we obtain a basic logic of obligations *modulo* **L**.

Our schema (O),

$A \in O_{X,t}$ iff X accepts at t that it ought to be that A,

suggests an intimate relation between norm sets and acceptance sets of a certain kind: the norm set of an agent comprises his beliefs, and his ensuing commitments, about what ought to be the case. It can come as no surprise then that norm sets are closed under a notion of commitment in the same sense in which acceptance sets are thus closed. Thus, norm sets are adjunctive, detached, and closed under provable implication. That is, if, according to X at t, it ought to be that A and it also ought to be that B, then, according to X at t, it ought to be that $A \& B$. Similarly, if $A \rightarrow B$ is in $O_{X,t}$, then if it ought to be that A (according to X at t), then it ought to be that B (according to X at t). And if A logically implies B, then X's commitment to 'A ought to be true' incurs a further commitment to 'B ought to be true'. The basic logic of obligation thus turns out to be isomorphic to the basic acceptance logic $\mathbf{L.R}^a$. This basic logic of obligation, $\mathbf{L.R}^o$ may be axiomatised by adding to postulates for **L** the following schemas concerning the operator o:

RM. $\dfrac{A \rightarrow B}{oA \rightarrow oB}$

oC. $oA \& oB \rightarrow o(A \& B)$

oI. $o(A \rightarrow B) \rightarrow .oA \rightarrow oB$.

Schotch and Jennings (1981) have argued against oC on the following grounds. From RM and oC we can derive

oC! $o(A \& B) \leftrightarrow oA \& oB$.[17]

But given oC!, the consistency postulate

[17] Jackson and Pargetter (1986) have presented *prima facie* counterexamples to the right-to-left direction of oC!. Accordingly, they have proposed to reject the rule RM for the obligation operator. But, as Jackson (1988) has since shown, these alleged counterexamples to the converse of oC depend either on tense shifts or on misidentifications

oD. $oA \rightarrow {\sim}o\,{\sim}A$

and the principle that the impossible cannot be prescribed,

oF. ${\sim}o\,(A\&{\sim}A)$,

are classically equivalent. (Proof: ${\sim}o\,(A\&{\sim}A)$ iff ${\sim}(oA\&o\,{\sim}A)$ [by oC! and replacement] iff ${\sim}oA \vee {\sim}o\,{\sim}A$ [by DeMorgan] iff $oA \rightarrow {\sim}o\,{\sim}A$ [by the classical equivalence of ${\sim}A \vee B$ and $A \rightarrow B$].) Schotch and Jennings want to retain RM, oF and the classical basis of deontic logic; they want to reject oD. So in order to block the proof of the equivalence of oD and oF, they reject oC. Of course, if this line of thought is to persuade anyone to give up oC, independent reasons for its rejection will be needed. Schotch and Jennings attempt to provide such reasons by motivating their "non-Kripkean" semantics for deontic logic as follows:

> "Perhaps the most direct approach to the semantics of deontic logic is to avoid [oC] by making the 'ought' operator ambiguous. This matches the diagnosis of conflicts of obligation as the result of employing two or more distinct theories to evaluate ought sentences. Thus we might commit ourselves to several moral theories at once. Alternatively, it might be the case that our moral and political (and also perhaps religious) views compete in some cases in the evaluation of oughts." ((1981), p. 156)

Schotch and Jennings' rejection of oC is thus based on a fragmentation approach (in the sense of sec. 14.3) to an agent's total theory of obligations: an agent's total norm set is broken up into different fragments, according to their sources. A formula oA of the deontic language La^{o} is, according to Schotch and Jennings, better understood as a disjunctive (or existential) claim, viz. that A ought to be the case according to *some* fragment $O^{i}{}_{X,t} \subseteq O_{X,t}$. It is easy to see now how oC may fail on this view: while A may be in $O^{1}{}_{X,t}$ and B may be in $O^{2}{}_{X,t}$, $A\&B$ may not be in any fragment of the total norm set $O_{X,t}$.

The deontic fragmentation theory is subject to essentially the same objections adduced earlier against Lewis' and Stalnaker's fragmentation theory pertaining to acceptance sets. Firstly, the observation that agents draw their views as to what ought to be from a variety of sources and that they evaluate oughts by means of possibly incongruous considerations, has no bearing on their commitments once these views are formed. As in

of the subjects of obligations. Thus, given the requisite care with which norm sentences in natural languages are to be formalised, there is no need to reject RM.

the case of acceptance generally, so in the case of norm-acceptance: if an agent accepts both that it ought to be that A and also that it ought to be that B, then 'it ought to be that $A\&B$' is *ipso facto* accepted too – no use to plead fragmentation or, less euphemistically, schizophrenia.[18] Secondly, the fragmentation strategy does not even allow us to intuitively separate oD from oF without the additional assumption that all fragments of total norm sets are consistent. This assumption is indeed made by Schotch and Jennings. But in the absence of any arguments for the assumption, we may comment by way of quoting a remark of Schotch and Jennings' about the thesis that *total* norm sets are always consistent:

> "It is mere stipulation to insist that of two apparently conflicting obligations one will finally emerge as absolute and override the other, *prima facie*, one. As Russell remarks: 'the method of postulating what we want has many advantages; these are the same as the advantages of theft over honest toil'. " ((1981), p. 155)

When intuitively motivating the idea of fragmentation, Schotch and Jennings suggest, like Lewis and Stalnaker, that fragmentation occurs for reasons other than latching onto an inconsistency. In fact they suggest that norm sets are fragmented according to the sources – like legal codes, social norms, articles of faith, etc. – from which an agent draws his oughts. That these sources may be inconsistent in themselves has been pointed out earlier. But when finally trying to cash in on the fragmentation approach (e.g. by claiming that oD may be rejected while maintaining the "core principle" oF), it emerges that fragments are to be consistent *by definition* (the method of stipulation!) – whereupon all initial motivation of the fragmentation theory evaporates.

15.2. The choice of a base logic

It may be speculated that the standard self-imposed restriction to considering consistent systems of norms only, is hardly due to a disdain for systems of norms as they actually are. Rather, there is a sense in

[18] Strangely enough, Schotch and Jennings themselves insist on the commitment ingredient (explicated in terms of logical consequence) in the notion of obligation. In defence of the rule RM they write: "On the deontic interpretation the principle reflects the fact that logically necessary conditions of sentences which ought to be the case also ought to be the case. This seems right and indeed useful in moral philosophy, for by means of this axiom we may persuade moral agents that they are committed to the logical consequences of their moral principles. Without at least this much, moral philosophy would be a very curious endeavour." ((1981), p. 151)

which it is the best policy to follow, once one has made up one's mind that deontic logic is to be based on classical logic. For, suppose the restriction to consistent norm sets were lifted by rejecting the schema oD. Given classical logic and a minimum of modal principles we can quickly derive very strange assertions about what ought to be the case according to a norm set that is inconsistent with respect to some sentence: everything ought to be the case! For,

*I. $oA \& o \sim A \rightarrow oB$

is a theorem of any logic of obligation based on **K** and containing the schemas

RE. $$\frac{A \leftrightarrow B}{oA \leftrightarrow oB}$$

and

oC!. $oA \& oB \leftrightarrow o\,(A \& B)$

(for a proof see sec. 13.2.). So classically based deontic logics are better not applied to inconsistent norm sets.

It would be patently absurd to stick to classical logic and maintain that agents whose views as to what ought to be the case do not discriminate between what ought to be and what ought not to be. Thus, under the descriptive interpretation, a logic of obligation had better be based on a paraconsistent logic. I shall now argue for the stronger claim that a logic of obligation should be based on a *relevant* logic, that is, a logic which contains neither principles, like $A \rightarrow . \sim A \rightarrow B$, that trivialise inconsistent theories, nor principles, like $A \rightarrow B \rightarrow A$ (K), that offend a relevant theory of good inference.

For an argument against K, consider Vera, the passionate but compassionate truth-teller. Vera is in a conflict. Fred has died most painfully in a car accident. Coming to know the details of his death will make his parents suffer. Hence, compassion requires that Vera should spare Fred's parents the details of his death. But his parents have asked Vera for the details. So she decides that

(1) she ought to tell the truth.

Of course, her decision in favour of truth-telling is not a consequence of her other, conflicting obligation to shield Fred's parents from needless pain, i.e.

(2) it is not the case that if Fred's parents ought not suffer needless pain, then she ought to tell the truth.

Now, (1) is of the form oA and (2) is of the form $\sim(oB \rightarrow oA)$. But in virtue of K, every classical deontic logic contains the schema

K^o. $\quad oA \rightarrow .oB \rightarrow oA$.

So we may use K^o to infer from (1):

(3) if Fred's parents ought not suffer needless pain, then Vera ought to tell the truth,

contradicting (2)! Thus K may be instantiated to false principles about deductive relations between oughts.

K^o is simply an instance of the schema K; it does not depend at all on any specific postulates governing the operator o. Thus, a counterexample to K^o is simply a counterexample to the positive paradox K in a deontic setting. (The presence of the operator o helps nevertheless to make the paradoxical character of K more vivid.) By contrast, we shall present in the next section an argument against a schema of classical deontic logic that does not simply instantiate a schema already contained in **K**. The schema under attack will be $oA \rightarrow o(B \rightarrow A)$. Since it derives in one step by RM from K, I shall argue that K ought to blamed.

15.3. Extensions of L.R^o and further deontic operators

Whereas it seems natural to extend the basic acceptance logic **L.R**^{a} by the necessitation rule RN and certain principles concerning the iteration of the acceptance operator, notably a4 and a5, any such extensions of the basic logic of obligation **La**^{o} has sparked serious controversy.

Von Wright, for example, writes about the deontic version of the rule RN, from $\vdash A$ infer $\vdash oA$:

> "This always seemed to me highly counterintuitive, sheer nonsense. Most logicians, however, seem willing to swallow the absurdity – presumably for reasons of formal elegance and expediency. I cannot regard this as an acceptable ground." ((1981), p. 8)

RN is also omitted from Lemmon's first system **D2**, and the system he later favoured, i.e. **D2** without oD.[19] In a classical deontic logic, however,

[19] In our nomenclature, the system **D2** would be referred to as **K.RD**. The system occurs as a logic of obligation in both Prior (1955) (pp. 220ff.) and Lemmon (1957). It is easily shown equivalent to Hansson's Standard System without the deontic necessitation rule RN. Lemmon (1965) has proposed to weaken **D2** by dropping the schema oD. Thus, Lemmon's final proposal for a logic of obligation, the system **K.R**, is a basic system in our sense, based on classical logic **K**.

there is little to be gained by rejecting RN. For even the most minimal systems of deontic logic are closed under the monotonicity rule RM. Now let $\top$ be an arbitrary tautology. In virtue of the schema K (and MP), each formula of the form $A \to \top$ is a theorem of classical logic. Hence by RM, a classical deontic logic contains all instances of

oL. $oA \to o\top$.

That is to say, if anything at all ought to be, according to some agent, then that agent is committed to requiring that all tautologies ought to be – which is presumably the kind of "nonsense" von Wright has objected to. So if logical truths are obligatory on the minimal assumption that some state of affairs is obligatory, then we may as well adopt RN right away and say that even agents who have no substantial views as to what ought to be, are at least committed to hold that logical truths ought to be. The reason is simply that it is difficult to see how a case could be made against RN which would not equally disqualify oL as a thesis of deontic logic.

I believe that a case can be made against RN. It is simply that there is no *logical* route from the logical truth of a sentence A to a commitment on the part of every agent to always require that A. Here the close relationship between acceptance sets and norm sets breaks down. We have argued for the regularity of acceptance sets on the grounds that if the commitments of an agents arise in virtue of the theorems of a logic L, then the agent is *ipso facto* committed to accept these theorems as true. But it does not follow from the way commitments are determined by means of theorems of logic that an agent is therefore committed to maintaining that these theorems *ought* to be true (thus supporting obligations to the effect that one ought to see to it that logical truths be true!). As Lemmon, Prior, von Wright, and others have noted, it is very odd, to say the least, to say of something that is inevitable that it ought to be. The soundest deontic attitude to the inevitable seems to be no attitude at all; that is, to not require what is and will be true anyway. We should therefore not rule out as a matter of logic that the "no-attitude attitude" be taken, that sentences like $\sim o(A \vee \sim A)$ are sometimes true descriptions of norm sets.[20]

Our case against RN is equally a case against oL. Although it is worthwhile to point out that formally a distinction can be drawn between

[20] That deontic logics should allow for the "no-attitude attitude" seems also to underly von Wright's (1951) Principle of Deontic Contingency: "A tautologous act is not necessarily obligatory, and a contradictory act is not necessarily forbidden." (p. 11).

RN and Lo, to maintain that this distinction is of any philosophical substance would be pure sophistry: we cannot coherently reject RN without also rejecting oL. But oL is an immediate consequence of the schema K and the rule RM. RM is a prerequisite for making moral debate a worthwhile enterprise – without it we loose an important part of the grounds on which we may criticise agents for the moral theories they hold. Thus the case against RN is, by *modus tollens*, a case against K, i.e. against classical and for relevant deontic logics.

In contrast to the rule RN, controversies about the validity of principles in which the obligation operator is iterated are likely to remain inconclusive. This is so because even less than in the case of acceptance do we have any firm intuitions about what sentences such as

It ought to be that it ought to be that *p*

entail and what other sentences they are entailed by. We have a good grasp of what Joan says when she says

(1) John ought to accept that he ought not to put his feet on the desk.

But it is less clear what is involved when asserting a sentence like

(2) I ought to accept that I ought not to put my feet on the desk.

From the point of view adopted here, the problem with principles like

o4. $oA \rightarrow ooA$ and

o5. $\sim oA \rightarrow o \sim oA$

is that they appear to postulate logical relations between an agent's ethics and his meta-ethics. But such postulates clash with the descriptive conception of deontic logic. According to o4, it is a necessary condition for an agent to accept a certain norm that this norm is required according to the agent's meta-ethics. And according to o5, a norm can only be rejected, if the rejection is required according to the agent's meta-ethical standards. But we have argued above that a logic of obligation should take seriously the possibility of moral conflicts. A special kind of moral conflicts are conflicts between practical oughts and the kind of oughts (meta-ethical oughts) that are used to assess practical oughts. o4 and o5 rule out such conflicts between ethics and meta-ethics. Thus I am inclined to the view that both o4 and o5 should be rejected.

We shall now extend the logic of obligation to a deontic logic proper, that is, a formal system that formalises not only the logical properties of oughts but also the logical relations between oughts and sentences involving other deontic words, such as 'forbidden', 'permitted', or 'indifferent'. We first observe that ought and negation combine to four basic modes of obligation:

$$o\ ,\ {\sim}o\ ,\ o{\sim}\ ,\ {\sim}o{\sim}.$$

Some of these modes have more natural readings in English than a combination of 'not' and 'ought'. Thus, we say that A is *forbidden* just in case A ought not to be. And we say that A is *permitted* just in case there is no obligation to the contrary.[21] So we may introduce new operators f (for prohibitions) and p (for permissions) by defining them in the way just suggested:

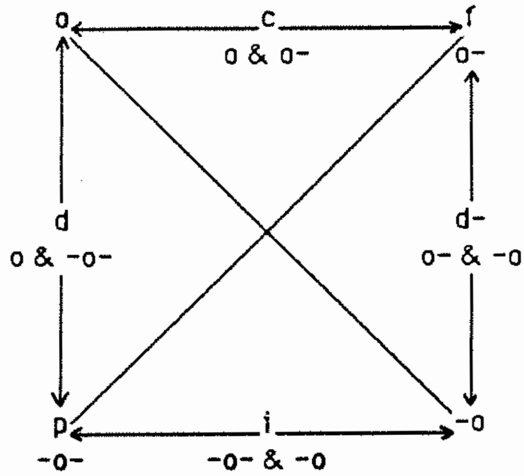
Dp. $\quad pA := {\sim}o{\sim}A,$

Df. $\quad fA := o{\sim}A,$

The basic logical relations between these and other, "intermediate", modalities are charted out in the diagram below.[22]

[21] Von Wright has disputed the interdefinability of 'ought' and 'permitted': "This question is in fact a classic problem of legal philosophy and theory. Do permissions (rights) have an independent in relation to prohibitions (obligations), or not? I think it is correct to say that opinions continue to be *very much* divided on this issue." ((1981), p. 7). I take this division of opinions as a sign that the notion of a permission is likely to be ambiguous. However the notion has to be disambiguated in a particular context, there is certainly one natural sense in which something is permitted just in case there is no obligation to the contrary.

[22] The diagram – isomorphic to our earlier epistemic square of oppositions – is an adaptation of Prior's (1955)(p. 220) deontic square of oppositions. Arrows indicate provable implications and diagonals connect contradictories. Prior's square does not include "intermediate" modalities and, given that he accepts the schema oD, the relations from o to p and from f to ${\sim}o$ are, to use the traditional term, subalternations.



If A is within the scope of a c-modality, then the agent in question is in a *conflict* as to whether A or $\sim A$ ought to be. Moral indifference is expressed by means of the i-modality: an agent holds no views as to whether A or $\sim A$ ought to be the case, i.e. he is *indifferent* with respect to A just in case neither A or $\sim A$ ought to be according to that agent's set of norms. A *definitely* ought to be if and only if A ought to be and $\sim A$ ought not be the case. Thus, apart from p and f we may define the further natural deontic operators c, i, and d as follows:

Dc.	$cA := oA \mathbin{\&} o\sim A$
Di.	$iA := \sim o\sim A \mathbin{\&} \sim oA$
Dd.	$dA := oA \mathbin{\&} \sim o\sim A$.

THEOREM 15.1.

Where **L** is a logic at least as strong as **DW**, all instances of the following schemas are derivable in $\mathbf{L.R}^o$ with the operators p, f, c, i, and d defined as above:

(1)	$cA \rightarrow oA$	(2)	$cA \rightarrow fA$
(3)	$dA \rightarrow oA$	(4)	$dA \rightarrow pA$
(5)	$d\sim A \rightarrow fA$	(6)	$d\sim A \rightarrow \sim oA$
(7)	$iA \rightarrow pA$	(8)	$iA \rightarrow \sim oA$
(9)	$pA \leftrightarrow \sim fA$	(10)	$oA \leftrightarrow f\sim A$
(11)	$iA \leftrightarrow \sim(fA \vee f\sim A)$	(12)	$iA \leftrightarrow pA \mathbin{\&} p\sim A$
(13)	$iA \leftrightarrow i\sim A$	(14)	$i\sim A \rightarrow \sim oA$
(15)	$cA \leftrightarrow fA \mathbin{\&} f\sim A$	(16)	$oA \rightarrow \sim(pA \mathbin{\&} p\sim A)$

(17) $dA \leftrightarrow oA \& pA$ (18) $\sim d \sim A \leftrightarrow \sim (o \sim A \& p \sim A)$

(19) $dA \rightarrow \sim d \sim A$ (20) $dA \& dB \rightarrow d(A \& B)$

(21) $d(A \rightarrow B) \rightarrow . dA \rightarrow dB$ (22) $\dfrac{A \rightarrow B}{dA \rightarrow dB}$

(23) $\dfrac{A \rightarrow B}{pA \rightarrow pB}$ (24) $\dfrac{A \rightarrow B}{fB \rightarrow fA}$

Proof. See T14.1. ■

The theorem may serve to show that the introduction of the defined deontic operators can turn a rather austere logic of obligation into a rich logic of several natural deontic notions. Among the deontic notions that are definable in **L.R**^{o} is that of a consistent obligation, represented by the "definitely ought" operator d. The theorem reveals that in **L.R**^{o}, a deontic logic in which no assumptions about the consistency of norm sets are made, we can actually choose to focus on consistent obligations and reason about *these* obligations in just the way recommended by the Standard System (of Hansson and *alia*) – modulo the now relevant base logic, of course. The schemas (19) to (22) provide exactly the axiomatic characterisation of the operator d that one would expect when consistent norm systems only are considered.

Chapter IV

The logic of theory change

16. Three kinds of change: expansion, contraction, and revision

16.1. An overview

Good theories are subject to quality control: they should be consistent with what we believe to be true, and they should not leave out reliable information that is pertinent to their particular subject matter. As new information comes in, our theories need updating: they change.

Sometimes we radically change our mind about how to account for the phenomena. Such changes of mind may issue in a wholesale abandoning of the conceptual structure of the theory we have developed so far. Such changes - exchanges - of theories ("paradigm shifts") will not be investigated here. Instead, I shall be concerned with piecemeal adjustments: changing a theory sentence by sentence in the light of new information, while leaving as much of what is already in the theory undisturbed as is compatible with accounting for the new information.

For the time being, let us think of a theory as a set of sentences closed under logical consequence. (I shall argue presently that for the purpose of developing a formal theory about how theories ought to change, this may be a too austere conception of a theory. But for the classification of types of change, to be given now, adopting the austere conception will do no harm.) We define the closure of a set of sentences under a consequence operation Cn as follows:

Definition 16.1. Consequence operation

Let **L** be a logic (in the sense of Part One) and let Δ be a set of sentences. For each sentence A,
$A \in Cn_{\mathbf{L}}(\Delta)$ iff there is a set $\{A_1, \ldots, A_n\} \subseteq \Delta$ such that $A_1 \& \cdots \& A_n \rightarrow A \in \mathbf{L}$ (for $1 \leq n < \omega$).

We shall suppress the subscript **L** to Cn whenever we talk generally about sets closed under *any* consequence operation induced by a logic, or when the class of consequence operations under consideration is clear from the context.

Suppose now that we have come into the possession of a piece of information A and that the theory T we have endorsed so far is incomplete with respect to our new piece of information, i.e. both $\sim A \notin T$ and $A \notin T$. Then we may just add A to T and close the result under logical consequence again. Such a simple change will be called an *expansion* (of T by A). We shall use + for the expansion operation and define

(Def +) $\quad +(T,A) := Cn(T \cup \{A\})$

By an harmless equivocation of notation, I shall write '$T+A$' for the value of $+(T,A)$.

Suppose next, that $A \in T$ and that we have obtained information that induces us no longer to accept A as true. In such a situation we should want to retract A from the theory T. Just as the expansion operation, as just defined, is always successful, in the sense that $A \in T+A$, and produces a theory $T+A$, so we should want a *contraction* operation, $-$, to produce a new theory $T-A$ which no longer contains A. But it is easy to see that the definition of such a contraction operation can not be as straightforward as the definition of +. For, T may contain sentences $A_1,...,A_n$ which entail A, and so at least some of these sentences will have to be removed together with A, if $T-A$ is to be a theory (i.e. closed under Cn) not containing A. Furthermore, some sentence B may be in T just because A is in T. But if B's subsistence in T is parasitic on A being in T, then B should be removed from T together with A.[1]

An expansion operation on a theory results in a richer theory: sentences are added, nothing is taken away from the original theory. By contrast, a contraction operation cuts a theory down to a subset of the original theory; sentences are withdrawn from and no new sentences are added to the original theory. But frequently we want to add a sentence to a theory while removing all sentences incompatible with the sentence to be added from the theory. When information becomes available which is incompatible with certain pieces of information in our theory, and when we decide that this recent piece of information overrides all conflicting earlier information, then the theory in question needs to be revised. The *revision* of a theory is obviously a composite change-operation: it involves both subtraction from and addition to a theory. If $\sim A$ is in T

[1] This latter claim is not completely uncontentious and I shall defend it in a moment when recommending a filtering condition on contractions.

and if we want to revise our theory to include A, then we should first remove $\sim A$ from A and then add A to T. Thus it appears natural to define the revision operation * by means of the Levi identity,

(LI) $T{*}A := (T - {\sim}A) + A$.[2]

Expansions are obviously not more problematic than the choice of the underlying consequence operation by which they are defined. Revisions I take to be decomposable into contractions and expansions, as specified by the definition (LI). Thus, the proper focus for a theory of how theories ought to change in the light of new information is the notion of contraction.

16.2. The Gärdenfors postulates for contraction

There are two ways in which we may proceed in an attempt to shed some light on the formal properties of the contraction operation. To everyone who is engaged in theorising, contraction is – unfortunately – a familiar operation. This familiarity supports the expectation that we should intuitively recognise a certain set of assertions involving the notion of contraction as essential to an understanding of that notion. Thus, we may try to *implicitly characterise* the operation of contraction by capturing our intuitions in a list of postulates. The following postulates for contraction have been proposed by Gärdenfors.[3] We assume T to be an arbitrary theory, i.e. $T = Cn(T)$.

Postulates for contraction

(-1)	$T - A$ is a theory	(Closure)
(-2)	$T - A \subseteq T$	(Inclusion)
(-3)	If $A \notin T$, then $T \subseteq T - A$	(Vacuity)
(-4G)	If $\nvdash A$, then $A \notin T - A$	(Success)
(-5)	If $Cn(A) = Cn(B)$, then $T - A = T - B$	(Preservation)
(-6)	$T \subseteq (T - A) + A$	(Recovery)

[2] That revisions ought to be decomposable into contractions and expansions, as laid down in (LI), was first suggested in Levi (1977).

[3] See Gärdenfors (1982). I adopt the numbering of the postulates as they appear in the 1985 survey article by Alchourron, Gärdenfors and Makinson.

Before we move to an alternative approach to determining the properties of the contraction operation, a few remarks on these postulates may be in order. (For a more extensive discussion of these postulates see Gärdenfors (1988), section 3.4.)

The closure postulate (-1) requires that the contraction operation should map a theory not into any old set of sentences but again into a theory.

Inclusion, (-2), encodes the idea that a contraction operation does not add anything to a given theory.

If the sentence to be retracted from T is not in T, then the contraction operation is vacuous, i.e. $T-A=T$; thus, the vacuity postulate (-3).

If A and B are logically equivalent, then, for any theory T, they are entailed by exactly the same set S of sentences in T. Now, in order to avoid commitment to either A or B, we shall have to remove from T some of the sentences in S. The preservation postulate (-5) requires that our choice of sentences in S that we should want to give up, ought not to depend on whether we want to avoid commitment to A or to B.

Contractions should be successful – but, so (-4G) says, only if the sentence to be removed is not a logical truth; logical truths can not be retracted. This limitation to the success of a contraction operation is due to the definition of the consequence operation used by Alchourron, Gärdenfors and Makinson (henceforth, AGM). AGM define theories as sets of sentences closed under a *regular* consequence operation induced by classical logic **K**. The definition of a regular consequence operation, RCn, results from our definition D16.1 by an apparently minute alteration.

DEFINITION 16.2. Regular consequence operation

Let **L** be a logic (in the sense of Part One) and let Δ be a set of sentences. For each sentence A,
$A \in RCn_{\mathbf{L}}(\Delta)$ iff there is a set $\{A_1, \ldots, A_n\} \subseteq \Delta$ such that $A_1 \& \cdots \& A_n \rightarrow A \in \mathbf{L}$ (for $0 \leq n < \omega$).

The difference between Cn and RCn resides in the last clause determining the range of the index n in the respective definitions. In D16.2, n is allowed to be 0, in which case membership in **L** suffices for membership in $RCn(\Delta)$ for any set Δ, including the empty set. So if theories are defined as closures under a regular **L**-consequence operation, then theories always include all theorems of **L**. Thus the postulate (-1) forces a qualification of the requirement that contraction operations be successful: $A \notin T-A$ only if A is not a theorem of the underlying logic.

Given that AGM take **K** as the logic theories are closed under, there is, in one sense, not much to choose between a consequence operation in the sense of D16.1 and a regular consequence operation as defined in D16.2. For, in virtue of the classical schema $A \rightarrow \mathrm{T}$ (where T may be replaced by any classical tautology), the set of all classical tautologies is a subset of $Cn_{\mathbf{K}}(S)$ for any nonempty set S. This is the sense in which closing under **K**-consequence and closing under regular **K**-consequence makes no difference for "most" sets of sentences – that is, all but the empty set.

However, when subjecting theories to contraction operations, it does make a marked difference whether theories are closed under $Cn_{\mathbf{K}}$ or under $RCn_{\mathbf{K}}$. In contrast to closure under $RCn_{\mathbf{K}}$, if we close under $Cn_{\mathbf{K}}$, then we can embrace the unconditional

(-4) $\qquad A \notin T-A$

even when A is a classical tautology. When A is a thesis of **K**, then contracting T by A requires contracting to the empty set, since any sentence whatsoever implies A according to **K**. Thus, trivially $A \notin T-A = \varnothing$.

The reverse side of this classical coin – viz. that contractions by tautologies must be contractions to the nullset – is that successful contractions of inconsistent theories must be contractions to consistent theories. Suppose T is an inconsistent set of sentences closed under $Cn_{\mathbf{K}}$. Since any sentence A is a **K**-consequence of any inconsistent pair of sentences $B, {\sim}B$, a contraction by A can only be successful once all inconsistencies have been eliminated from the theory in question. Just as closure under $Cn_{\mathbf{K}}$ prevents removing tautologies from a theory without removing everything, so it prevents removing any sentence from a theory without changing the theory to an absolutely consistent one.

Throughout this chapter we shall assume that theories are closed under an operation $Cn_{\mathbf{L}}$, where **L** may be any logic in the sense of chapter one. In fact, for most of the formal results – that is, for all positive results – logics even weaker than our minimal relevant logic **BM** will suffice: all we shall need occasionally, are elementary (lattice-) properties of $\&$ and $\vee$. For some examples, however, we shall assume a stronger notion of theory according to which theories are detached: is

$A \in T$ and $A \to B \in T$, then $B \in T$. Closure under a logic containing the schema

WI. $A \& (A \to B) \to B$.

will ensure that property.

Although theories closed under classical consequence thus fall within the scope of the present investigation, we recommend – in view of the just mentioned anomalies – that theories be closed under a logic which contains neither the schema $A \to .B \to A$ nor $A \to .{\sim}A \to B$ or any of their cognates.

To revert to our discussion of the Gärdenfors postulates, the recovery postulate (-6) is perhaps the least plausible in Gärdenfors list. In conjunction with (-2), it requires that T should be recoverable from $T{-}A$ by adding A to $T{-}A$ and closing the result under logical consequence. The postulate is implausible because in order to retract A, we may have had to retract a sentence B which is strictly stronger than A, where B is strictly stronger than A in T just in case $B \to A$ is in T or a theorem of our logic while neither is $A \to B$ in T nor is it a logical theorem. In such a situation there seems to be no prospect for getting B back by adding A to $T{-}A$ and closing the result under logical consequence. For Gärdenfors, however, Recovery reflects the maxim that a contracted theory should be as large a subset of the original theory as possible under the circumstances; contractions should be *strictly minimal changes*. In particular, a contraction operation should always leave as much in the contracted theory as to guarantee complete recovery after adding the contracted sentence again. But such a guarantee, as we shall see, can not in general be given without invoking essentially some of the paradoxical aspects of classical logic.

16.3. Explicit definitions of the contraction operation and the intuitive process of contraction

The alternative to an implicit definition of the contraction operation by means of a set of postulates, is to provide an *explicit definition* which allows us to construct $T{-}A$ for any theory T and sentence A. Again, our practice of theorising suggests that we do have some intuitions about *how* to contract a theory. I shall presently outline what I take to be the intuitive process of contraction. It may, however, be questioned whether it is required, helpful, or even possible to make an explicit definition of a contracted theory conform with our intuitions about the process of contraction. Consider an analogy with grammatical theory. A correct grammar for a language is a set of rules which generates those and only those expressions which a competent speaker of that language judges as well-formed. Intuitive judgements about well-formedness by a native speaker constitute the set of data to be accounted for by the grammar. It is an altogether different – and deep – issue whether a particular grammar is "psychologically real", that is, makes explicit a speaker's tacit knowledge of the rules of the language he speaks, or whether the grammar merely "saves the phenomena". Indeed, a grammar may offend a speaker's intuitive understanding of how he forms expressions of the language, while delivering nonetheless just the right criterion for well-formedness. Similarly, we may give a recipe for how to contract theories which pays no attention at all to – and may even be inconsistent with – our intuitive understanding of that process. Such a recipe, it may be argued, would be vindicated by its satisfying our postulates for contraction, and by delivering counterexamples to postulates which we emphatically reject, and, perhaps, by satisfying only those postulates we have explicitly endorsed.

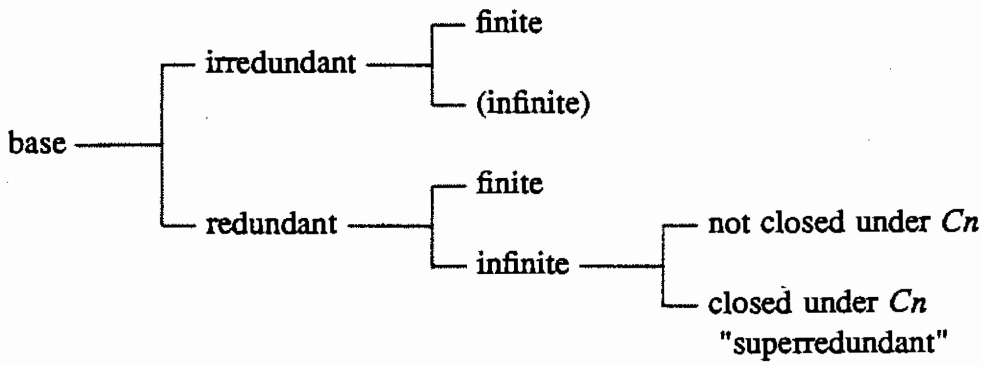
But if this is the only way by which we may vindicate the construction, then we can have no more confidence in the construction than we have in the postulates which it satisfies. We have seen, however, that at least one of the Gärdenfors postulates, namely Recovery, is eminently disputable. Ideally then, we should want a construction of the contraction operation which is supported independently of – and perhaps even better than – the more dubious of Gärdenfors' postulates. Thus, we are well-advised to strive for intuitive plausibility when devising an explicit recipe for how to contract theories.

What intuitions we have about theory change, do not pertain to infinite sets of sentences closed under logical consequence. Infinite sets are simply too big to survey. And survey we need when we are asked to pick out those sentences in a set which we are prepared to give up.

When talking about theories in any practicable manner, we usually mean sets of sentences generated from a finite base by means of the logical consequence operation. To hold a theory is to accept some surveyable set of sentences and to accept commitment to any sentence that is entailed by that set.

To outline the intuitive process of contraction it will be useful to have some precise terminology and a few pieces of notation at hand. By a *theory* I shall mean, as before, a set of sentences closed under logical consequence. A *base* for a theory T is any subset S of T such that

$Cn(S)=T$. Where T is any theory, I shall use 't' to denote a base for T. An *irredundant base* is a set of logically independent sentences: for each sentence A in an irredundant base t, $A \notin Cn(t—\{A\})$ (the long horizontal '—' stands for set subtraction). It is well-known that there are theories which can be irredundantly but not finitely axiomatised.[4] Such theories I shall leave out of consideration here: irredundant bases will be assumed to be finite. Redundant bases, however, may be finite or infinite. The limiting case of an infinite redundant base for a theory T is T itself. A base closed under logical consequence, I shall call *superredundant*, which is just another but sometimes convenient way of saying that such a base is a theory. This classification of bases is summarised in the diagram below.



Human limitations require that we explore the theories we are prepared to accept from the vantage point of sufficiently small bases. 'Sufficiently small' must mean at least finite; ideally it means irredundant. The theory generated from such a base is largely *terra incognita* and to a very large extent even completely uninteresting. A may be an important piece of information which we want to incorporate into some base t. We may then go on to explore some of the consequences of t extended by A. But it is unlikely that we would care to deduce any of the disjunctions $AvA_1, AvA_2, \ldots, AvA_1vA_2v \cdots$.

Thus, in sketching out our intuitions about the process of contraction, we need to focus upon theories generated from small bases. For, we do not have, indeed can not have, any intuitions about the contraction of theories as such. It will have to be seen then, if and how these intuitions can be extended to theories generated from large bases, including infinite

[4] Certain many-valued logics are examples of such theories. The observation that there are classical theories which are not finitely axiomatisable is originally due to Lindenbaum (cf. Tarski (1930), p.88).

and in particular superredundant bases.

It must be emphasised that a definition of the contraction operation for theories on small bases should not be regarded as a mere stepping stone towards defining contraction for theories on superredundant bases. For all *practical purposes* the former is what we need, while the latter is a special case of contractions on infinite bases, a notion for which no useful applications are yet in sight.

Among the practical purposes of providing a recipe for constructing contracted sets, I count the design of a program that updates databases. The recent interest in a theory of theory change has been much enhanced by the fact that essentially the same questions to be answered by such a theory arise in certain areas of artificial intelligence, such as knowledge representation. A database implemented on a machine is a finite ensemble of data structures representing what the machine "believes" at a particular point of time. (Usually a database is split up into a set of interrelated files. But this aspect of real-life databases need not interest us here.) Information encoded in a database may be retrieved by posing queries. If a database is coupled with an "inference engine", then the database can supply information which is not explicitly stored in it. A database together with an inference engine is thus just like a based theory, the base being the database, and the theory being the total set of information retrievable by means of the database *cum* inference engine.[5] And just as theories need updating when new information becomes available, so do databases. Not only will it be heuristically helpful but also useful in view of potential applications, to develop a theory of theory change with the problem of database updating in mind.

When T is generated from a base t, I shall say that T is a theory on t. Suppose then that some sentence A is a member of a theory T on a base t (that A is a consequence of t) and that we have decided to remove A from t. As a first step towards determining which sentences we should remove from t in order to avoid commitment to A, we should try to establish the strongest results about how A can be derived from t. That is, we are looking for derivations $<A_1, \cdots, A_k, A>$ such that $A_1, \ldots, A_k \in t$ and each of $A_1, \ldots, A_k$ is essential (important, irredundant, relevant, etc.) for the derivation of A. Before we proceed any further with our outline of the intuitive contraction process, it will be useful to define some concepts.

[5] Cf. section 14.1.

A sentence A is derivable from a set of sentences S just in case $A \in Cn(S)$. Any set S such that $A \in Cn(S)$ will be called an *A-consequence set.* A sentence B is *essential* in an A-consequence set if and only if $A \notin Cn(S—\{B\})$. I shall call a set of sentences S an *entailment set* for A if and only if all sentences in S are essential for the derivation of A, i.e. no subset of S will suffice to derive A.

DEFINITION 16.3. Entailment set in t

A set of sentences S is an *entailment set for A* in a base t if and only if

(a) $S \subseteq t$, and

(b) $A \in Cn(S)$, and

(c) for all $S' \subset S$, $A \notin Cn(S')$.

I shall denote the set of all A-consequence sets in t and the set of all entailment sets for A in t by $D(A)_t$ and $E(A)_t$ respectively. (The reference to the base t will be omitted when misunderstandings can safely be precluded.) Note that the definition of an entailment set for A in t is equivalent to that of a *minimal A-consequence set.* A set S is minimal in $D(A)_t$ just in case (i) $S \in D(A)_t$ and (ii) for all $S' \in D(A)_t$, if $S' \subseteq S$, then $S \subseteq S'$ (i.e. $S'=S$).

Proof. Clearly, (i) entails (a) and (b) of D16.1 and *vice versa.*

(c) $\Rightarrow$ (ii):

(1)	$(\forall S')(S' \in D(A)\ \&\ S' \subset S \Rightarrow A \notin Cn(S'))$	Assumption
(2)	$M \in D(A)$	Assumption
(3)	$M \subseteq S$	Assumption
(4)	$A \in Cn(M) \Rightarrow not(M \subset S)$	contraposing (1), M/S'
(5)	$not(M \subset S)$	(2),(4)
(6)	$S = M$	(3),(5).

(ii) $\Rightarrow$ (c):

(1)	$(\forall S')(S' \in D(A)\ \&\ S' \subseteq S \Rightarrow S \subseteq S')$	Assumption
(2)	$M \subset S$	Assumption
(3)	$M \notin D(A)$ or $not:(M \subseteq S)$	(1),(3)
(4)	$M \notin D(A)$	(2),(3)
(5)	$A \notin Cn(M)$	(4), Def. D(A). ■

I resume now the description of the intuitive process of contraction. Suppose we want to give up A and we have determined the set $E(A)_t$ of all entailment sets for A in t. Since we are seeking an intuitive description of the kind of contraction processes we can be familiar with, t must be assumed to be finite; hence, $E(A)$ will be a finite collection of finite sets. Let '$\cup E(A)$' stand for the set of all sentences in t that are in some entailment set for A, i.e. $\cup E(A) := \{B : \exists S \in E(A) \,\&\, B \in S\}$. Clearly, every sentence in $\cup E(A)$ is in principle a candidate for retraction. Some sentences, however, we are prepared to give up more easily than others; they have a higher degree of *retractibility* than others. For example, high-level, tentative generalisations may be rejected in favour of undeniable brute facts. On the other hand, high-level generalisations make a theory powerful and interesting. Thus, we may be inclined to retain such generalisation at the expense of conceding error with respect to some sentence expressing a low-grade observation. If some sentence has surfaced repeatedly in entailment sets for undesirable sentences, then this sentence has presumably accumulated a high degree of suspiciousness and we should assign it accordingly a high degree of retractibility. Whatever the criteria by which we judge retractibility are, we shall have to make a choice among the rejection candidates in $\cup E(A)$. Formally, we can represent this discrimination with respect to retractibility among sentences in $E(A)$ by means of a *comparative retractibility ordering*. In practice, we need not bother with ordering the sentences in a base with respect to their comparative retractibility before the need for contraction arises; when contracting a base by A, the ordering of sentences outside $\cup E(A)$ is irrelevant for the purpose at hand. Accordingly, a "user-friendly" contraction program would minimise the ordering task by asking the user to order the sentences in $\cup E(A)$ only (the program would thus have to be interactive). For the purpose of developing our formal theory, however, we shall assume that the sentences in a base are ordered according to their comparative retractibility.

Suppose that we have chosen out of $\cup E(A)$ those sentences which we are prepared to give up. Call the set of these sentences 'the *reject set* for A in t' (notation: $R(A)_t$). If we subtract $R(A)$ from the base t, we obtain a new base $t—R(A)$. If the theory T is determined by the base t, then the contraction of T by A, $T-A$, is determined by the base $t—R(A)$, i.e. $T-A := Cn(t—R(A))$.

The stages of this procedure are summarised in the flowchart diagram below.

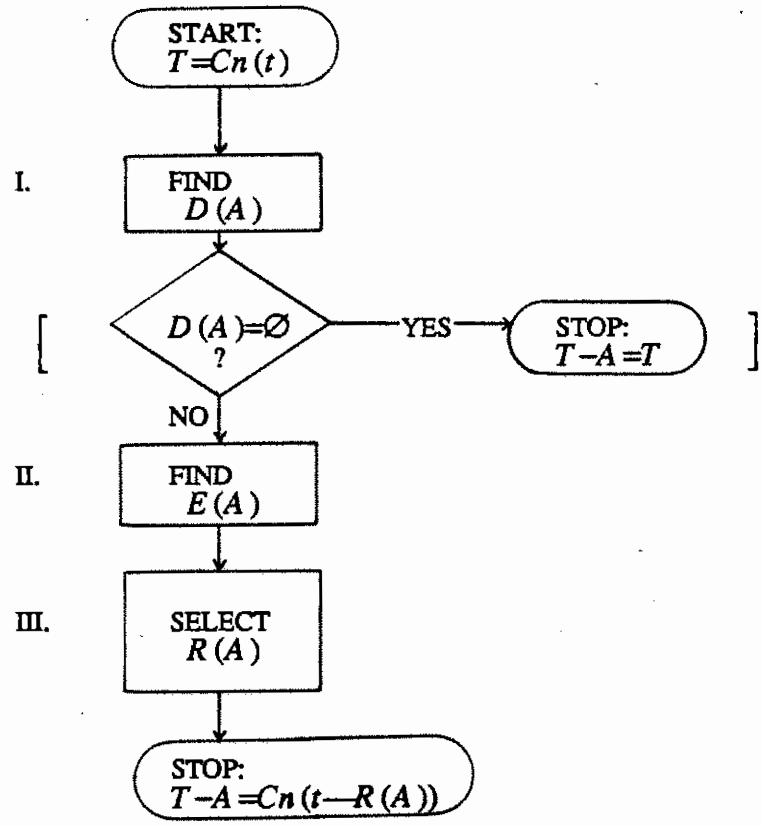


FIGURE 16.4
Given T on t, find $T-A$.

The decision diamond between the processes I and II is just an optional speed-up device. If there is no subset of t of which A is a consequence, that is, if $A \notin T$, then the procedure may be exited with the result that $T-A$ is set to T. This early exiting is just a matter of convenience, since, if $D(A)$ is empty, so are $E(A)$ and $R(A)$; thus, the result is the same, whether we take the shortcut after step I or whether we continue with steps II and III.

The steps I and II of the above procedure are fully determined given the consequence operation Cn. Only step III allows for some discretion to be exercised. I shall consider now two plausible constraints on determining the reject set $R(A)$.

First, we naturally want $R(A)$ large enough to block derivations of A. Thus, we need to impose the following condition.

SUCCESS

For each entailment set S in $E(A)$, at least one sentence in S must be rejected.

This condition will trivially be met by any construction in which $R(A)$ is quite naturally determined to be the set of all most retractible elements in each A-entailment set. For, given an ordering on the base t under consideration, each subset of t contains at least one minimal (under the ordering) element.

But we should also endeavour to keep the loss of information in t at a minimum, that is, we need to keep $R(A)$ as small as we can, *modulo*, of course, the SUCCESS condition. How small we can keep $R(A)$ will depend on how finely we can discriminate among the sentences in $\cup E(A)$ with respect to their (comparative) retractibility. Quine's ((1950), p.14) example of Bizet and Verdi may serve to illustrate this point. Suppose that we want to retract from some theory T the sentence 'Bizet and Verdi were not compatriots' (let A stand for this sentence). Suppose further that $\cup E(A)$ contains the two sentences 'Bizet was French' (B) and 'Verdi was Italian' (C), and that $\cup E(A)$ is such that it suffices to retract one of B or C in order to block all derivations of C (in the base for T). But such an ideal situation may not always obtain. In fact, I, like most other people, find it difficult to make up my mind as to whether I should have more confidence in the truth of the one or the other sentence. As far as their retractibility is concerned, they are thus *en par*. So they should both be removed (followed, perhaps, by an expansion by BvC). Ideally, however, we should want the assignment of retractibility to all rejection candidates in $\cup E(A)$ be sufficiently fine-grained to retract *at most* one sentence from each entailment set for A. But it makes little sense to legislate that the ideal case always obtains. Nevertheless, we can issue the advice to make sure that the ideal is realised "as much as possible". In terms of an ordering $\leq$ of comparative retractibility we state this piece of advice as follows.

ADVICE

Compare (in terms of $\leq$) all sentences in a base and avoid ties.

The first part of the advice ("compare all sentences!") recommends that the ordering be connected.

CONNECTEDNESS

For all sentences A, B in a base t, either $A \leq B$ or $B \leq A$.

It is clear, however, that unless the second part of ADVICE ("no ties!") is also *fully* met, CONNECTEDNESS does not in general result in smaller reject sets. Thus connectedness will play no role in what follows.

There are conditions which would yield even smaller reject sets than the ones that are generated by retracting $\leq$-minimal elements only from each entailment set. In practice, we are usually well-advised to regard those sentences as prime rejection candidates which are common to all entailment sets. Thus, it may be tempting to propose the following condition.

(*) For each pair of entailment sets (S, S') such that $S \cap S' \neq \varnothing$, consider only sentences in $S \cap S'$ as candidates for rejection.

But a moment's reflection shows that (*) is too strong a requirement. For, even all (and not only some) derivations of A may proceed by means of some completely uncontroversial principle C. Thus, C would be in the intersection of all entailment sets for A while C is at the same time virtually immune to retraction. The strategy of assigning those sentences high retractability which occur in more than one entailment set yields sensible results only after we have already filtered out those sentences which are "beyond dispute" (usually tacitly assumed "background" principles in the sort of derivations we are presented with in practice).

Considerations like those *a propos* (*) show that minimality, as measured by set inclusion ("strict minimality"), is of no virtue. We are not interested in reducing the base for a theory by just as few sentences as are needed to secure that some unwanted sentence is no longer derivable. The sentences in a base must be thought of as associated with some measure of their importance for theorising, and we should be prepared to retain a highly important sentence at the expense of giving up more than one sentence of less importance.

This undermines those arguments for the recovery postulate, $T - A + A = T$, which proceed from the assumption that strict minimality should be an essential characteristic of contraction.[6] But quite apart from

[6] Such arguments for recovery have been advanced by Makinson (1977), pp. 383 and 391f.

any considerations to the effect that strict minimality is a misguided demand on a contraction operation, we can actually show how Recovery fails by means of simple counterexamples within the framework developed so far.

Let T be generated from the base $t = \{A, A \rightarrow B\}$ and let the target be $T-B$. Then t itself is the only entailment set (in t) for B and so either A or $A \rightarrow B$ (or both) has to be retracted from t. (We assume that the logic providing the closure operation for T contains the schema WI; thus T is detached. We also assume that one of $\{A, A \rightarrow B\}$ is more retractible than the other. For the example to be a counterexample to Recovery, however, this assumption is of no consequence.)

$$t = \{A, A \rightarrow B\}$$
$$\text{target: } T-B$$
$$E(B) = \{\{A, A \rightarrow B\}\}$$

(1)	(2)
$R(B) = \{A\}$	$R(B) = \{A \rightarrow B\}$
$T-B = Cn(A \rightarrow B)$	$T-B = Cn(A)$
but: $A \notin T-B+B$	but: $A \rightarrow B \notin T-B+B$

Both contraction strategies, (1) and (2), yield reduced bases which, after adding B, do not result in bases for the original theory T. If we followed strategy (1), we would need $B \rightarrow A$ in $T-B$ or as a theorem of our background logic **L** in order to recover A in $T-B+B$. If we followed strategy (2), we would need $A\&B \rightarrow A \rightarrow B$ in $T-B$ or **L** for $A \rightarrow B$ to be in $T-B+B$. Since $A\&B \rightarrow A \rightarrow B$ is a classical tautology, Recovery may be rescued for the second strategy by requiring **L** to be **S**. But for the first strategy Recovery is beyond such rescue operation: not even classically is A a consequence of B and $A \rightarrow B$. Note that this counterexample to recovery depends on two very weak assumptions only: (a) that contraction operations on theories proceed via bases which are not themselves closed under Cn (we do not even need to assume that bases must be finite), and (b) that contractions must be successful (in the sense of SUCCESS).

Given the existence of such simple counterexamples to Recovery, the reader may wonder how Recovery has ever found its way into Gärdenfors' postulates for contraction. The reason is that there are certain not completely unnatural ways of defining a contraction operation on classically closed theories which do satisfy recovery. Forget about bases for the moment and assume that theories are closed under classical logic. For each sentence $A \in T$, $A \vee {\sim}B \in T$ (that much holds also relevantly). Now consider all maximal subsets of T which do not have B as a consequence; $A \vee {\sim}B$ is in each such set (Lemma 2.1 in Alchourron, Gärdenfors and Makinson (1985)). No matter how you define $T - A$ in terms of these maximal subsets of T not entailing B (whether by taking the intersection of them all ("full meet contraction"), or by picking out a preferred set ("maxichoice contraction"), or by intersecting a number of equally preferred sets ("partial meet contraction")) – $A \vee {\sim}B$ will survive in $T - B$, for any A in T. But classically (though not relevantly) $A \vee {\sim}B$ is equivalent to $B \to A$; hence, $A \in Cn(T - B \cup \{B\})$.

This argument for Recovery depends on the denial of (a) above and the requirement that theories be classically closed. From the methodological viewpoint adopted here (and argued for earlier), rejecting (a) – and indeed *requiring* that the base from which a theory is generated should play no distinctive role in the process of contraction – subverts any intuitive grip we may have on the notion of contraction. The process of generalisation should go from the familiar (contraction of theories generated from surveyable bases) to the unfamiliar (contraction of theories generated from superredundant bases) – not *vice versa*. It has moreover the consequence that sentences, such as $A \vee {\sim}B$, which intuitively should have been retracted in the transition from T to $T - B$, remain in $T - B$.

To illustrate this latter point, consider a theory T generated from $t = \{A\}$. For any sentence B, $A \vee B$ is in T. Now suppose we want to contract by A. Clearly, the base for the new theory $T - A$ must be the empty set.[7] Thus, none of the disjunctions $A \vee B$ will be in the theory generated from the reduced base $t - \{A\}$ – just as we should expect. For, these disjunctions came into T just because A was in the base, and after the removal of A from the base, there is no reason why any of these disjunctions should remain in $T - A$. Indeed it would be positively misleading, and potentially dangerous, to keep these "inferential danglers" in the theory in the absence of any sentences from which they can be inferred. For if a theory contains the information that $A \vee B$ is true without containing either A or B, then one should be able to infer that $A \vee B$ has been included in T on non-logical grounds, that it is not just a trivial consequence of the stronger pieces of information that A or that B. This, however, can not safely be inferred if a contraction operation tolerates

[7] This particular example will satisfy the recovery postulate. But for the present purpose we do not need an example for which Recovery fails.

inferential danglers. The possibility of such danglers undermines the very point of allowing for non-prime theories (i.e. theories that may contain a disjunction without containing either disjunct).

Thus, we should make it a third adequacy condition that contracted theories be free from dangerous litter.

FILTERING

> If B has been retracted from a base t in order to bar derivations of A from t, then $T-A$ should not contain any sentences which were in T "just because" B was in T.

(FILTERING is, in a sense, a condition *dual* to SUCCESS. The success condition is concerned with pruning those sets of sentences of which the sentence to be contracted is a consequence; the filtering condition requires a pruning of those sentences which are consequences of the sentence to be contracted.)

It will presently emerge that the precise meaning of the phrase 'just because' must depend on whether the base in question is redundant or not (and if not, in which sense the base can be said to be irredundant). As a first approximation, however, we can offer the following explication in terms of the notion of 'dependency in a based theory'.

DEFINITION 16.5. Dependency in a based theory

> For any sentences B, C and theory T based on t:
> C is *dependent in T (T-dependent)* on B if and only if $C \in Cn(t)$ and $C \notin Cn(t - \{B\})$.

It stands to reason that this the right sort of explication of the phrase 'just because' when dealing with irredundant bases. For such bases then, the filtering condition takes the form

(F) If $B \in R(A)_t$ then $T-A$ should not contain any sentence C which T-depends on B; i.e. $T-A$ should not contain any sentence C such that $C \in T$ and $C \notin t - R(A)_t$.

In this form the filtering condition is not an extra requirement on the definition of the contraction operation but a quite trivial fact about the definition already offered. If the base in question is irredundant, then our definition of the contraction process ensures that the demand formulated in FILTERING will be met.

But in the case of redundant bases, (F) may not be strict enough to satisfy the intuitive import of FILTERING. Consider for example the base $t=\{A, A\vee B\}$ and suppose we remove A from t. It appears that we can take either one of two stances about the question whether $A\vee B$ should be in the A-reduced base. Which stance we take will depend on how we answer the question: would we have included $A\vee B$ in the base, if we had not included A? If the answer is No, $A\vee B$ should not be in t (and hence, not in T) after A has been removed; if the answer is Yes, $A\vee B$ should stay on. In section 18 a unified approach will be proposed in which both possibilities can be accommodated.

17. Contraction

17.1 Contraction on theories generated from irredundant bases

We recall that a base t for a theory T is irredundant if and only if for any sentence $A\in t$, $A\notin Cn(t-\{A\})$, or equivalently:[8] if and only if no proper subset of t is a base for T.

Let the notions of a *consequence set* and an *entailment set* be as before; i.e.

$$D(A)_t := \{S\subseteq t : A\in Cn(S)\}$$
$$E(A)_t := \{S\in D(A)_t : (\forall S'\subset S)(A\notin Cn(S'))\}$$

Now consider an ordering $\leq$ on t. Intuitively, $A\leq B$ should mean that A is at least as retractible as B (or, to borrow the terms of Gärdenfors (1984), A is not more *epistemically important* or *entrenched* than B). Given such an ordering on the members of a base, we define the *reject set for A (in a base t)* as follows:

$$R(A) := \{B : (\exists S\in E(A))(B\in S \text{ and } (\forall C\in S)(\text{if } C\leq B \text{ then } B\leq C))\}.$$

In English: $R(A)$ contains all those members from each entailment set S for A which are minimal (under $\leq$) in S.

[8] Theorem 32 in Tarski (1930), pp. 84f.

We can now define the contraction by A of the theory T as generated from the base t:

$$T-A := Cn(t - R(A))$$

A *base contraction* operation $\dot{-}$ is an operation on pairs of sets of sentences and sentences such that

$$\dot{-}(t,A) := t - R(A)$$

(As in the case of the sign '$-$' for the contraction operation on theories, I shall denote the value of $\dot{-}$ applied to $\langle t,A\rangle$ as '$t\dot{-}A$'.) As a preliminary to establishing some of the properties of the contraction operation on theories, it will be of interest to prove some facts about base contractions, as just defined.

THEOREM 17.1.

The operation of base contraction, $\dot{-}$, satisfies the following conditions (for any set of sentences t and any sentences A,B).

($\dot{-}$1) If t is irredundant, so is $t\dot{-}A$.

($\dot{-}$2) $t\dot{-}A \subseteq t$.

($\dot{-}$3) $A \notin Cn(t) \Rightarrow t \subseteq t\dot{-}A$.

($\dot{-}$4) $A \notin t\dot{-}A$.

($\dot{-}$5) $Cn(A)=Cn(B) \Rightarrow t\dot{-}A = t\dot{-}B$.

Proof.
Ad ($\dot{-}$2). Obvious.
Ad ($\dot{-}$1). From the definition of 'irredundant set' it follows immediately that every subset of an irredundant set is irredundant. The proposition follows then by ($\dot{-}$2).
Ad ($\dot{-}$3). If $A \notin Cn(t)$, then $R(A)=\varnothing$; hence, $t\dot{-}A=t$.
Ad ($\dot{-}$4). Suppose $A \notin t$. Then $A \notin t\dot{-}A \subseteq t$. Suppose $A \in t$. Then $\{A\}$ is an entailment set for A in t; moreover, A is minimal under $\leq$ in $\{A\}$. So $A \in R(A)$; hence, $A \notin t\dot{-}A$.
Ad ($\dot{-}$5). Suppose $Cn(A)=Cn(B)$. Then $E(A)=E(B)$ and so $R(A)=R(B)$. Hence, $t\dot{-}A=t\dot{-}B$. ■

Base contraction versions of Recovery, like $t \subseteq t\dot{-}A+A$ or even $t \subseteq Cn(t\dot{-}A+A)$, can be shown to fail by means of counterexamples like the one already adduced against Recovery in the last section.

Given these results about base contraction, it is now mostly trivial to show that the contraction operation – on based theories satisfies the postulates (-1) to (-5).

THEOREM 17.2.

For any theory T and sentences A, B, the contraction operation – satisfies the postulates

(-1) $T-A$ is a theory

(-2) $T-A \subseteq T$

(-3) $A \notin T \Rightarrow T \subseteq T-A$

(-4) $A \notin T-A$

(-5) $Cn(A)=Cn(B) \Rightarrow T-A=T-B$

Proof. (-1) holds in virtue of the definition of $T-A$ ($= Cn(t\dot{-}A)$). Given T17.1, (-2), (-3), and (-5) follow immediately from the fact that if $M \subseteq N$ then $Cn(M) \subseteq Cn(N)$.

For (-4) assume for reductio that $A \in T-A$, i.e. $A \in Cn(t—R(A))$. Then there must be some entailment set S for A in $t—R(A) \subseteq t$. Now, S contains some $\leq$-minimal element B. By the definition of $R(A)$ then, $B \in R(A)$. But $S \subseteq t$ and so $B \in t$. Hence, S can not be a subset of $t—R(A)$ – contradiction. ■

In addition to (-1) to (-5) and the recovery postulate (-6), Gärdenfors has proposed two further postulates for contraction (which he calls 'supplementary postulates'):

(-7) $T-A \cap T-B \subseteq T-(A\&B)$

(-8) If $A \notin T-(A\&B)$, then $T-(A\&B) \subseteq T-A$

Postulate (-7) is intuitively impeccable: to contract a theory by a conjunction requires the removal of at least one conjunct. In case both conjuncts can be traced back to equally retractible rejects, we may have to remove both conjuncts. Removing both conjuncts, however, is the most that can reasonably be asked for.

In contrast to (-7), postulate (-8) is certainly less evident than (-7). In fact, it may even appear to be quite daring. In view of our explicit definition of contractions, the following problem about (-8) may arise.

Assume, according to the antecedent of (-8), that each entailment set for A (in a base t generating T) contains some element that is minimal in some entailment set for $A\&B$. Now, the obvious way to show that the consequent of (-8) obtains, is to show that $R(A) \subseteq R(A\&B)$. But why should it be (and why should it even be desirable) that every sentence that is minimal in some entailment set for A should also be minimal in some entailment set for $A\&B$? Could it not be that entailment sets for $A\&B$ either do not contain members of $R(A)$ or that they always contain elements which are strictly less (in the sense of the ordering) than any of those members of $R(A)$ which they do contain? This is indeed the problem one is faced with when attempting to prove that our definition of contractions satisfies (-8). Neither have I been able to solve this problem, nor have I been able to find a counterexample to (-8).

> *Open Problem.* Does the present definition of the contraction operation satisfy the postulate (-8)? If the definition in the present form does not satisfy (-8), is there a non-trivial condition (presumably on the ordering $\leq$) which will ensure that (-8) holds?

Towards a solution of this problem I note two facts. First, we can show that given the antecedent of (-8), $R(A)$ is a subset of all those sentences in t that are minimal in some consequence set for $A\&B$.[9] We would need to show, however, that $R(A)$ is a subset of all those sentences in t that are minimal in some entailment set for $A\&B$, i.e. that $R(A) \subseteq R(A\&B)$. The second fact is perhaps rather discouraging. Alchourron and Makinson (1985) have defined the notion of a *safe contraction* which shares many characteristics with the contraction operation proposed here; the chief difference being that safe contraction proceeds from theories and not from bases (arbitrary sets of sentences). But in showing that safe contraction satisfies (-8), Alchourron and Makinson need the fact that t is a theory (in the present terminology: that theories are generated from superredundant bases). Moreover, their proof makes also essential use of the assumption that t is a *classical* theory.

Given that (-8) is certainly less transparently true than the other postulates, one should perhaps feel not too disconcerted if (-8) failed for the contraction operation as just defined. Gärdenfors' principal motivation

[9] The argument is essentially the same as the one for observation 6.2 in Alchourron and Makinson (1985), pp. 415f.

for (-8) is, that it allows to derive a corresponding revision postulate, (*8), which he finds desirable. However, Gärdenfors derivation of (*8) from (-8) via the Levi identity is riddled with applications of the relevantly unacceptable disjunctive syllogism. And, unless theories are assumed to be closed under disjunctive syllogism, our definition of contractions affords simple counterexamples to (*8). Such a counterexample will be displayed in section 19.

It would be simply awful, however, if the contraction operation failed to satisfy the convincing postulate (-7). Fortunately, this is not the case (T17.7). In order to show that (-7) is validated by the construction offered here, we first prove a few lemmas.

LEMMA 17.3.

If $S \in E(A\&B)$, then there are sets $S_1 \in E(A)$ and $S_2 \in E(B)$ such that $S = S_1 \cup S_2$.

Proof. Assume that $S \in E(A\&B)$. Then $S \in D(A)$ and $S \in D(B)$. So there must be subsets S_1 and S_2 of S such that $S_1 \in E(A)$ and $S_2 \in E(B)$. Clearly, $S_1 \cup S_2 \subseteq S$. It is also clear that $S_1 \cup S_2 \in D(A\&B)$. Now suppose for contradiction that $S_1 \cup S_2 \subset S$. Then $S_1 \cup S_2$ but not S is an entailment set for $A\&B$, i.e. $S \notin E(A\&B)$ – contradicting our hypothesis. So $S_1 \cup S_2 = S$. ■

LEMMA 17.4.

Either $R(A\&B) \subseteq R(A)$ or $R(A\&B) \subseteq R(B)$.

Proof. Assume that $C \in R(A\&B)$. Then there exists some set $S \in E(A\&B)$ with C $\leq$-minimal in S. By L17.3, $S = S_1 \cup S_2$ for some $S_1 \in E(A)$ and $S_2 \in E(B)$. So (a) $C \in S_1$ or (b) $C \in S_2$. Assume that (a) is the case. Then since C is $\leq$-minimal in S, C must be $\leq$-minimal in S_1. Hence, $C \in R(A)$. If on the other hand (b) is the case, then $C \in R(B)$. So either $C \in R(A)$ or $C \in R(B)$, as required. ■

COROLLARY 17.5.

$R(A\&B) \subseteq R(A) \cup R(B)$.

THEOREM 17.6.

The operation of base contraction satisfies the condition

($\dot{-}$7) $(t \dot{-} A) \cap (t \dot{-} B) \subseteq t \dot{-} A\&B$.

Proof. By set theory,

(1) $(t—R(A))\cap(t—R(B))\subseteq t—(R(A)\cup R(B))$.

By C17.5,

(2) $t—(R(A)\cup R(B))\subseteq t—R(A\&B)$.

Hence, from (1) and (2) applying the definition of $\dot{-}$,

(3) $(t\dot{-}A)\cap(t\dot{-}B)\subseteq t\dot{-}A\&B$. ■

THEOREM 17.7.

For any theory T and sentences A,B, the contraction operation $-$ satisfies the postulate

(-7) $(T-A)\cap(T-B)\subseteq T-(A\&B)$.

Proof. Assume $C\in(T-A)\cap(T-B)$, i.e. $C\in Cn(t—R(A))\cap Cn(t—R(B))$. Then there are $A_1,\ldots,A_m\in t—R(A)$ and $B_1,\ldots,B_n\in t—R(B)$ such that $A_1\&\cdots\&A_m\rightarrow C\in \mathbf{L}$ and $B_1\&\cdots\&B_n\rightarrow C\in\mathbf{L}$; i.e.

(1) there are $A_1,\ldots,A_m\in t$ such that
$A_1,\ldots,A_m\notin R(A)$ and
$A_1\&\cdots\&A_m\rightarrow C\in\mathbf{L}$,

and

(2) there are $B_1,\ldots,B_n\in t$ such that
$B_1,\ldots,B_n\notin R(B)$ and
$B_1\&\cdots\&B_n\rightarrow C\in\mathbf{L}$.

Now, by L17.4, either

(i) $R(A\&B)\subseteq R(A)$ or

(ii) $R(A\&B)\subseteq R(B)$.

Suppose (i) is the case. It follows from (1) that $A_1,\ldots,A_m\notin R(A\&B)$. Hence,

(3) there are $C_1,\ldots,C_k\in t$ (namely $A_1,\ldots,A_m$!) such that
$C_1,\ldots,C_k\notin R(A\&B)$ and
$C_1\&\cdots\&C_k\rightarrow C\in\mathbf{L}$.

Thus $C\in Cn(t—R(A\&B))$, i.e. $C\in T-A\&B$, as required. (The case for (ii) is similar, using (2).) ■

17.2 Contraction on theories generated from redundant bases

In proving the formal results of the last section, nowhere have we made use of the assumption that the theories to be contracted are generated from irredundant bases. Thus, were it not for the filtering condition, we could extend the definition of the last section to theories generated from all kinds of bases: irredundant, redundant, or even superredundant. So, how can the filtering condition be violated if we extend the approach of the last section without change to redundant bases?

The answer to this question has already been indicated at the end of section 16. Suppose that A is a redundant element in some base t: A can be deduced from proper subsets t' of t. Then, on the definition of $T-C$ proposed in the last section, A will be in $T-C$ as long as the reject set $R(C)_t$ does not contain A.

This may not always be what we want. Consider for example a database system in which answers to queries are added to the database from which the answers have been deduced. Thus, the database grows with every successful query. (Such a system would have the advantage of speeding up query processes as the database expands.) Clearly, if A is in the database and, as a consequence, $A \vee B$ is in the database, then a query ($A \vee B$?) should not result in an affirmative answer after the database has been contracted by A.

On the other hand, the definition of the last section may give us exactly the right sort of results for certain redundant bases. So-called aesthetic considerations apart, logically independent axiomatisations of theories are important for many purposes. For example, when arguing that a theory "fits" a particular model, it would be wasteful to verify a redundant base for the theory in the model, when an irredundant base is available. And not among the least important reasons for preferring irredundant bases is the fact that they are usually easier to contract. But a logically independent base may not always be a good "representative" of a theory:

- We expect the base of a theory to give us some information as to how the theory treats its key concepts. The key concepts of a logic, for example, are the logical connectives of its underlying language. Axiomatisations which can be split into groups of postulates, each group containing some key information about a particular connective, are thus an important aid for the purposes of comparison and deriving further theorems. However, the demand for informativeness frequently clashes with that for economy.

- For an empirical theory to be successful, it is a necessary condition that it allows to derive a significant range of sentences which square well with the data to be accounted for. If the theory in question is indeed successful, such data will be part of the theory in, as it were, two modes: as brute facts and by inference. But those brute facts which are derivable from high-level generalisations will not be included in an irredundant base for the theory. Intuitively, however, such data should be included in the base; they are, as it were, the very starting point and not *merely* – albeit also – consequences of the theory.

Our strategy has been to reduce the problem of theory-contraction to the problem of set-theoretic subtraction on a distinguished subset of the theory in question (the base of the theory). It now turns out that, in order to obtain adequate results, we need to focus attention not on any odd base for a theory – only those bases should be the proper starting points for a contraction process which are "good representatives" of the theory to be contracted. Being a logically irredundant set of sentences is neither a sufficient nor a necessary condition for a base to be such a representative. I shall suggest now that the contraction process should start from *virtually irredundant bases.*

We distinguish between the sentences of a theory with respect to their *pedigree* (or warrant). Some sentences do not depend for their inclusion in a theory on the presence of certain others. Such sentences are *independently warranted.* By contrast, there are also sentences (in fact, infinitely many) in a theory which do depend on the presence of certain others; their warrant for membership in the theory is *by inference* from ultimately independently warranted sentences. Such sentences must be retracted from a theory whenever the sentences on which they depend are retracted. A *virtually irredundant base* for a theory T contains exactly the set of independently warranted sentences (or *basic sentences*, as we shall say henceforth) in T. A simply irredundant base is thus a special case of a virtually irredundant base: a set of basic sentences which are logically independent.

Now consider again the problem that may arise when extending the approach of the last section beyond (logically) irredundant bases. To take a simple example, let $t=\{A, A \vee B\}$ and suppose we remove A from t. Should $A \vee B$ be in $T-C$ (where the contraction by C requires the removal of A from t)? Our reply will depend on how we answer the following question: Would we have included $A \vee B$ in the base if we had not included A? If AvB is independently warranted, the answer must be Yes. In this case t is virtually irredundant and we should keep AvB in $T-C$. If, on the other hand, AvB is warranted by inference (from A), then t is not virtually irredundant (i.e. t is virtually redundant) and AvB should be removed from $T-C$ together with its warrant A. Thus, the approach of the last section can safely be extended to theories generated from virtually irredundant bases.

The simplest course we may take now is to require, for the purpose of defining a contraction operation on based theories, that theories be generated from virtually irredundant bases. Such a requirement would not completely be *ad hoc*. For, virtually redundant bases are defective in an important sense. When putting forward a set of sentences as the base for a theory, the act of asserting a sentence as "basic" for the theory under consideration carries an implicature to the effect that the sentence is independently warranted; thus, proposing a base for a theory usually implies that the base thus put forward is virtually irredundant.

The simplest course, however, is better avoided. First, there is no reason why we should not – if in fact we can – extend our theory of contraction to theories generated from virtually redundant bases. Secondly, we still want to apply our construction to theories "generated from" superredundant bases. But there is certainly something very odd in saying that superredundant bases are virtually irredundant.

The construction of the last section proceeded by narrowing down the set of candidates for rejection in a theory: from the whole theory to its base, and from the base to the union of all entailment sets in the base. We shall insert now a further step. Let t be a base generating the theory T. Instead of requiring outright that a base t be virtually irredundant, we shall henceforth assume that a virtually irredundant base can be uniquely reconstructed from t. (In the case of databases, we may assume that independently warranted sentences in a base are somehow marked.) The set of all independently warranted members of t will be denoted by 't^*'. Thus, each base t has associated with it a unique virtually irredundant (sub)base t^*; if t is virtually irredundant, then, of course, $t^*=t$. Now the process of contraction proceeds as before, with t^* in place of t. The amended procedure is summarised in the diagram on the next page.

Thus, the contraction by A of a theory T generated from the base t is now redefined as the set of consequences of the virtually irredundant subbase t^* of t without the set of A-rejects in t^*, i.e.

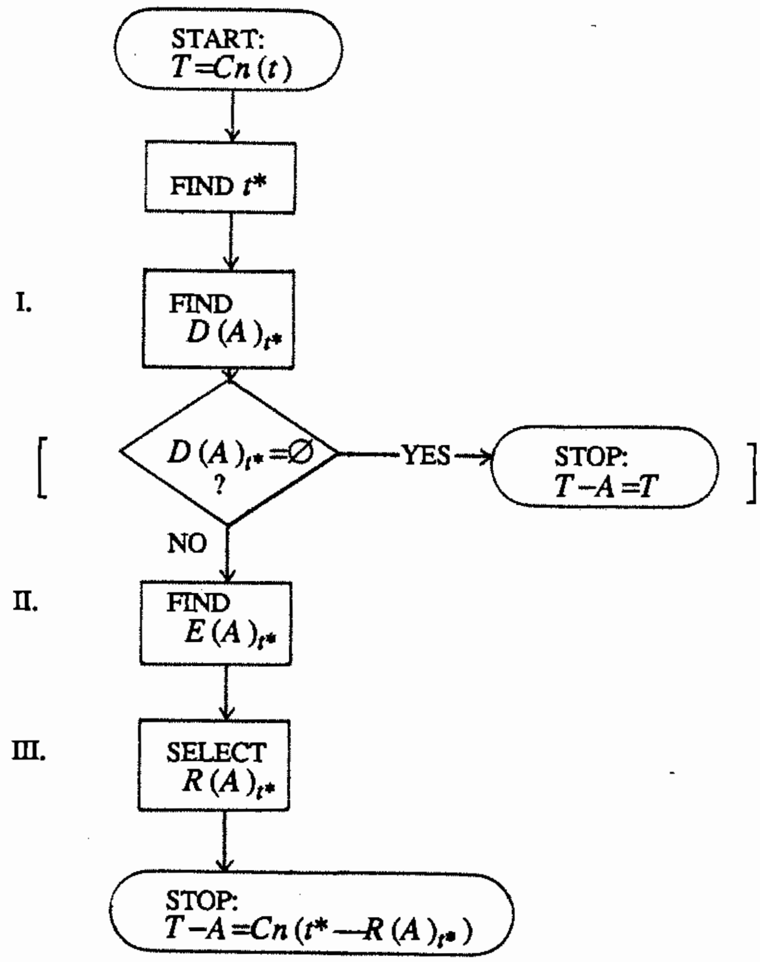


FIGURE 17.8
Given T on t, find $T{-}A$.

$$T{-}A := Cn(t^{*}{-\!\!-}R(A)_{t^{*}}), \quad \text{for all } T, t \text{ such that } T{=}Cn(t).$$

As noted earlier, the formal results in section 17 were quite general: they did not depend on any assumptions about the character of the bases t relative to which consequence sets, entailment sets, and rejection sets were defined. Thus, in the proofs of these results t^* may replace t *salva veritate.*

If T is generated from an *infinite base* t, $T{-}A$ is still well defined (according to the definition of the last paragraph) and the contraction operation satisfies the Gärdenfors postulates, except for Recovery. But unless t can be reduced to a finite virtually irredundant base t^* for t, the proposed algorithm ceases to be effective. The effectiveness of the algorithm breaks down already at the first subroutine: since t^* is infinite, 2^{t^*} contains infinitely many A-consequence sets. Even if we skip the process [FIND $D(A)$!], 2^{t^*} may contain infinitely many A-entailment sets; and if t^* is superredundant, then there will be *in general* infinitely many A-entailment sets in t^*. (For, let t^* be a superredundant base containing A. Then t^* is closed under finite conjunction. Since t^* is infinite, there are infinitely many conjunctions in t^* with A as a conjunct. For each such conjunction $\&(A)_i$ $(0 \leq i \leq \omega)$, the singleton set $\{\&(A)_i\}$ is an A-entailment set in t^*.)

18. Multiple contraction

We shall now generalise our theory of contraction to cover the case of contracting a theory by a set of sentences rather than a single sentence.

Writing tentatively '$T-S$' for the contraction of the theory T by the set of sentences S, we note at once that we may have two rather different expectations about $T-S$. We may either require that S must not be a subset of $T-S$, or that no sentence in S should be a member of $T-S$. In the former case it suffices that at least one member of S be retracted from T; in the latter case we need to retract all members of S from T. Thus, we need to disambiguate the expression $T-S$ as follows.

$T-[A_1 \cdots A_n]$: the contraction of T by all A_i;
$T-\langle A_1 \cdots A_n \rangle$: the contraction of T by some A_i;
$(1 \leq i \leq n)$.

Multiple contractions of the first kind will be called *(multiple) meet contractions*; multiple contractions of the second kind will be referred to as *(multiple) choice contractions*. Obviously, when S is a singleton set $\{A\}$, the meet contraction $T-[S]$ coincides with the choice contraction $T-\langle S \rangle$: they reduce to the simple contraction $T-A$ treated in the preceding section. It should also be obvious that choice-contracted theories are always subsets of their corresponding meet-contracted theories.

Both meet and choice contractions are not only of theoretical interest as generalisations – into two directions – of simple contractions. They are also of practical importance.

We frequently need to retract sentences from a theory *simultaneously*. Now, the simple contraction operation on single sentences is not commutative: $(T-A)-B$ is not in general the same as $(T-B)-A$. Thus, the simultaneous retraction of a collection of sentences from a theory is not definable as a sequence of simple contractions, since the order in which sentences are retracted may make a difference. It would seem rather odd that when having to retract a number of sentences, we should first sit down and contemplate in which order the sentences should be retracted. We want to retract all sentences in one swoop – and this is what meet contraction is for.

Choice contractions, on the other hand, are made in response to an invitation of the form: 'You have to give up at least one of A_1 or ... or A_n – take your pick!' A choice contraction by a set of sentences is the right operation to perform on a theory, whenever the sentences in the set do not "fit together", are incoherent. 'Incoherent' may not just mean inconsistent, but that is at least one important way in which a set of sentences may be incoherent. Thus, suppose we have derived two sentences $A, \sim A$ from a base t and want to restore the consistency of T with respect to A. Then we shall have to give up either A or $\sim A$ – that is to say, we shall have to choice-contract T by the set $\{A, \sim A\}$. Consistency (and, more generally, coherence) is an important goal in theorising; hence, it is important to know how these properties can be restored.

As in the preceding section, I shall first offer explicit definitions of (and recipes for constructing) meet and choice contractions. It will then have to be seen how the Gärdenfors postulates can be generalised to yield interesting properties of these two multiple contraction operations.

18.1. The definition of meet contraction

As in the previous section, we are considering theories T generated from bases t. A function $*$ is defined on each base t which maps t into its virtually redundant subbase t^*. t^* is ordered by means of a comparative retractibility relation $\leq$. The definitions of 'consequence set', 'entailment set', and 'reject set' are as before.

The task of finding a contracted theory $T-[A_1 \cdots A_n]$ – where T is generated from t – can be split up into six subroutines:

1. Find the virtually irredundant subbase t^* of t.
2. For each A_i $(1 \leq i \leq n)$, find all A_i-consequence sets in t^*.
3. For each A_i, cut $D(A_i)$ down to the set $E(A_i)$ of all A_i-entailment sets in t^*.
4. For each A_i, collect all sentences which are minimal under $\leq$ in some A_i-entailment set in the reject set $R(A_i)$.
5. Let $R([A_1 \cdots A_n]) = \bigcup_{i=1}^{n} R(A_i)$.
6. Define $T-[A_1 \cdots A_n] := Cn(t - R([A_1 \cdots A_n]))$.

Again, this procedure lends itself to the design of a meet contraction program as set out in the diagram on the next page.

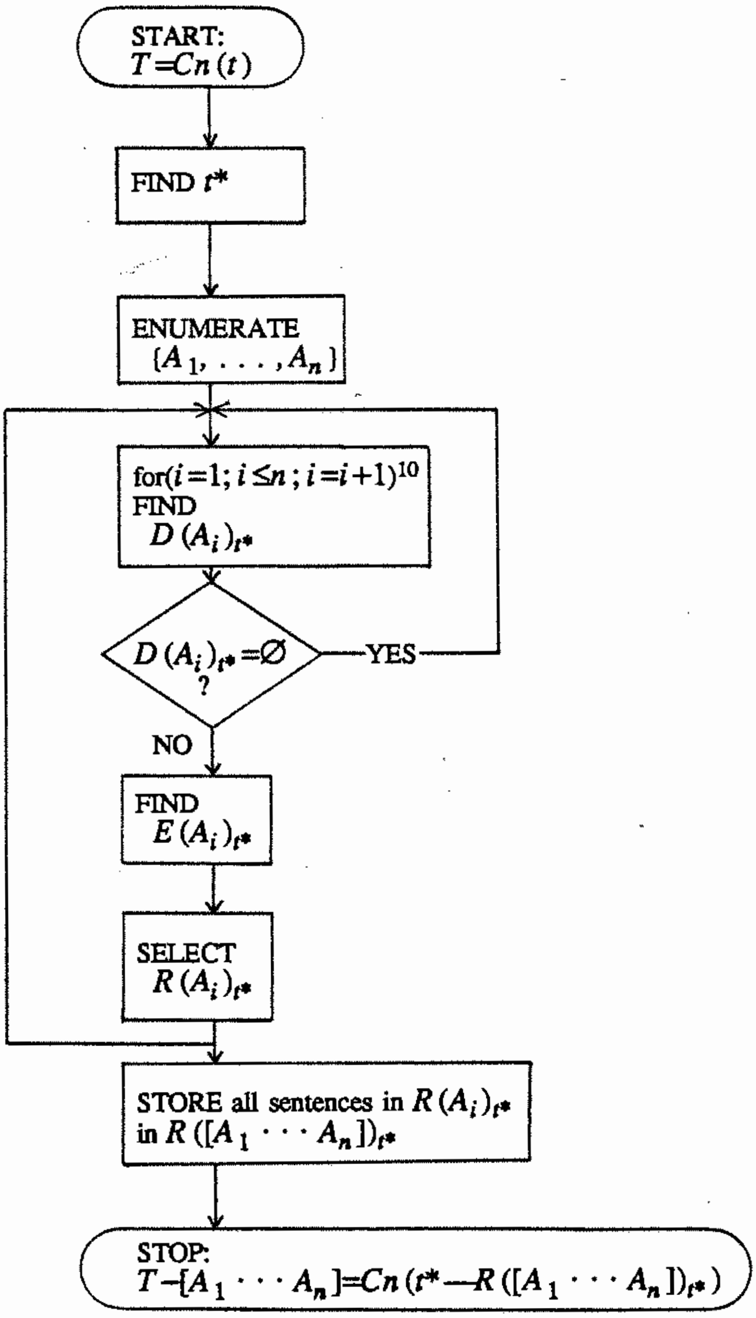


FIGURE 18.1
Given T on t, find $T-[A_1 \cdots A_n]$.

[10] The for-statement sets the initial value of i to 1, increments the value of i by 1 every time a loop has been carried out, and stops the program when i has reached the value n.

18.2. The definition of choice contraction

It has been emphasised earlier that when defining change operations on theories gratuitous loss of information should be avoided. Thus, when facing a choice as to which one(s) out of a set of sentences to retract, one should select the one(s) that keep the loss of information contained in the original theory at a minimum. Though, strictly speaking, the purpose of a choice contraction by S will be fulfilled if one sentence only is retracted from the theory under consideration, we shall sometimes have to retract more than one sentence out of S. This will happen whenever a contraction by two or more sentences incurs an equally minimal loss of information (remember Bizet and Verdi!). On the other hand, sometimes a choice contraction may require no incision into a theory at all. When S contains a sentence A which is not in T, then it is already the case that S is not contained in T and so the choice contraction of T by S ought to be vacuous: $T-\langle S\rangle=T$.

So far we have only considered an ordering of comparative retractibility on the members of the base of a theory. We now need some way of assessing for *any* two sentences of our language whether the one is comparatively more retractible from a given theory T than the other – whether retracting the one sentence does less damage to T than retracting the other. That is, for each theory T on t, we need some way of ordering all sentences of the language with respect to their comparative retractibility from T (on t). The key to defining such an ordering $\leq^t$ on all sentences is the fact that each sentence A can be traced back to a reject set $R(A)_t$ (for any base t generating a theory T). If A is not a member of $Cn(t)$, then $R(A)_t$ will be empty; otherwise $R(A)_t$ will be structured by the ordering $\leq$ on t. Suppose we want to find the set $T-\langle A,B\rangle$, where T is generated from t. For each of A,B those sets of sentences we should delete from the base t in order to avoid commitment to A and B respectively are uniquely determined: these are the reject sets in t for A and B respectively, $R(A)$ and $R(B)$. So the problem of choosing between contracting T by B or by A comes down to a choice between subtracting $R(A)$ or $R(B)$ from the base. This choice should be determined by how the retractibility of sentences in the one reject set compares with that in the other. Clearly, if, say, $R(A)$ is empty, then A is easier to retract than B. So in this case it ought to be that $A\leq^t B$. For the non-vacuous case we need some way of "uplifting" the ordering $\leq$ on t to an ordering $\leq^2$ between subsets of t so as to compare reject sets with respect to their comparative retractibility. Given the right kind of ordering $\leq^2$ between reject sets (i.e. one that satisfies *inter alia* the requirement that the empty set is always minimal under $\leq^2$), we may then define $A \leq^t B$ to hold just in case $R(A)_t \leq^2 R(B)_t$, for any sentences A, B.

There are various ways in which an ordering on a set may be used to define an ordering on its powerset. One for the present purpose particularly promising way of doing this, is to let $\leq^2$ be the *power ordering* of $\leq$.[11]

DEFINITION 18.2. Power ordering

If S is a set ordered by $\leq$, then the power ordering $\leqq$ induced by $\leq$ is that relation between subsets of S such that for any subsets X, Y of S

$$X \leqq Y \text{ iff } (\forall x \in X)(\exists y \in Y)(x \leq y) \text{ and } (\forall y \in Y)(\exists x \in X)(x \leq y).$$

So we tentatively define:

$$(*) \qquad X \leq^2 Y \text{ iff } X \leqq Y.$$

The definition (*) requires that for a set X to be at least as retractible as a set Y, two conditions must be satisfied: first, each sentence in X is "below or at the same level with" some sentence in Y, and secondly, each sentence in Y is "above or at the same level with" some sentence in X.

By and large, this definition of $\leq^2$ gives the right results. Consider, for example, two reject sets, one for A and one for B. Let $R(A)=\{A, A'\}$ and $R(B)=\{B, B'\}$. The following figure shows Hasse diagrams of four ways of linearly pre-ordering the set $\{A, A', B, B'\}$.

[11] The concept of a power ordering is due to Brink. Brink (1986) contains some general observations about power orderings. In Brink (1987) the concept of power ordering is applied to improve upon Popper's definition of comparative verisimilitude among theories.

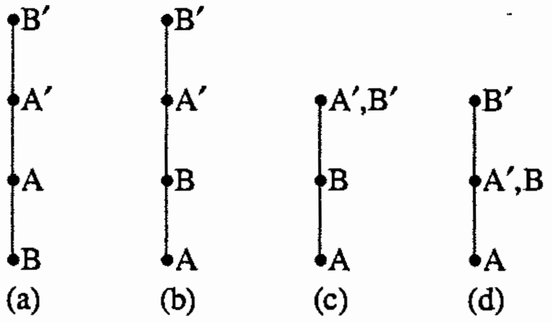


FIGURE 18.3

A moment of reflection shows that in cases (b) to (d) it clearly ought to be that $R(A) \leq^2 R(B)$ while $R(B) \not\leq^2 R(A)$, and this is indeed what we get, when $\leq^2$ is defined as in (*). In case (a), $R(A)$ and $R(B)$ are incomparable with respect to $\leq^2$. On the one hand, there is no element in $R(A)$ which is less or equal to B; thus $R(A) \not\leq^2 R(B)$. On the other hand, B' is maximal in the chain; thus the first condition for $R(B) \leq^2 R(A)$ fails. This result may raise an eyebrow: 'Suppose B' is *much more* important than A'. Would it not be better then to dispose of A and A' rather than removing B' (and B)?' The question points towards a limitation of the present approach. In the formal framework offered in this chapter we simply lack the means of expressing that some sentence is "much more" important ("much less" retractible) than another. But even if we exchanged the ordering $\leq$ for, say, a real-valued function i assigning degrees of immunity to retraction between 0 and 1 to sentences, it would still not be clear whether case (b) should be resolved in favour of $R(B)$ if B' is much less retractible than A. Consider for example the following assignments of immunity to retraction: $i(B)=0.02$, $i(A)=0.4$, $i(A')=0.6$, and $i(B')=0.97$. There is a good sense in which B' is much less retractible than A'. Unfortunately, however, B' has a highly retractible companion, B. If the retractibility of a set is determined by averaging the retractibility of its members, as it seems natural to do, then $R(B)$ will in fact be (slightly) more retractible than $R(A)$. On the other hand, if we slightly decrease the immunity value of either A or A' or slightly increase the value of $i(B)$, then $R(A)$ rather than $R(B)$ will have to go. Thus, even where a quantitative ordering preserves a qualitative ordering of comparative retractibility on sentences, their extrapolations to an ordering on sets of sentences may lead to conflicting results. Since a quantitative ordering is more discerning, it would be bad policy to resolve the conflict by applying the definition (*) rather than the quantitative averaging method. However, on the assumption that only an admittedly crude qualitative ordering is available, we should let the definition (*) – or some definition like it – take its course and treat the results with the requisite caution. To reject these results by introducing quantitative orderings, would amount to a change of topic.

Note that where $A, B \in t$, we can not generally infer $R(B) \leq^2 R(A)$ from $B \leq A$. Case (a) illustrates this fact: although $B \leq A$, $R(B) \nleq^2 R(A)$. Thus the ordering $\leq^t$ on all sentences does not in general *extend* the ordering $\leq$ on sentences in the base t. However, $\leq$ is *embedded* in $\leq^2$: if $A \leq B$ then $\{A\} \leq^2 \{B\}$. (As a consequence, $\leq^t$ does extend the ordering $\leq$ on sentences in t, if t is irredundant. For, suppose that t is irredundant, that $A, B \in t$, and that $A \leq B$. Then $R(A) = \{A\}$ and $R(B) = \{B\}$, since t is irredundant. Hence, $R(A) \leq^2 R(B)$ and so $A \leq^t B$.)

Despite these attractive features, the definition (*) has a serious drawback. When having a choice between retracting A or retracting B, and $R(A)$ is a subset of $R(B)$, then $R(A)$ should obviously be at least as retractible as $R(B)$ – so, we want the following to hold:

(†) if $R(A) \subseteq R(B)$, then $R(A) \leq^2 R(B)$.

But, given (*), (†) can be shown to fail in a most unpleasant way. Consider the reject sets $R(A) = \{A\}$ and $R(B) = \{A, B\}$ and suppose that $B \leq A$. Since removing $R(A)$ does less damage to the theory than removing $R(B)$, we should expect $R(A)$ to be more retractible than $R(B)$. But applying (*), it is quickly verified that $R(A) \nleq^2 R(B)$ and – what is worse! – that $R(B) \leq^2 R(A)$. So, (*) can force us to make a larger incision into a theory than necessary.

In one respect, such a result is not at all surprising. The power ordering = on a power base 2^t is defined solely in terms of an ordering $\leq$ on t. By contrast, the relation of set inclusion among sets in 2^t pays no attention at all to the ordering $\leq$. Thus, $\nLeftarrow$ and $\subseteq$ are completely independent. If we want some dependency between $\subseteq$ and $\leq^2$ to obtain, we have to say so. Thus, the improved definition of $\leq^2$ is as follows.

DEFINITION 18.4.

If t is a base ordered by $\leq$, then the comparative retractibility relation on $2^t \times 2^t$ is the smallest set $\leq^2$ such that

(a) if $S_1 \subseteq S_2$, then $<S_1, S_2> \in \leq^2$;

(b) if $S_1 \not\subseteq S_2$, then $<S_1, S_2> \in \leq^2$ iff $S_1 \Leftarrow S_2$.

This new definition gives us the same verdicts as the tentative definition (*) on our earlier example with variations on the ordering $\leq$ as displayed in figure 18.3. But in contrast to (*), D18.4 satisfies the condition (†). An ordering $\leq^t$ of comparative retractibility from t on all sentences of the language may now be defined thus:

DEFINITION 18.5.

For every base t ordered by $\leq$ and sentences A, B:

$$A \leq^t B \text{ iff } R(A)_t \leq^2 R(B)_t .$$

As an immediate consequence of of D18.4.(a) and D18.5 it follows that

THEOREM 18.6.

for every base t and sentences A, B: if $A \notin Cn(t)$, then $A \leq^t B$.

Intuitively, the theorem states that if a sentence A is not a member of the theory T generated from t, then A is minimal under $\leq^t$, i.e. A is not less retractible from T than any other sentence of the language.

We have gathered now all prerequisites for defining choice contractions on theories generated from ordered bases. As usual, the definition will be stated in a step-by-step fashion and summarised in a flowchart diagram below. Suppose $T=Cn(t)$, where t is ordered by $\leq$. The choice contraction $T-<A_1 \cdots A_n>$ is constructed in six steps (steps 1 to 4 are as before).

1. Find the virtually irredundant subbase t^* of t.
2. For each A_i $(1 \leq i \leq n)$, find all A_i-consequence sets in t^*.
3. For each A_i, cut down $D(A_i)$ to the set $E(A_i)$ of all A_i-entailment sets in t^*.
4. For each A_i, collect in the reject set $R(A_i)$ all sentences which are minimal under $\leq$ in some A_i-entailment set.
5. $R(<A_1 \cdots A_n>) = \bigcup_{i=1}^{n} [R(A_i) \colon A_i \text{ is } \leq^t\text{-minimal in } \{A_1, \ldots, A_n\}]$.

6. Define $T-\langle A_1 \cdots A_n \rangle := Cn(t - R(\langle A_1 \cdots A_n \rangle))$.

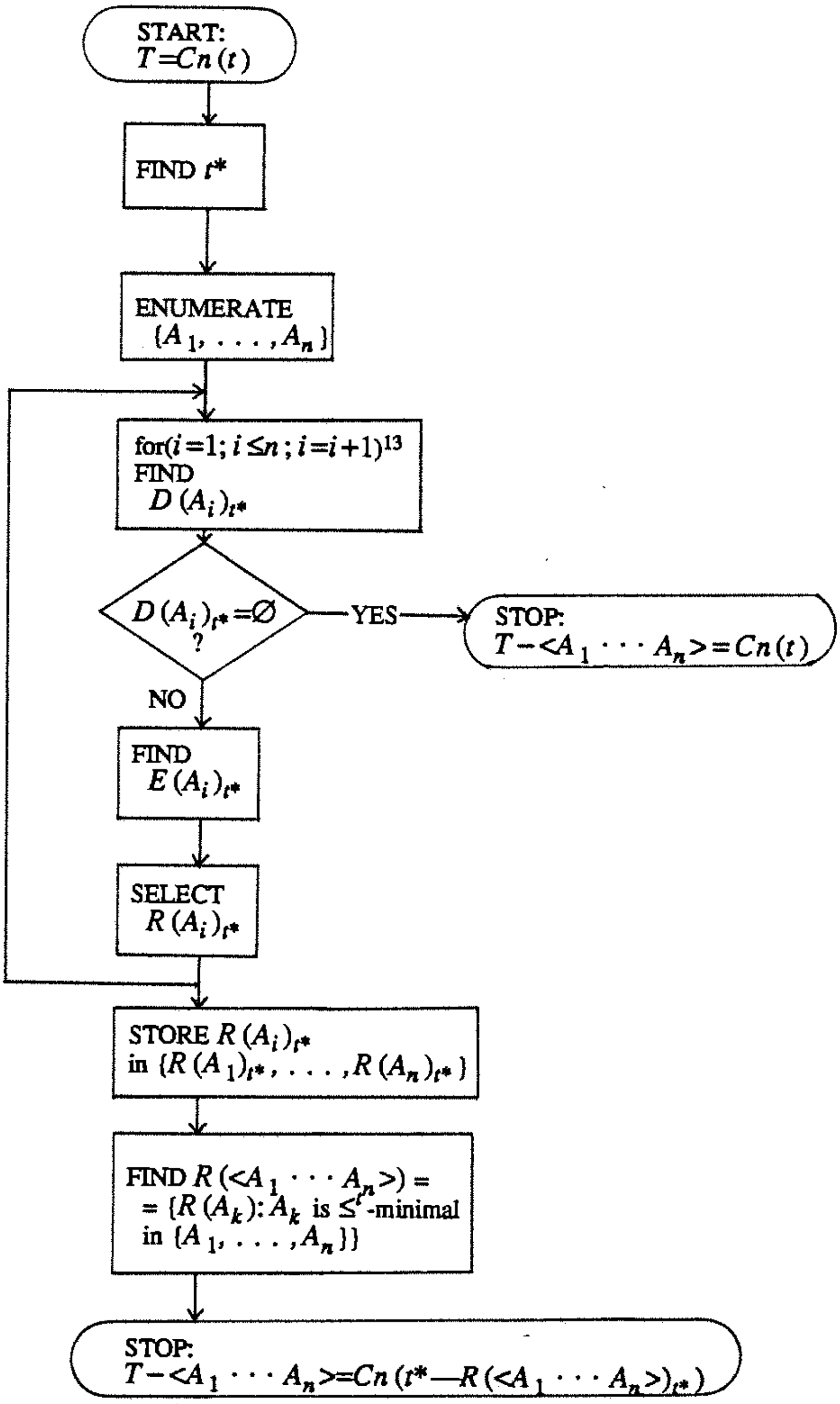


FIGURE 18.7
Given T on t, find $T-\langle A_1 \cdots A_n \rangle$.

18.3. The generalised Gärdenfors postulates for multiple contractions

We shall now match the explicit definitions of multiple contractions with a set of postulates. Just as the definitions of meet and choice contractions can be viewed as generalisations of the definition of the single contraction operation, so postulates for multiple contractions should result by appropriately generalising the Gärdenfors postulates.

The obvious generalisations of the postulates (-1) and (-2) are

[-1] $T-[S]$ is a theory

<-1> $T-\langle S\rangle$ is a theory

and

[-2] $T-[S]\subseteq T$

<-2> $T-\langle S\rangle\subseteq T$

respectively.

The postulate (-3) tells us under what condition a contraction is vacuous. For multiple contractions such conditions are as easily found as for single contractions. In the case of meet contraction we need to remove all elements of S from T. So, the contraction will be vacuous, if no member of S is in T, i.e.

[-3] $T\cap S=\varnothing \Rightarrow T\subseteq T-[S]$.

A choice contraction of T by S retracts from T all those and only those members of S which can be removed at minimal cost. If S contains sentences that can be removed from T at *no* cost, then only those sentences should be "removed"; thus,

<-3> $S\not\subseteq T \Rightarrow T\subseteq T-\langle S\rangle$.

The postulates [-4] and <-4> state that multiple contractions must be successful:

[-4] $S\cap(T-[S])=\varnothing$

<-4> $S \not\subseteq T - \langle S \rangle$.

Some care needs to be exercised in generalising (-5) to principles about multiple contractions. What appear to be the most straightforward generalisations of (-5) fail, namely,

*[-5] $Cn(S)=Cn(S') \Rightarrow T-[S]=T-[S']$ and

*<-5> $Cn(S)=Cn(S') \Rightarrow T-\langle S\rangle=T-\langle S'\rangle$.

(For a counterexample to the somewhat more plausible *[-5], let $t=\{A, AvB\}$, $S=\{A\}$, $S'=\{A, AvB\}$. Then $E(A)=\{\{A\}\}$ and $E(AvB)=\{\{A\},\{AvB\}\}$. Hence, $R(A)=\{A\}$ and $R(AvB)=\{A, AvB\}$. Clearly, $Cn(S)=Cn(S')$. But $T-[S]=Cn(t—R(A))\doteq Cn(AvB)$ is a proper superset of $T-[S']=Cn(t—(R(A)\cup R(AvB)))=Cn(\varnothing)$.)

The logical equivalence of S and S' does not suffice to guarantee the equivalence of multiple contractions by S and by S'. The condition we need is that S and S' must be *pairwise equivalent*: for each sentence in S there is a logically equivalent sentence in S' and *vice versa.*

DEFINITION 18.8. Pairwise equivalence

Two sets of sentences S, S' are pairwise equivalent (*modulo* the consequence operation Cn) just in case

$$(\forall A\in S)(\exists A'\in S')(Cn(A)=Cn(A')) \text{ and}$$
$$(\forall A'\in S')(\exists A\in S)(Cn(A')=Cn(A)) .$$

Then the correct generalisations of (-5) to multiple contractions are

[-5] if S and S' are pairwise equivalent, then $T-[S]=T-[S']$;

<-5> if S and S' are pairwise equivalent, then $T-\langle S\rangle=T-\langle S'\rangle$.

Gärdenfors supplementary postulates, (-7) and (-8), are perhaps best be motivated as conditions on *multiple* contractions.[14] For, there is an obvious correspondence between retracting a disjunction and meet contraction on the one hand and retracting a conjunction and choice

[14] This is indeed how Gärdenfors motivates the "corresponding" revision postulates (*7) and (*8) which will be discussed in the next section: "My aim is to formulate a generalisation of (*3) and (*4) that apply to *iterated* [simultaneous?] changes of belief." (1988), p.61.

contraction on the other: to contract by a disjunction, all disjuncts have to be retracted, and to contract by a conjunction, some conjunct has to be retracted. Thus, (-7) translates into a postulate for choice contraction,

<-7> $$T-<S_1>\cap\cdots\cap T-<S_n>\subseteq T-<S_1\cup\cdots\cup S_n>\ ,$$

while the corresponding postulate for meet contraction can be strengthened to an identity (as we shall presently show):

[-7.1] $$T-[S_1\cup\cdots\cup S_n]\subseteq T-[S_1]\cap\cdots\cap T-[S_n]\ .$$

[-7.2] $$T-[S_1]\cap\cdots\cap T-[S_n]\subseteq T-[S_1\cup\cdots\cup S_n]\ ,$$

The analogy between <-7> and (-7) is brought out more sharply by the following instance of <-7>.

$$(T-A)\cap(T-B)\subseteq T-<A,B>\ .$$

The disjunctive version of (-7), viz.

(-7v) $$T-(A\vee B)\subseteq(T-A)\cap(T-B)$$

corresponds dually to [-7.1]:

$$T-[A,B]\subseteq(T-A)\cap(T-B)\ .$$

Similarly, we shall now interpret the postulate

(-8) $$A\notin T-(A\&B)\Rightarrow T-(A\&B)\subseteq T-A$$

and its dual

(-8v) $$A\vee B\notin T-A\ \Rightarrow T-A\subseteq T-(A\vee B)$$

as concerned with multiple contractions. Thus understood, these postulates become now *partial vacuity* principles, stating under which condition the multiple contraction by a superset $S\cup S'$ of S does not result in a smaller set than contracting by S only (adding on sentences to S is, in this sense, "vacuous").

<-8> $$S'\not\subseteq T\ \Rightarrow T-<S>\subseteq T-<S\cup S'>\ .$$

[-8] $$S'\cap T=\varnothing\ \Rightarrow T-[S]\subseteq T-[S\cup S']\ ,$$

These partial vacuity postulates are in fact generalisations of the vacuity postulates <-3> and [-3], as shown by the next theorem.

THEOREM 18.9.

(a) <-8> entails <-3>;

(b) [-8] entails [-3].

Proof. For (a), <-3> results by letting S be the empty set in <-8>. For (b), [-3] results by letting $S=S'$ in [-8]. ■

Moreover, in the presence of [-2] and [-7], [-3] and [-8] are equivalent, and so are <-3> and <-8>, given <-2> and <-7> (T18.11). Hence, the multiple -8 postulates are redundant.

LEMMA 18.10.

(a) <-2>, <-3> and <-7> entail <-8>;

(b) [-2], [-3] and [-7.2] entail [-8].

Proof. Ad (a). Assume

(1) $S' \not\subseteq T$.

Then, by <-3>,

(2) $T \subseteq T - \langle S' \rangle$.

From <-2> we have

(3) $T - \langle S \rangle \subseteq T \cap T - \langle S \rangle$,

and by (2),

(4) $T \cap T - \langle S \rangle \subseteq T - \langle S' \rangle \cap T - \langle S \rangle$.

Now we apply <-7>,

(5) $T - \langle S' \rangle \cap T - \langle S \rangle \subseteq T - \langle S \cup S' \rangle$,

to obtain from (3), (4), and (5)

(6) $T - \langle S \rangle \subseteq T - \langle S \cup S' \rangle$,

as required.

Ad (b). Assume

(1) $S' \cap T = \varnothing$.

It follows by [-3] that

(2) $T \subseteq T - [S']$.

By [-2],

(3) $T - [S] \subseteq T$,

whence,

(4) $T-[S]\subseteq T-[S]\cap T$.

Hence, from (2) we can infer

(5) $T-[S]\subseteq T-[S]\cap T-[S']$.

Applying [7.2], we have

(6) $T-[S]\cap T-[S']\subseteq T-[S\cup S']$.

Hence, from (5) and (6),

(7) $T-[S]\subseteq T-[S\cup S']$,

as required. ■

THEOREM 18.11.

(a) If <-2> and <-7>, then <-3> iff <-8>;

(b) if [-2] and [-7.2], then [-3] iff [-8].

Proof. From T18.9 and L18.10. ■

A choice contraction by $\{A,B\}$ must be either a contraction by A, or by B, or by both A and B:

<-V>

$$T - < A,B > = T-[A,B] \text{ or }$$
$$T - < A,B > = T-A \text{ or }$$
$$T - < A,B > = T-B \ .$$

More generally, each choice contraction by a set S must be identical to a meet contraction by some subset of S. So we shall add this requirement as an extra postulate to the generalised Gärdenfors postulates.

{-9} $\quad T - < S > = T-[S']$, for some $S'\subseteq S$.

THEOREM 18.12.

(a) The operation of meet contraction, as defined above, satisfies the postulates [-1] to [-7].

(b) The operation of choice contraction, as defined above, satisfies the postulates <-1> to <-7>.

(c) Meet and choice contractions satisfy the postulate {-9}.

Proof.

Ad (a). [-1] and [-2] are trivial.

For [-3] suppose $T\cap S=\varnothing$. Then $R([S])=\varnothing$. Hence, $T-[S]=Cn(t—R([S]))=Cn(t)=T$.

For [-4] assume for reductio that $S\cap(T-[S])\neq\varnothing$. Then for some A_i, $A_i\in S$ and $A_i\in Cn(t—R([S]))$. Let $S=\{A_0,\ldots,A_n\}$. Then $R([S])=R(A_0)\cup\cdots\cup R(S_n)$ and so $R(A_i)\subseteq R([S])$. Hence, $t—R([S])\subseteq t—R(A_i)$ and consequently $Cn(t—R([S]))\subseteq Cn(t—R(A_i))$. But since $A_i\in Cn(t—R([S]))=T-A_i$, $A_i\in Cn(t—R(A_i))=T-A_i$, contradicting (-4).

For [-5] assume that S and S' are pairwise equivalent. Where $S=\{A_1,\ldots,A_n\}$ and $S'=\{A'_1,\ldots,A'_n\}$ it will suffice to show that

(*) $R(A_1)\cup\cdots\cup R(A_n)=R(A'_1)\cup\cdots\cup R(A'_n)$.

By the definition of R, if $A\leftrightarrow A'\in \mathrm{L}$, then $R(A)=R(A')$. By the pairwise equivalence of S and S', there is for each $A\in S$ an $A'\in S'$ such that $R(A)=R(A')$ and *vice versa.* Thus the equation (*) holds in virtue of the fact that each reject set on the left-hand-side can be paired off with an extensionally equivalent reject set on the right-hand-side and *vice versa.*

For [-7] it will suffice to show that

(5) $T-[A_1\cdots A_n]=T-A_1\cap\cdots\cap T-A_n$.

By the definition of R for meet contractions, we have

(1) $R([A_1\cdots A_n])=R(A_1)\cup\cdots\cup R(A_n)$,

from which it follows by set theory that

(2) $t—R([A_1\cdots A_n])=t—(R(A_1)\cup\cdots\cup R(A_n))$.

Also by set theory we have

(3) $t—(R(A_1)\cup\cdots\cup R(A_n))=t—R(A_1)\cap\cdots\cap t—R(A_n)$.

Hence, from (2) and (3),

(4) $t—R([A_1\cdots A_n])=t—R(A_1)\cap\cdots\cap t—R(A_n)$,

which gives us (5) by the monotonicity of Cn and the definition of a meet-contracted theory.

Ad (b). <-1> and <-2> are trivial.

For <-3> suppose that $S\not\subseteq T$. Then $R(<S>)=\varnothing$. Hence, $T-<S>=Cn(t—R(<S>))=Cn(t)=T$.

For <-4> assume for reductio that $S\subseteq T-<S>$, i.e. that for all $A\in S$, $A\in Cn(t——R(<S>))$. Let $R(A_i)$ be $\leq^2$-minimal in $\{R(A):A\in S\}$. Then $Cn(t—R(<S>))\subseteq Cn(t—R(A_i))$. Since $A_i\in Cn(t—R(<S>))$, $A_i\in Cn(t—R(A_i))$, i.e. $A_i\in T-A_i$ which contradicts (-4).

Given the earlier proof of [-5], it suffices for <-5> to observe that where $A \leftrightarrow A' \in \mathbf{L}$ and $B \leftrightarrow B' \in \mathbf{L}$, $A \leq^t B$ iff $A' \leq^t B'$.

To prove <-7>, let $A_i, \ldots, A_k$ be the $\leq^t$-minimal elements in $S = S_1 \cup \cdots \cup S_n$. Each of $A_i, \ldots, A_k$ must be $\leq^t$-minimal in some of the sets $S_1, \ldots, S_n$. So by the definition of $R(\langle S \rangle)$,

(*) $R(\langle S \rangle) \subseteq R(\langle S_1 \rangle) \cup \cdots \cup R(\langle S_n \rangle)$.

Assume now for some A_0 that $A_0 \in T - \langle S_1 \rangle \cap \cdots \cap T - \langle S_n \rangle$.[15] Then

(1) there are $A_1, \ldots, A_j \in t$ such that
$A_1, \ldots, A_j \notin R(\langle S_1 \rangle)$ and
$A_1 \& \cdots \& A_j \rightarrow A_0 \in \mathbf{L}$; and

(2) there are $B_1, \ldots, B_k \in t$ such that
$B_1, \ldots, B_k \notin R(\langle S_2 \rangle)$ and
$B_1 \& \cdots \& B_k \rightarrow A_0 \in \mathbf{L}$; and

. . .

(n) there are $D_1, \ldots, D_m \in t$ such that
$D_1, \ldots, D_m \notin R(\langle S_n \rangle)$ and
$D_1 \& \cdots \& D_m \rightarrow A_0 \in \mathbf{L}$.

By (*), either $R(\langle S \rangle) \subseteq R(\langle S_1 \rangle)$ or $\cdots$ or $R(\langle S \rangle) \subseteq R(\langle S_n \rangle)$. Pick an arbitrary disjunct, i.e. suppose for some S_i that $R(\langle S \rangle) \subseteq R(\langle S_i \rangle)$. By the assumption

(i), there are $C_1, \ldots, C_l \in t$ such that
$C_1, \ldots, C_l \notin R(\langle S_i \rangle)$ and
$C_1 \& \cdots \& C_l \rightarrow A_0 \in \mathbf{L}$.

Since by hypothesis $R(\langle S \rangle) \subseteq R(\langle S_i \rangle)$ it follows that $C_1, \ldots, C_l \notin R(\langle S \rangle)$. Hence, $A_0 \in Cn(t - R(\langle S \rangle)) = T - \langle S \rangle$ as required.

Ad (c). Let $S = \{A_1, \ldots, A_n\}$. Then $T - \langle S \rangle = Cn(t - (R(A_i) \cup \cdots \cup R(A_k)))$ where $A_i, \ldots, A_k$ are exactly the $\leq^t$-minimal sentences in S. Thus, by the definitions of R for meet and choice contractions, $T - \langle S \rangle = T - [A_i \cdots A_k]$ and $\{A_i \cdots A_k\} \subseteq S$, as required. ■

It follows immediately from {-9} that for singleton sets $\{A\}$, $T - \langle A \rangle = T - [A]$. For the contraction of T by a single sentence A, we shall go on to write $T - A$ on the understanding that $T - A$ is defined as either $T - [A]$ or $T - \langle A \rangle$; in view of the just mentioned consequence of

[15] Compare the proof of T17.7.

{-9}, it does not matter which definition we choose. The next theorem states that the Gärdenfors postulates for single contractions are special cases of the postulates for multiple contractions.

THEOREM 18.13.

For $1 \leq n \leq 5$, the meet contraction postulate $[-n]$ entails the single contraction postulate $(-n)$, and the choice contraction postulate $\langle -n \rangle$ entails the single contraction postulate $(-n)$.

Proof. By instantiating S, S' in the $[-n]$ and $\langle -n \rangle$ postulates to singleton sets $\{A\}, \{B\}$. ■

18.4. Choice-contraction and conjunction-contraction, meet-contraction and disjunction-contraction, and iterated contraction

We have remarked above on the analogy between choice contraction and retracting a conjunction on the one hand, and between meet contraction and retracting a disjunction on the other hand. The question naturally arises as to whether these analogies can be translated into formal results. We shall consider now in some detail the relations between

- contraction by a conjunction,
- contraction by a disjunction,
- choice contraction,
- meet contraction, and
- iterated contraction.

The following is a list of (almost) all those inclusion relations between these kinds of contractions which can claim at least some *prima facie* plausibility. (One "missing link", (&C), will be discussed below.)

(MC)	$T-[A,B] \subseteq T-\langle A,B \rangle$
(C&)	$T-\langle A,B \rangle \subseteq T-(A \& B)$
(M&)	$T-[A,B] \subseteq T-(A \& B)$
(MI)	$T-[A,B] \subseteq T-A-B$
(vM)	$T-(A \vee B) \subseteq T-[A,B]$
(IM)	$T-A-B \subseteq T-[A,B]$
(vC)	$T-(A \vee B) \subseteq T-\langle A,B \rangle$

(&C)	$T-(A\&B) \subseteq T-\langle A,B \rangle$
(IC)	$T-A-B \subseteq T-\langle A,B \rangle$
(I&)	$T-A-B \subseteq T-(A\&B)$
(Iv)	$T-A-B \subseteq T-(A \vee B)$
(vI)	$T-(A \vee B) \subseteq T-A-B$
(v&)	$T-(A \vee B) \subseteq T-(A\&B)$
(II)	$T-A-B \subseteq T-B-A$

As it turns out, only the top four inclusion relations are valid – for the rest we can find counterexamples. We argue as follows.

LEMMA 18.14.

If $t \subseteq t'$, then $E_t(A) \subseteq E_{t'}(A)$ (for arbitrary bases t, t' and sentences A, B).

Proof. Suppose $S \in E_t(A)$. By the definition on an entailment set, that means that

(1) $A \in Cn(S)$,

(2) $S \subseteq t$, and

(3) $(\forall S')(S' \subset S \supset A \notin Cn(S'))$.

Since, by assumption, $t \subseteq t'$, we can infer from (2) that

(2') $S \subseteq t'$,

whence by the conjunction of (1), (2') and (3), S is an entailment set in t' for A, i.e. $S \in E_{t'}(A)$, as required. ■

COROLLARY 18.15.

If $t \subseteq t'$, then $R_t(A) \subseteq R_{t'}(A)$.

Proof. Suppose $t \subseteq t'$. Then, by the preceding lemma,

(*) $E_t(A) \subseteq E_{t'}(A)$.

So assume $B \in R_t(A)$. Then there is some $S \in E_t(A)$ with B minimal in S. By (*), $S \in E_{t'}(A)$; hence, $B \in R_{t'}(A)$. ■

THEOREM 18.16.

For any theory T on t and sentences A, B:

(MI) $T-[A,B] \subseteq T-A-B$.

Proof. Since $t \dot{-} A \subseteq t$, it follows by C18.15 that

(1) $R_{t\dot{-}A}(B)\subseteq R_t(B)$.

Hence, by set theory from (1),

(2) $R_t(A)\cup R_{t\dot{-}A}(B)\subseteq R_t(A)\cup R_t(B)$, and so

(3) $t—(R_t(A)\cup R_t(B))\subseteq t—(R_t(A)\cup R_{t\dot{-}A}(B))$, i.e.

(4) $t—(R_t(A)\cup R_t(B))\subseteq (t\dot{-}A)—R_{t\dot{-}A}(B)$.

By the monotonicity of Cn it follows from (4) that

(5) $Cn[t—(R_t(A)\cup R_t(B))]\subseteq Cn[(t\dot{-}A)—R_{t\dot{-}A}(B)]$,

as required. ■

THEOREM 18.17.

For any theory T on t and sentences A,B:

(C&) $\quad T-\langle A,B\rangle\subseteq T-(A\&B)$.

Proof. It will suffice to show that

$R(A\&B)\subseteq R(\langle A,B\rangle)$.

Note two facts.

Fact 1. $R(A\&B)\subseteq R(A)$ or $R(A\&B)\subseteq R(A)$
(and hence, $R(A\&B)\subseteq R(A)\cup R(B)$) (by L17.4 and C17.5).

Fact 2. $R(\langle A,B\rangle)=R(A)$ or $R(\langle A,B\rangle)=R(B)$ or
$R(\langle A,B\rangle)=R(A)\cup R(B)$ (by the def. of R for choice contraction).

Now assume $C\in R(A\&B)$. By Fact 2 we need to show that

(2) $C\in R(A)$ or $C\in R(B)$ or $C\in R(A)\cup R(B)$.

From the assumption it follows by Fact 1 that

(1) $C\in R(A)$ or $C\in R(B)$ and $C\in R(A)\cup R(B)$,

which entails (2). ■

COROLLARY 18.18.

For any theory T on t and sentences A,B:

(M&) $\quad T-[A,B]\subseteq T-(A\&B)$.

Proof. From [-7.1] and <-7> we obtain (MC), which in conjunction with (C&) (T18.17) entails (MC). ■

So much for the positive results: (MI), (C&), (MC), and (M&) are valid contraction principles. We now show that all of the remaining principles in the list on p.234 fail.

Example 1.

$t = \{r \to p, r, r \to q\}$; $r \to p \le r \le r \to q$.

It follows that ...

$E_t(p) = \{\{r, r \to p\}\}$, hence, $R_t(p) = \{r \to p\}$;

$E_t(q) = \{\{r, r \to q\}\}$, hence, $R_t(q) = \{r\}$;

$E_t(p \& q) = \{t\}$, hence, $R_t(p \& q) = \{r \to p\}$;

$t \dot{-} p = \{r, r \to q\}$; $t \dot{-} q = \{r \to p, r \to q\}$;

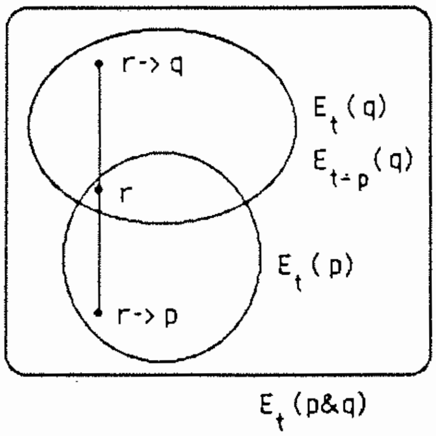
$E_{t \dot{-} p}(q) = \{r, r \to q\}$, hence, $R_{t \dot{-} p}(q) = \{r\}$;

$E_{t \dot{-} q}(p) = \varnothing$, hence, $R_{t \dot{-} p}(p) = \varnothing$;

$T - p - q = Cn(r \to q)$;

$T - q - p = Cn(r \to p, r \to q)$;

$T - (p \& q) = Cn(r, r \to q)$.



Example 1 shows how (II) and (I&) may fail. It is easily verified that all of (vM), (vI), (Iv), (v&), and (vC) fail for the next example.

Example 2.

$t = \{pvq, pvq \rightarrow p, pvq \rightarrow q\}$; $pvq \rightarrow p \approx pvq \rightarrow q \leq pvq$.

It follows that ...

$E_t(p) = \{\{pvq, pvq \rightarrow p\}\}$, hence, $R_t(p) = \{pvq \rightarrow p\}$;

$E_t(q) = \{\{pvq, pvq \rightarrow q\}\}$, hence, $R_t(q) = \{pvq \rightarrow q\}$;

$E_t(pvq) = \{\{pvq\}\}$, hence, $R_t(pvq) = \{pvq\}$;

$E_t(p\&q) = \{t\}$, hence, $R_t(p\&q) = \{pvq \rightarrow p, pvq \rightarrow q\}$;

$E_{t\dot{-}p}(q) = \{\{pvq, pvq \rightarrow q\}\}$, hence, $R_{t\dot{-}p}(q) = \{pvq \rightarrow q\}$;

$R_t([p,q]) = R_t(p) \cup R_t(q) = \{pvq \rightarrow p, pvq \rightarrow q\}$;

$R_t(p) \approx^2 R_t(q)$,

hence, $R_t(<p,q>) = R_t(p) \cup R_t(q) = \{pvq \rightarrow p, pvq \rightarrow q\}$;

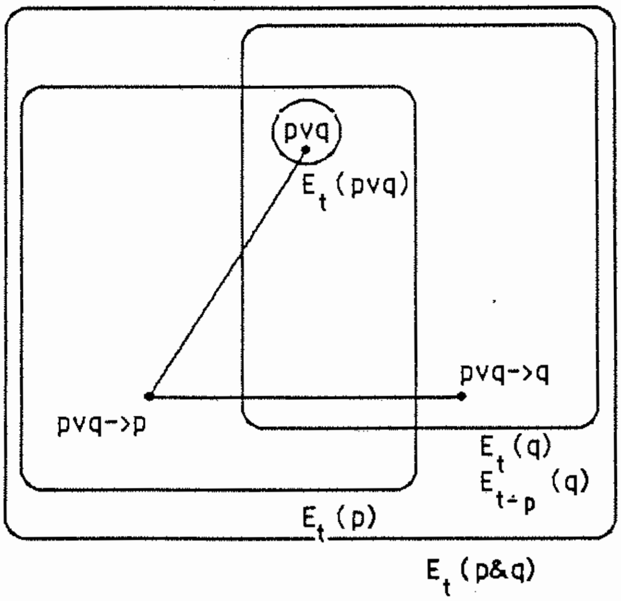
$T-(pvq) = Cn(pvq \rightarrow p, pvq \rightarrow q)$;

$T-(p\&q) = Cn(pvq)$;

$T-p-q = Cn(pvq)$;

$T-[p,q] = Cn(pvq)$;

$T-<p,q> = Cn(pvq)$;



It remains to show that (IM) and (IC) are not valid contraction principles. Note, that

(1) (IM) and (M&) entail (I&) , and

(2) (IC) and (C&) entail (I&) .

But we know that (I&) is false, and we also know that (M&) and (C&)

are true. Hence, neither (IM) nor (IC) are acceptable as valid contraction principles.

About the two examples it is interesting to observe that they both involve *logically irredundant bases* which are *linearly pre-ordered*. Thus, the results just obtained hold even for theories which are generated from bases satisfying the strongest assumptions considered so far.

Extending T18.13, we observe that the Gärdenfors postulate (-7) is derivable from either [-7.2] and (M&), or <-7> and (C&).

Finally, we note that meet contractions may be represented in terms of single contractions:

THEOREM 18.19.

The following contractions represent the same set:

(a) $T-[A,B]$,

(b) $(T-A-B)\cap(T-B-A)$, and

(c) $(T-A)\cap(T-B)$.

Proof. The inclusion (a)⊆(b) follows immediately from (MI). For (b)⊆(c) it suffices to observe that

$$R_t(A)\cup R_t(B) \subseteq (R_t(A)\cup R_{t\dot{-}A}(B))\cup(R_t(B)\cup R_{t\dot{-}B}(A)).$$

And (c)⊆(a) is an instance of [-7.2]. Hence, (a)=(b)=(c). ■

The results of this subsection are summarised in the figure below. (An arrow from one point to another stands for set inclusion in that direction.)

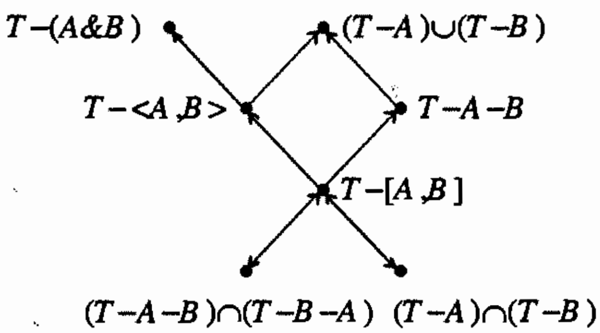


FIGURE 18.20

18.5. Is (finite) choice-contraction conjunction-contraction?

In the preceding section one not entirely implausible principle about the relation between conjunction-contraction and choice contraction had been omitted from the list of inclusion relations, viz.

(&C) $\quad T-(A\&B) \subseteq T-\langle A, B\rangle$.

The reason for this omission was that the status of (&C) remains an

Open Problem. Is (&C) a valid contraction principle?

If this question can be answered affirmatively, then (since (C&) is valid) we have for choice contractions by finite sets a result analogous to T18.19: finite choice contractions can be represented by single contractions, namely

$$T-\langle A_1 \cdots A_n\rangle = T-(A_1\& \cdots \&A_n) \text{ (for finite } n\text{).}$$

A positive solution to the Open Problem would yield at once some interesting consequences. Following the paths in Fig.18.20, we obtain the "covering" condition (cf. Alchourron, Gärdenfors and Makinson (1985), p.525)

(-C) $\quad T-(A\&B) \subseteq T-A \;$ or $\; T-(A\&B) \subseteq T-B$,

and, even stronger, the "ventilation" condition (cf. Alchourron, Gärdenfors and Makinson (1985), p.525)

(-V) $\quad T-(A\&B)=T-A$ or
$T-(A\&B)=T-B$ or
$T-(A\&B)=(T-A)\cap(T-B)$

(from the corresponding fact about choice contraction, i.e.

<-V> $\quad T-\langle A, B\rangle=T-A$ or
$T-\langle A, B\rangle=T-B$ or
$T-\langle A, B\rangle=(T-A)\cap(T-B)$.)

There is, however, a certain tension between (&C) and another principle – discussed earlier – whose status is uncertain, viz. (-8). For, from the identity $T-(A\&B)=T-\langle A,B\rangle$ (from (&C) and (C&)), we obtain by <-V> from (-8),

(-S) if $A \notin T-B$, then $T-B \subseteq T-A$.

But counterexamples to (-S) are easily construed.

Example 3.

$t = \{q \to p, r \to q, r\}$; $q \to p \leq r \to q \leq r$.

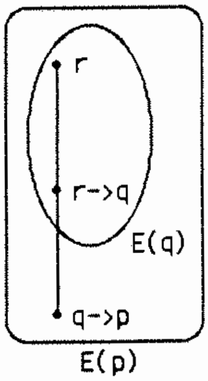
It follows that ...

$E(p) = \{\{r, r \to q, q \to p\}\}$, hence, $R(p) = \{q \to p\}$;

$E(q) = \{\{r, r \to q\}\}$, hence, $R(q) = \{r \to q\}$;

$t \dot{-} p = \{r, r \to q\}$; $t \dot{-} q = \{r, q \to p\}$;

$T-p = Cn(r, r \to q)$; $T-q = Cn(r, q \to p)$.



Thus, assuming that (C&) and <-V> are unobjectionable, one of (&C) or (-8) can not be a valid contraction principle. That is to say, at least one of the open problems posed in this section – concerning the validity of (-8) and that of (&C) – must be answered in the negative. The logical dependency relations among the principles presently under considerations are charted out in the following diagram. (Arrows stand for entailment. An arrow to a box indicates that all principles in the box are entailed by the principle(s) at the origin of the arrow; and, similarly, an arrow from a box indicates that the emboxed principles jointly entail the principle(s) pointed at.)

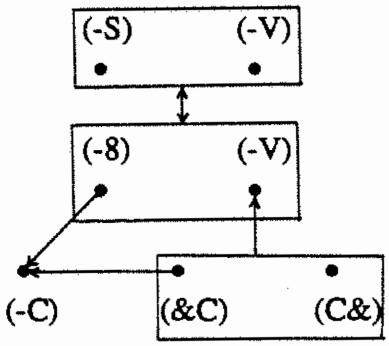


FIGURE 18.21

(Given <-V> as background principle, these entailment relations are easily established. For (-8)$\Rightarrow$(-C) use Success: $A\&B \notin T-(A\&B)$. Then either $A \notin T-(A\&B)$ or $B \notin T-(A\&B)$. So, by (-8), either $T-(A\&B) \subseteq T-A$ or $T-(A\&B) \subseteq T-B$, as required.)

19. Contractions and revisions

In section 16.1 we have observed that revisions ought to be treated as composite change operations. When one wishes to consistently add a sentence A to a theory, then one should first retract all conflicting information from the theory and then expand the result by A. One would expect that a contraction by ${\sim}A$ is an operation that opens the way for consistently adding A. So it appears natural to define the revision operation $*$ by means of the Levi identity

(LI) $$T{*}A := (T-{\sim}A)+A \ .$$

Of course, if $A \rightarrow {\sim}A$ is a theorem of our logic, then $T-{\sim}A+A=T{*}A$ is bound to contain both A and ${\sim}A$ in virtue of its closure under logical consequence. I shall say that a contraction $T-{\sim}A$ is *mind-opening* just in case A may be consistently added to $T-{\sim}A$ provided that $A \rightarrow {\sim}A \notin \mathrm{L}$. And a revision $T{*}A$ will be said to be *normal* if and only if ${\sim}A \notin T{*}A$, again, provided that $A \rightarrow {\sim}A \notin \mathrm{L}$. Thus a contraction of T by ${\sim}A$ is mind-opening just in case the revision of T by A is normal – that is to say, just in case the following condition is satisfied:

(*5) if $A \rightarrow {\sim}A \notin \mathbf{L}$, then ${\sim}A \notin T-{\sim}A+A = T^*A$.

The principal result of this section is that, depending on the choice of **L**, contractions may fail to be mind-opening and, hence, that revisions defined by means of (LI) may fail to be normal. Whatever the purpose of the contraction operation investigated in the preceding sections may be, it cannot generally be to safely pave the way for consistently adding A, even if $A \rightarrow {\sim}A$ fails to be a logical theorem.

Before turning to the question as to how a generally mind-opening contraction operation can be defined, let us first explore some of the properties of revisions as defined by the Levi identity. Like contractions by single sentences, single revisions will emerge (in section 19.3) as special cases of multiple revisions. So we shall consider the more general case directly. As in the case of multiple contractions, multiple revisions fall into either one of two kinds. A *meet revision* of a theory by some set S results in a theory that contains all members of S. By contrast, a *choice revision* of a theory by S may result in a theory to which not all but some sentences in S have been added. Both kinds of multiple revisions will now be discussed in turn.

19.1. Meet revision

Let S be a set of sentences. The set $\tilde{S}$ (the contradictory of S) is obtained from S by prefixing every sentence in S with the negation sign ${\sim}$.

$$\tilde{S} =: \{{\sim}A : A \in S\} .$$

Now we define the operation of meet revision by generalising the Levi identity (LI) as follows.

[LI] $T^*[S] := T-[\tilde{S}]+S$

Thus, the properties of the operation of meet revision are completely determined by those of the meet contraction operation (together with the consequence operation, defining $+$). Accordingly, we can derive from the postulates for meet contraction corresponding postulates characterising meet revisions via [LI] (given some background facts about the expansion operation $+$).

THEOREM 19.1.

If the operation of meet revision is defined by [LI] from the operation of meet contraction, then the following conditions hold for the meet revision of any theory T by any sets of sentences S, S'.

[*1] $T^*[S]$ is a theory;

[*2] $S \subseteq T^*[S]$;

[*3] $T^*[S] \subseteq T+S$;

[*4] if $T \cap \tilde{S} = \varnothing$, then $T+S \subseteq T^*[S]$;

[*6] if S and S' are pairwise equivalent, then $T^*[S] = T^*[S']$;

[*7] $T^*[S' \cup S] \subseteq T^*[S']+S$;

[*8] if $T \cap \tilde{S} = \varnothing$, then $T^*[S']+S \subseteq T^*[S' \cup S]$.

Proof. [*1] and [*2] are trivial.

For [*3] we need to show that

$$T-[\tilde{S}]+S \subseteq T+S$$

which follows immediately from [-2] and the monotonicity of +.

For [*4] we need to show that

$$T \cap \tilde{S} = \varnothing \Rightarrow T+S \subseteq T-[\tilde{S}]+S.$$

From the antecedent it follows by [-3] and [-2] that $T = T-[\tilde{S}]$, whence $T+S \subseteq T-[\tilde{S}]+S$.

[*6] follows immediately from [-5].

For [*7] observe that

$$T-[\tilde{S'} \cup \tilde{S}]+(S' \cup S) \subseteq (T-[\tilde{S'}] \cap T-[\tilde{S}])+S'+S$$

(by [-7.1] and the monotonicity of +), whence the proposition follows by set theory.

[*8] is immediate from [-8]. ■

I take it that the meet revision principles [*1] to [*8] express intuitive constraints on a revision operation. A revision operation should map theories into theories ([*1]). It should be successful in the sense that the revised theory should contain the target sentences ([*2]). As revision may require incision, revisions result in subsets of their corresponding expansions ([*3]). However, if the sentences to be added are consistent with the original theory, then no incision is needed – in which case revision reduces to expansion ([*4]). (We shall presently see that [*4] is not as robust a condition on revisions as it may appear on first sight.) That deductively equivalent sentences should be equivalent with respect to the revision operation ([*6]) is immediate from the corresponding facts about contractions and expansions. The supplementary principles [*7] and [*8] generalise [*3] and [*4] respectively; the latter result from the former by letting S' be the empty set.

THEOREM 19.2.

(a) [*7] implies [*3];

(b) [*8] implies [*4].

In labelling the principles in T30.2 we have left a gap. For [*5] we would like to have

$$\text{if } (\forall A \in S)(A \rightarrow \sim A \notin \mathbf{L}), \text{ then } T^*[S] \cap \tilde{S} = \varnothing.$$

In English: when T is revised to include all of S, then $T^*[S]$ ought to be consistent with respect to S; no negation of a sentence in S ought to be in $T^*[S]$ – except when S includes some sentence A which implies its negation according to our logic. A revision operation for which [*5] generally holds will be called *normal.* In section 19.4 it will be shown that if $*$ is defined as in [LI], then $*$ may fail to be normal in this sense.

19.2. Choice revision

The notion of a choice revision is perhaps less natural than that of a meet revision and useful applications are presumably harder to find. Considerations concerning the concept of a choice revision will therefore be kept comparatively brief, issuing only in a suggestion as to how the notion may be defined by way of a suitable generalisation of the Levi identity (LI).

Choice revisions are asked for in response to a request of the form: "Revise your theory so as to include at least one of $A_1, \ldots, A_n$!" The question is of course: which one? As in the case of choice contraction, we resolve this question by appeal to the principle that changes in a theory should not incur gratuitous loss of information. Thus, if we can add some A_i $(1 \leq i \leq n)$ without loosing any information contained in the original theory then that sentence should be added to the theory. But perhaps none of $A_1, \ldots, A_n$ can thus simply be added to the original theory – some incision may be needed in order to add at least some A_i. In that case we should add only those A_i which require an incision that is qualitatively minimal. We have to retract some sentences from the original theory; so let us retract only those sentences which are, by comparison, the least important (the most retractible) ones. Thus, in order to obtain the choice revision of T by S, $T^*\langle S\rangle$, we first retract from T those sentences in the contradictory $\tilde{S}$ of S which are "easiest" to retract, and then expand the result by those sentences in S whose negations we have decided to retract. This intuitive idea finds formal expression in the Levi identity for choice revision,

$$\langle LI\rangle \qquad T^*\langle S\rangle := T-\langle\tilde{S}\rangle+(S\cap\{A: {\sim}A\notin T-\langle\tilde{S}\rangle\}).$$

A simple example may serve to show how this definition works. Suppose that ${\sim}A\notin T$ and ${\sim}B\in T$ and that we want to choice-revise T by $\{A,B\}$. Then we first choice-contract T by $\{{\sim}A,{\sim}B\}$: $T-\langle{\sim}A,{\sim}B\rangle$. Since ${\sim}A\notin T$, there is nothing we need to remove from T in order to have ${\sim}A$ not in T. Thus, $R({\sim}A)=\varnothing$ – in contrast to $R({\sim}B)$ which is nonempty. Hence, $R({\sim}A)\leq^2 R({\sim}B)$ and so ${\sim}A\leq^t{\sim}B$. So we choose ${\sim}A$ to retract, retaining ${\sim}B$, i.e. $T-\langle{\sim}A,{\sim}B\rangle=T-[{\sim}A]$. If the contraction by ${\sim}A$ has been mind-opening, we can now add A – but not B! – to $T-\langle{\sim}A,{\sim}B\rangle$ to obtain the choice revision $T^*\langle A,B\rangle=T-[{\sim}A]+A$.

19.3. Revision by single sentences and Gärdenfors' supplementary postulates

As special cases, we obtain from the principles [*1] to [*6] for meet revisions the following principles for revision by single sentences.

(*1) T^*A is a theory

(*2) $A\in T^*A$

(*3) $T^*A\subseteq T+A$

(*4) if ${\sim}A\notin T$ then $T+A\subseteq T+A$

(*6) if $Cn(A)=Cn(B)$ then $T^*A=T^*B$

To obtain the full list of Gärdenfors' basic postulates for the revision operation, we would have to add

(*5G) if $\nvdash{\sim}A$, then T^*A is absolutely consistent.

However, we have developed our theory of theory change under the assumption that theories are not generally regular: they need not contain all theses of the logic under which they are closed. Logical theorems may be successfully removed from a theory and hence the theoremhood of $\sim A$ does not preclude that $\sim A \notin T^*A$.[16] But we certainly want T^*A to include A (*2) and we also want T^*A to be a theory (*1). Thus, if $A \rightarrow \sim A$ is a theorem of our logic, then $\sim A$ must be in T^*A. Given then that theories may fail to be regular, the antecedent of (*5G) ought to be replaced by the (relevantly, though not classically) weaker condition that $\nvdash A \rightarrow \sim A$.

But, secondly, we have recommended that theories be closed under a paraconsistent logic. If this recommendation is followed, then it is of comparatively little interest to know whether a theory is *absolutely* consistent. The picture we have advocated is that a theory may be *locally* inconsistent with respect to a number a sentences and that we ought to be in a position to remove these local inconsistencies separately, one by one, if we wish to do so. If a theory T is to be changed so as to consistently include A, then we should expect that T^*A be locally consistent with respect to A (provided that $A \rightarrow \sim A \notin \mathbf{L}$). But there is no reason to expect that T^*A will be locally consistent with respect to any sentence whatsoever; in order to add consistently to T a particular sentence A, we need not first make T absolutely consistent. So instead of (*5G) we should want the condition

(*5) if $A \rightarrow \sim A \notin \mathbf{L}$, then $\sim A \notin T^*A$

to hold. But, as we shall see in the next section, (*5) – the normality condition for revisions – may fail for the revision operation * defined from – and + by means of (LI).

In addition to the basic postulates (*1) to (*8) Gärdenfors has proposed two supplementary postulates:

(*7G) $T^*(A\&B) \subseteq T^*A + B$,

(*8G) if $\sim B \notin T^*A$ then $T^*A + B \subseteq T^*(A\&B)$.

However, given our explicit definition of the contraction operation, there are simple counterexamples to both (*7G) and (*8G), if the variable T occurring in the formulation of these principles ranges over relevantly closed theories. To show the invalidity of (*7G), we first provide a

[16] Cf. the discussion of (-4G) and (-4) in section 16.2.

counterexample to

(-7D) $T-(AvB) \subseteq T-A$

(the "dual" to (-7), $(T-A) \cap (T-B) \subseteq T-(A\&B)$).

Example 1.

$t = \{q, q \rightarrow p\}$; $q \rightarrow p \leq q$.

It follows that ...

$E(pvq) = \{\{q\}\}$, hence, $R(p) = \{q\}$;

$E(p) = \{\{q, q \rightarrow p\}\}$, hence, $R(pvq) = \{q \rightarrow p\}$;

$T-(pvq) = Cn(q \rightarrow p)$ and $T-p = Cn(q)$.

Example 1 refutes the validity of (-7D). Now expand both sides of (-7D) by $\sim A \& \sim B$. Example 1 may still serve to provide a counterexample to the thus obtained inclusion relation

(1) $T-(AvB)+(\sim A\& \sim B) \subseteq T-A+(\sim A\& \sim B)$.

(For, $T-(pvq)+(\sim p\& \sim q) = Cn(q \rightarrow p, \sim p, \sim q)$ is not a subset of $T-p+(\sim p\& \sim q) = Cn(q, \sim p, \sim q)$!) But (1) is equivalent to

(2) $T-(\sim\sim Av \sim\sim B)+(\sim A\& \sim B) \subseteq T-\sim\sim A+\sim A+\sim B$

which, by (LI), is the same as

(3) $T^*(\sim A\& \sim B) \subseteq T^* \sim A+\sim B$

– an instance of (*7G).

For a similar counterexample to (*8G), we shall look at the following instance of (*8G).

(4) if $B \notin T^* \sim A$ then $T^* \sim A+\sim B \subseteq T^*(\sim A\& \sim B)$.

Using (LI) and $T+A+B = T+(A\&B)$, (4) may be spelt out to

(5) if $B \notin T-A+\sim A$ then $T-A+(\sim A\& \sim B) \subseteq T-(AvB)+(\sim A\& \sim B)$.

The second example will be a counterexample to

(-8D) if $B \notin T-A$ then $T-A \subseteq T-(AvB)$

that can be extended so as to refute the "expanded version" (5) of (-8D).

Example 2.

$t = \{p, q, q \rightarrow p, pvq\}$; $q \leq q \rightarrow p$ (other $\leq$-pairs *ad libitum*, except that not($q \rightarrow p \leq q$)).

It follows that ...

$E(pvq) = \{\{p\}, \{q\}, \{pvq\}\}$, hence, $R(pvq) = \{p, q, q \rightarrow p\}$;

$E(p) = \{\{p\}, \{q, q \rightarrow p\}\}$, hence, $R(p) = \{p, q\}$;

$T-(pvq) = Cn(q \rightarrow p)$ and $T-(p) = Cn(q \rightarrow p, pvq)$.

Thus, $T-(pvq)$ is a proper subset of $T-p$ while the antecedent condition of (-8D), $q \notin T-p$, is satisfied. Moreover,

$$q \notin T-p+(\sim p \& \sim q)$$

while

$$pvq \in T-p+(\sim p \& \sim q)$$

and yet

$$pvq \notin T-(pvq)+(\sim p \& \sim q),$$

thus counterexampling (5) and, *a fortiori*, (*8G). (Note that this extension of Example 2 depends on the unavailability of the disjunctive syllogism (DS). For, if theories were closed under DS, then it would be the case that $q \in T-p+(\sim p \& \sim q)$, since both pvq and $\sim p$ are in $T-p+(\sim p \& \sim q)$.)

Gärdenfors (in (1988), section 3.3) motivates the inclusion of (*7G) and (*8G) among the postulates for revision as follows.

> "My aim is to formulate a generalisation of (*3) and (*4) that apply to *iterated* changes of belief. The idea is that if T^*A is a revision of T and T^*A is to be changed by adding further sentences, such a change should be made by using expansions of T^*A whenever possible. More generally, the minimal change of T to include both A and B, i.e. $T^*(A\&B)$ [!], ought to be the same as the expansion of T^*A by B, as long as B does not contradict the beliefs in T^*A."

The basic idea of the requirement expressed in the last sentence seems a most reasonable one and we may lay it down more formally in the two postulates (each an instance of [*7] and [*8] respectively)

(*7) $T^*[A, B] \subseteq T^*A+B$ and

(*8) if $\sim B \notin T^*A$ then $T^*A+B \subseteq T^*[A, B]$.

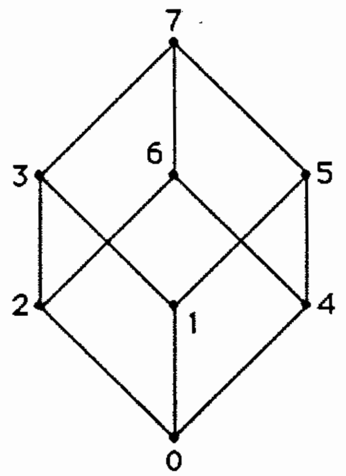
According to our account of revisions however, "the minimal change of T to include both A and B" had better not be represented by $T^*(A\&B)$ lest the intuitive constraint on such changes – mentioned in the above quotation – will not generally be met, as the counterexamples to (*7G) and (*8G) show. Instead, a "minimal change of T to include both A and B" is, naturally enough, a minimal change of T to include the set $\{A, B\}$, resulting in the revised theory $T^*[A,B]$; and that theory is indeed equivalent to T^*A+B whenever B is consistent with T^*A.

19.4. Mind-opening contractions and normal revisions

Assume that theories are closed under a relevant logic containing the schema

WI. $\qquad A\&(A \to B) \to B.$

Now consider a basis $t = \{p\&q \to {\sim}p, q\}$ and let T be the theory generated from that basis, i.e. $T = Cn(t)$. Note that ${\sim}p \notin T$. For the formula $q\&(p\&q \to {\sim}p) \to {\sim}p$ is not a theorem of any of the fully relevant logics of chapter one as the following matrix set, which satisfies **R**, shows:



$\to$	0	1	2	3	4	5	6	7	$\sim$
0	7	7	7	7	7	7	7	7	7
1	0	1	2	3	4	5	6	7	6
2	0	0	1	1	0	0	3	7	3
3	0	0	0	1	0	0	2	7	2
4	0	0	0	0	1	1	5	7	5
5	0	0	0	0	0	1	4	7	4
6	0	0	0	0	0	0	1	7	1
7	0	0	0	0	0	0	0	7	0

Designated values: 1, 3, 5, 7; $x\&y = min(x,y)$, $x \vee y = max(x,y)$.[17]

For $p=2$ and $q=4$, $q\&(p\&q \to {\sim}p) \to {\sim}p$ takes the undesignated value 0. Thus, since ${\sim}p$ is not derivable from t, $T - {\sim}p = T$. But if we expand $T = T - {\sim}p$ by p, then p is derivable from the resulting base $t \cup \{p\}$, since the formula

(1) $\qquad (p\&q)\&(p\&q \to {\sim}p) \to {\sim}p$

is an instance of WI. Furthermore, for the atom p, $p \to {\sim}p$ is not theorem of any of our logics. Thus we have produced a counterexample to the principle (*5), that for all sentences A and theories T,

[17] That the matrix satisfies **R** has been verified by the TESTER program.

(*5) if $A \rightarrow {\sim}A \notin \mathbf{L}$, then ${\sim}A \notin T^*A$.

In the terminology proposed earlier, since the contraction of T by ${\sim}p$ has failed to be mind-opening, the revision of T by p fails to be normal.

In contrast to revisions on theories closed under a relevant logic, the condition (*5) will always be met when theories closed under classical consequence $Cn_{\mathbf{K}}$ are under consideration.

THEOREM 19.3.

If T is closed under $Cn_{\mathbf{K}}$ then for all sentences A,

if $A \rightarrow {\sim}A \notin \mathbf{K}$, then ${\sim}A \notin T - {\sim}A + A$.

Proof. Assume that

(1) ${\sim}A \in T - {\sim}A + A$

and suppose for contradiction that

(2) $A \rightarrow {\sim}A \notin \mathbf{K}$.

Either (a) $T - {\sim}A = \varnothing$ or (b) $T - {\sim}A \neq \varnothing$. Suppose that (a) is the case. Then $T - {\sim}A + A = Cn(A)$. Hence, by (1), $A \rightarrow {\sim}A \in \mathbf{K}$, contradicting (2). So suppose that (b) is the case. Then it follows from (1) that

(3) $C \& A \rightarrow {\sim}A \in \mathbf{K}$

where C is the conjunction of some finite collection of members of $T - {\sim}A$. By Exportation we obtain from (3):

(4) $C \rightarrow (A \rightarrow {\sim}A) \in \mathbf{K}$.

But then it follows from (4) and Reductio, $A \rightarrow {\sim}A \rightarrow {\sim}A$, by transitivity that

(5) $C \rightarrow {\sim}A \in \mathbf{K}$

whence

(6) ${\sim}A \in Cn(T - {\sim}A) = T - {\sim}A$,

contradicting the principle (-4). ■

How ought we to interpret these results? On the one hand, it seems clear that we should want revisions to be normal. After all, it is intuitively the hallmark of revisions, in contrast to expansions, that they add *consistently* a sentence to a theory (provided that logic allows to do so, i.e. that $A \rightarrow {\sim}A \notin \mathbf{L}$). On the other hand, there is no reason to require that contractions be generally mind-opening; they only need to be mind-opening when made for the purpose of revising a theory. So the theory of contractions advanced in earlier sections need not be called into question in response to the above results. It does need supplementation though. When contracting a theory by $\sim A$ with a view to adding A to the contracted theory, then we ought to make use of a contraction operation which ensures that A can be added consistently, logic permitting. The contraction operation defined in earlier section does not generally have that property. Thus, for the purpose of ensuring the normality of revisions, we need a different, generally mind-opening contraction operation. I shall now offer a definition of such a *mind-opening contraction operation* and prove some basic – and, hopefully, attractive – results about this new operation.[18]

The basic idea of the following definition is that when T is to be contracted by $\sim A$ with a view to adding A, then the future addition of A should be anticipated in determining those sentences that ought to be deleted from the base t of T in order to block derivations of $\sim A$. We do not want sets to survive in $t \dot{-} \sim A$ which turn into entailment sets for $\sim A$ on adding A. Such entailment sets for A ought to be considered and pruned in advance. So for the purpose of determining the mind-opening contraction of $T = Cn(t)$ by $\sim A$, we should consider all entailment sets for $\sim A$ in the base t extended by A. However, in order to prune such entailment sets in the right way, some care has to be exercised with respect to how A is to be fitted into the ordering of comparative retractibility on t. After all, we do not want A to be the only minimal member of some entailment set for $\sim A$ in the A-extended base. For then the A-free part of such an entailment set may survive the contraction by $\sim A$ only to give us back $\sim A$ after A has been added again. The following definition of the *A-extension* of a base t is designed to prevent such cases.

[18] It should be noted that the need for a mind-opening contraction operation, as distinct from the contraction operation investigated earlier, does not only arise when non-classical theories are under consideration. For certain classical theories it can be shown that ordinary contractions fail to cut down theories to the size appropriate for the purpose of well-behaved revision; see Fuhrmann (1987).

DEFINITION 19.4.

Let t be a set of sentences (a base) ordered by a relation $\leq$ (of comparative retractibility). Let $\leq'$ be the restriction of $\leq$ to the set t—$\{A\}$. (Thus, if $A \notin t$, then $\leq = \leq'$.) The relation $\leq_A$ is the smallest set of pairs in $2^{t\cup\{A\}}$ such that

(a) $\leq_A \subseteq \leq'$, and

(b) for all $B \in t$—$\{A\}$, $B < A$ (i.e. $B \leq A$ and $A \not\leq B$).

The *A-extension* t/A of t is the set $t \cup \{A\}$ ordered by $\leq_A$.

The point of this definition will become clearer after we have stated the definition of a mind-opening contraction operation.

The (step-wise) definition of the mind-opening contraction operation $\ominus$ is as follows:

(We assume for simplicity that t is a virtually irredundant base. For $t/{\sim}A$ we shall henceforth write $t_{t/\bar{A}}$ and similarly $\leq_{\bar{A}}$ for $\leq_{({\sim}A)}$.)

(a) $D(A)_{t/\bar{A}} := \{S \subseteq t \cup \{{\sim}A\} : A \in Cn(S)\}$;

(b) $E(A)_{t/\bar{A}} := \{S \in D(A)_{t/\bar{A}} : (\forall S' \in D(A)_{t/\bar{A}})(\text{if } S' \subseteq S \text{ then } S \subseteq S')\}$;

(c) $R(A)_{t/\bar{A}} :=$
$\{B : (\exists S \in E(A)_{t/\bar{A}})(B \in S \text{ and } (\forall C \in S)(\text{if } C \leq_{\bar{A}} B \text{ then } B \leq_{\bar{A}} C))\}$;

(d) $T \ominus A := Cn(t - R(A)_{t/\bar{A}})$.

Not only do we now search for A-entailment sets in the $\bar{A}$-extension of t rather than in t, but sentences will now be deleted from the base t just in case they are minimal under the new ordering $\leq_{\bar{A}}$ in some entailment set for A in $t/\bar{A}$.

THEOREM 19.5.

For any base t,

$$\text{if } {\sim}A \rightarrow A \notin \mathbf{L}, \text{ then } {\sim}A \notin R(A)_{t/\bar{A}}.$$

Proof. Assume the antecedent and suppose for contradiction that ${\sim}A \in R(A)_{t/\bar{A}}$. Then there must be some set $S \in E(A)_{t/\bar{A}}$ with ${\sim}A$ minimal (under $\leq_{\bar{A}}$) in S. According to D19.4, since ${\sim}A$ is strictly maximal in $t/\bar{A}$, ${\sim}A$ must be strictly maximal in $S \subseteq t/\bar{A}$. Thus, ${\sim}A$ is minimal in S iff $S = \{{\sim}A\}$. But we have assumed that ${\sim}A \rightarrow A \notin \mathbf{L}$.

Hence, $S=\{\sim A\}\notin E(A)_{t/\overline{A}}$ – contradiction. ■

The theorem shows that, according to D19.4, $\sim A$ is fitted into $t/\overline{A}$ such that it can never be "blamed" for a derivation of A in the base t extended by $\sim A$. If $\sim A$ is not in t, such fitting is easy: we may just add $\sim A$ as a strictly maximal element to t, thus simply extending the order of comparative retractibility on t. However, if $\sim A \in t$, then we need to make sure that whenever $\sim A$ occurs in an entailment set S for A, then some element in S other than $\sim A$ – if there are any – will be minimal in S and hence deleted from the base. The new ordering $\leq_{\overline{A}}$ on $t\cup\{\sim A\}$ does exactly this by "shifting $\sim A$ to the top" while disturbing as little as possible the original ordering $\leq$ on t.

We can now prove that our new contraction operation $\ominus$ is indeed mind-opening and that the revision operation $\oplus$, defined as

(LI′) $$T\oplus A := T\ominus\sim A+A\ ,$$

is normal:

THEOREM 19.6.

If $A\rightarrow\sim A\notin \mathbf{L}$, then $\sim A\notin T\ominus\sim A+A$.

Proof. Assume that

(1) $A\rightarrow\sim A\notin \mathbf{L}$

and suppose for contradiction that

(2) $\sim A\in T\ominus\sim A+A$, i.e.

(3) $\sim A\in Cn((t—R(\sim A)_{t/A})\cup\{A\})$.

It follows from (1) by T19.5 that

(4) $A\notin R(\sim A)_{t/A}$.

Thus, by set theory from (3):

(5) $\sim A\in Cn(t\cup\{A\}—R(\sim A)_{t/A})$.

So there is some set $S\sqsubseteq t\cup\{A\}$ such that $\sim A\in Cn(S)$, i.e.

(6) $S\in D(\sim A)_{t/A}$, and

(7) $S\cap R(\sim A)_{t/A}=\varnothing$.

It follows from (6) that there is some subset S' of S such that

(8) $S' \in E(\sim A)_{t/A}$.

Let B be a $\leq_A$-minimal member of S'. Then $B \in R(\sim A)_{t/A}$. But also $B \in S' \subseteq S$. Hence,

(9) $S \cap R(\sim A)_{t/A} \neq \varnothing$,

contradicting (7). ■

The next theorem states that mind-opening contractions satisfy all but one of the other basic postulates for contractions.

THEOREM 19.7.

The operation $\ominus$ satisfies the following conditions for any theory T and sentences A, B :

($\ominus$1) $T \ominus A$ is a theory;

($\ominus$2) $T \ominus A \subseteq T$;

($\ominus$4) $A \notin T \ominus A$;

($\ominus$5) if $Cn(A) = Cn(B)$, then $T \ominus A = T o - B$.

Proof. ($\ominus$1) and ($\ominus$2) are immediate from the definition of $T \ominus A$.

For ($\ominus$4) assume for reductio that $A \in T \ominus A$, i.e. that

(1) $A \in Cn(t - R(A)_{t/\overline{A}})$.

Then there exists some entailment set S for A in $t - R(A)_{t/\overline{A}}$. So

(2) $S \subseteq t - R(A)_{t/\overline{A}}$

and since $t - R(A)_{t/\overline{A}} \subseteq t \subseteq t/\overline{A}$,

(3) $S \in E(A)_{t/\overline{A}}$.

Now S contains some $\leq_{\overline{A}}$-minimal element B. By the definition of $R(A)_{t/\overline{A}}$, $B \in R(A)_{t/\overline{A}}$. But since $B \in S$, it follows from (2) that $B \in t - R(A)_{t/\overline{A}}$, i.e. that $B \in t$ and that $B \notin R(A)_{t/\overline{A}}$ – contradiction.

($\ominus$5) follows, like (-5), from the fact that logically equivalent sentences determine identical reject sets. ■

The principle

($\ominus$3) if $A \notin T$, then $T \subseteq T \ominus A$

fails for mind-opening contraction. While A may not be deducible from T, A may become deducible once $\sim A$ has been added to T. Thus even if

A is not in T, there may be sentences in T that a mind-opening contraction by A will have to remove in order to make room for consistently adding $\sim A$. In such a case $T \ominus A$ will be a proper subset of T. The following counterexample to ($\ominus$3) is a variation on the example displayed at the opening of this section.

> Let $T = Cn(t)$ where $t = \{q \& \sim p \to p, q\}$ with $(q \& \sim p \to p) \leq q$. (Actually, any order on t may do for this example, including the empty order.) Then $p \notin Cn(t)$. Now
> $$t/\overline{p} = t \cup \{\sim p\}$$
> with
> $$(q \& \sim p \to p) \leq_{\overline{p}} q \leq_{\overline{p}} \sim p.$$
> Since $E(p)_{t/\overline{p}} = \{t/\overline{p}\}$, $R(p)_{t/\overline{p}} = \{q \& \sim p \to p\}$. So
> $$T \ominus p = Cn(t - R(p)_{t/\overline{p}}) = Cn(q, \sim p).$$
> But then $q \& \sim p \to p \in T$ while $q \& \sim p \to p \notin T \ominus p$.

The failure of ($\ominus$3) shows that a mind-opening contraction by A is not the right sort of operation to perform if one only wishes to change T as little as is necessary for A to be no longer in T. For, if A is not a member of T in the first place, then T need not change in any way in order to remove A from T. If the purpose of a contraction is merely the deletion of a sentence from a theory, then such a contraction ought to satisfy the vacuity condition (-3). If, on the other hand, contractions are made for the purpose of revision, then the vacuity condition is bound to fail in certain cases. A contraction operation that satisfies the vacuity condition (but may fail to be mind-opening) will henceforth be referred to as a *minimal contraction operation.*[19]

We have already shown that the operation $\oplus$, defined as

(LI′) $\qquad T \oplus A := T \ominus \sim A + A$

is normal (T19.6). In view of this result, we may refer to $\oplus$ as the *normal revision operation.* A further "nice" property of $\oplus$ is that it usually does not matter whether we define it by means of (LI′) or by

[19] The term is not meant to suggest that a minimal contraction operation (in our sense) satisfies Alchourron, Gardenfors and Makinson's requirements on minimal change. In particular, mininmal contraction operations need not satisfy the recovery postulate discussed earlier.

means of the commuted form of (LI′):

(LIc) $T \oplus A := T + A \ominus {\sim}A$.

In order to revise T by A, it (usually) does not matter whether we first remove ${\sim}A$ from T and then add A or whether we first add A and then remove ${\sim}A$. "Usually" here means always, except when $A \rightarrow {\sim}A$ is a theorem of our logic. For in the latter case ${\sim}A$ will be a member of $T \ominus A + A$ but – in view of ($\ominus$4) – not a member of $T + A \ominus {\sim}A$.

THEOREM 19.8.

For all theories T and sentences A:

(i) if $A \rightarrow {\sim}A \notin \mathbf{L}$, then $T \ominus {\sim}A + A \subseteq T + A \ominus {\sim}A$;

(ii) $T + A \ominus {\sim}A \subseteq T \ominus {\sim}A + A$.

Proof.

Ad (i). Assume that

(1) $A \rightarrow {\sim}A \notin \mathbf{L}$

and that for some sentence B,

(2) $B \in (t - R({\sim}A)_{t/A}) \cup \{A\}$.

By set theory from (2):

(3) ($B \in t$ and $B \notin R({\sim}A)_{t/A}$) or $B = A$.

By distribution from (3):

(4) $B \in t$ or $B = A$, and

(5) $B \notin R({\sim}A)_{t/A}$ or $B = A$.

Suppose $B = A$. It follows by T19.5 from (1) that $B \notin R({\sim}A)_{t/A}$; hence (5) simplifies to

(6) $B \notin R({\sim}A)_{t/A}$.

It follows from (4) and (6) that

(7) $B \in (t \cup \{A\}) - R({\sim}A)_{t/A}$.

Thus from (2) and (7):

(8) $B \in (t - R({\sim}A)_{t/A}) \cup \{A\} \subseteq B \in (t \cup \{A\}) - R({\sim}A)_{t/A}$.

Hence, by the monotonicity of Cn and the obvious identity $R({\sim}A)_{t/A} = R({\sim}A)_{t \cup \{A\}/A}$,

(9) $Cn[B \in (t - R({\sim}A)_{t/A}) \cup \{A\}] \subseteq Cn[B \in (t \cup \{A\}) - R({\sim}A)_{t \cup \{A\}/A}]$.

By definition, (9) may be replaced by

(10) $T\ominus{\sim}A{+}A\subseteq T{+}A\ominus{\sim}A$.

Ad (ii). Assume for some B that

(1) $B\in(t\cup\{A\})—R({\sim}A)_{t/A}$.

Then

(2) ($B\in t$ or $B{=}A$) and $B\notin R({\sim}A)_{t/A}$.

By distribution it follows from (2) that

(3) ($B\in t$ and $B\notin R({\sim}A)_{t/A}$) or $B{=}A$ whence

(4) $B\in(t—R({\sim}A)_{t/A})\cup\{A\}$.

So from (1) and (4),

(5) $(t\cup\{A\})—R({\sim}A)_{t/A}\subseteq(t—R({\sim}A)_{t/A})\cup\{A\}$.

From (5) we obtain, again by monotonicity of Cn and $R({\sim}A)_{t/A}{=}R({\sim}A)_{t\cup\{A\}/A}$,

(6) $Cn[(t\cup\{A\})—R({\sim}A)_{t\cup\{A\}/A}]\subseteq Cn[(t—R({\sim}A)_{t/A})\cup\{A\}]$

whence, by definition,

(7) $T{+}A\ominus{\sim}A\subseteq T\ominus{\sim}A{+}A$. ■

Corollary 19.9.

For all theories T and sentences A:

if $A\rightarrow{\sim}A\notin\mathbf{L}$, then $T\ominus{\sim}A{+}A={T}{+}A\ominus{\sim}A$.

We conclude this section with the observation that normal revisions are also "well-behaved" with respect to Gärdenfors basic postulates for revisions.

Theorem 19.10.

For all theories T and sentences A,B:

($\oplus$1) $T\oplus A$ is a theory;

($\oplus$2) $A\in T\oplus A$;

($\oplus$3) $T\oplus A\subseteq T{+}A$;

($\oplus$5) if $A\rightarrow{\sim}A\notin\mathbf{L}$, then ${\sim}A\notin T\oplus A$;

($\oplus$6) if $Cn(A){=}Cn(B)$, then $T\oplus A{=}T\oplus B$.

Proof. ($\oplus$1) and ($\oplus$2) are obvious from the definition of $T\oplus A$.

By ($\ominus$2), $T\ominus{\sim}A\subseteq T$. Hence, by the definition of $+$, $T\ominus{\sim}A{+}A\subseteq T{+}A$, i.e. $T\oplus A\subseteq T{+}A$.

For ($\oplus$5) see T19.6.

For ($\oplus$6) assume that $Cn(A)=Cn(B)$. Then $Cn(\sim A)=Cn(\sim B)$. Thus, by ($\ominus$5), $T\ominus\sim A=T\ominus\sim B$ whence (by the definition of $+$), $T\ominus\sim A+A=T\ominus\sim B+B$, i.e. $T\oplus A=T\oplus B$. ■

The condition

$$(\oplus 4) \qquad \text{if } \sim A\notin T\text{, then } T+A\subseteq T\oplus A$$

fails for essentially the same reasons for which ($\ominus$3) fails: while $\sim A\notin T$, T may contains sentences whose conjunction with A implies $\sim A$. Such sentences will have to be deleted in the move from T to $T\oplus A$.

20. Some comparisons with the AGM theory

By the AGM theory, I mean a cluster of explicit definitions of theory change operations – together with sets of postulates – put forward for consideration in a series of articles by Alchourron, Gärdenfors, and Makinson (for short: AGM). Although these definitions differ in important aspects, as we shall see, they also share certain features and it is these features that I shall focus on when comparing the proposals of this chapter with the AGM theory. As a background to the unifying aspects of the AGM approaches to theory change, it will be necessary to briefly chart out these approaches.

20.1. The AGM definitions of the contraction operation

The AGM theory subdivides into two branches, which I shall refer to as *'the partial meet branch'* and *'the safe branch'*.[20] Roughly, in the partial meet branch the contraction of a theory T by some sentence A is determined by considering maximally $\sim A$-consistent subsets of T; by contrast, in the safe branch the definition of the contraction operation proceeds by pruning minimal subsets of T from which A is derivable.

(a) The partial meet branch

[20] The relation between the two branches is investigated in Alchourron and Makinson (1986).

The partial meet branch of the AGM theory has been presented in a series of papers by AGM. An informal survey article is Makinson (1985). The main technical results can be found in Alchourron and Makinson (1982) and AGM (1985). For both formal results and informal discussion see Gärdenfors (1988).

Theories are sets of sentences closed under classical consequence. In order to define the contraction of a theory T by a sentence A, consider first the set $T \perp A$ of all maximal subsets S of T such that $A \notin Cn(S)$. If A is a tautology, $T \perp A$ will be empty. In that case we set $T-A$ to be T. In the principal case, where A is not a tautology, we let a selection function s choose "the best" sets in $T \perp A$ and define $T-A$ as the intersection of the chosen sets, i.e.

$$\text{(PMC)} \qquad T-A := \begin{cases} \cap(s(T \perp A)), \text{ if } T \perp A \neq \emptyset; \\ T \text{ otherwise.} \end{cases}$$

A contraction operation thus defined is a *partial meet contraction.* Various conditions on the choice function s may now be imposed, determining subclasses of partial meet contractions. The classes of contraction operations determined by two such conditions have been thoroughly investigated in the work of AGM.

- *Full meet contraction*: s selects *all* sets in $T \perp A$. Thus, $T-A$ is the intersection of all maximal subsets of T not yielding A (whenever there are such sets; $T-A = T$ otherwise). It turns out, however, that full meet contraction performs a rather radical excision on a theory: in the principal case of interest, where $A \in T$, $T-A$ reduces to $T \cap Cn(\sim A)$, i.e. only those consequences of $\sim A$ which are also in T (Observation 2.1 in Alchourron and Makinson (1982)).
- *Maxichoice contraction*: s selects a *unique* set out of $T \perp A$. (Again, if $T \perp A$ is empty, $T-A = T$.) The background assumption here is that there always is a "best" set in $T \perp A$. But on any interpretation of the selection function so far proposed, this assumption appears dubious.[21] Possibly worse, however, are the

[21] The Quinean example of Bizet and Verdi may serve again to cast intuitive doubt on the assumption underlying the definition of maxichoice contraction. The assumption that s should always yield a unique set bears a close analogy to a similar constraint on possible worlds models for counterfactual conditionals. In Stalnaker's semantics, a counterfactual $A>B$ is true at a world a just in case B is true at *the closest-to-a* world at which A holds. Lewis, on the other hand, has argued that several worlds may be equally

consequences of the uniqueness constraint on s for the concept of maxichoice revision, defined in the now familiar way by means of the Levi identity,

(LI) $T{*}A := T - {\sim}A + A.$

As shown in Alchourron and Makinson (1982) (Observation 3.1), maxichoice contraction satisfies the – already brow-raising – property

(1) $(A \vee B) \in T - A$ or $(A \vee {\sim}B) \in T - A$

for $A \in T$ and any B. But from (LM) and (1) we get the definitely unattractive

(2) $T{*}A$ is complete, whenever ${\sim}A \in T$.

(Alchourron and Makinson (1982), Observation 3.2).

Full meet contractions thus result in too small theories, while maxichoice contractions (and, in particular, revisions) blow up theories beyond reasonable limits. But plain partial meet contraction, i.e. partial meet contractions without the totality or the uniqueness condition on the selection function s, appears, by comparison, to be attractively sandwiched between full meet and maxichoice contraction. It can be argued, however, that this appearance is deceptive.

The intuitive justification for plain partial meet contraction is that it does not needlessly discard information while being attentive to the fact that at times we may find it impossible to linearly order all sentences in a theory with respect to their epistemic importance. Partial meet contraction is thus designed to reconcile the maximising aspect of maxichoice contraction with the need to retract sometimes more than what is *logically* required. To take a well-worn example, suppose both *Bizet is French (A)* and *Verdi is Italian (B)* are in T. We assume that we have no grounds to prefer A over B or B over A Then $s(T \perp (A\&B))$ will comprise both a maximal subset of T excluding A and a maximal subset of T excluding B. But B will be kept in the former, while A will be kept in the latter (by maximality). Hence, neither A nor B is in $\cap(s(T \perp (A\&B)))$, while $A \vee B$ is still in. So partial meet contraction seems to handle tie-cases like the one of Bizet and Verdi rather well.

But it is *only* for tie-cases that partial meet contraction (and its associated revision operation) avoids the properties (1) and (2). Whenever we can avoid retracting more than what is logically required –

close or similar to a given world. Formally, Lewis' semantics in terms of selection functions differs from Stalnaker's in waiving the condition that such a function must always pick out a unique world. See Stalnaker (1968) and Lewis (1973).

i.e. whenever no ties arise – the selection function will pick out a unique maximal subset of the theory in question: the contraction has in that case the properties of a maxichoice contraction. The refusal to impose the uniqueness condition on the selection function *tout court* cancels out the counterintuitive consequences of the maxichoice operation in some but not in all cases. It is little consolation that (1) and (2) fail to hold generally of partial meet contractions and revisions, when in fact they hold in all "normal" cases (i.e. cases where the balance always tips in favour of the one or the other sentence).

(b) The safe branch

A precursor to the concept of a *safe contraction* was first presented in Alchourron and Makinson (1981). Safe contractions proper are treated in Alchourron and Makinson (1985). The basic idea of safe contraction will already be familiar to the reader, since it is the same idea which also underlies the definition of the contraction operation proposed in this thesis.

Again, the definition of a safe contraction pertains to theories understood as classically closed sets of sentences.[22] But now we think of the members of a theory as ordered by some relation $<$; where $A<B$ intuitively expresses the idea that B is in some sense better (or less retractible, as we have said) than A. For a start, the only requirement on the relation $<$ is that it does not cycle: there are no $A_1, \ldots, A_n$ in T such that $A_1<A_2<\cdots<A_n<A_1$. A member B of T is *safe* (from removal) with respect to a sentence A just in case B is not minimal (under $<$) in any minimal (under $\subseteq$) subset S of T such that $A \in Cn(S)$. In the terminology of this thesis: B is safe in T with respect to A if and only if B is not $<$-minimal in any (classical) entailment set $S \subseteq T$ for A, i.e. B is not in the reject set for A in T. Let T/A be the set of all safe (w.r.t. A) elements in T, then

(SC) $\qquad T{-}A := Cn(T/A).$

[22] In Alchourron and Makinson (1985) the authors give also some consideration to safe contractions performed on arbitrary sets of sentences. (For arbitrary sets of sentences S, the definition (SC) below has to be slightly altered to $S\text{-}A := S \cap Cn(S/A)$.) The principal case investigated, however, is that of safe contraction as an operation on (classical) theories.

20.2. Contrasts

The theory offered in this thesis differs in three ways from both branches of the AGM theory.

(R) The units of theory change need not be sets of sentences closed under the consequence operation of classical logic. We have recommended use of a consequence operation induced by a *relevant logic*. If this recommendation is accepted, then inconsistent theories may be changed in the same piece-meal way in which consistent theories are changed, and theorems of logic can be retracted without retracting to the nullset.

(B) Theories are generated from *bases*. In the principal cases of interest, these bases are finite (and indeed "surveyable").

(M) Our theory focuses more generally on multiple change operations. Contractions and revisions by single sentences are simple instances of multiple contractions and revisions.

The last point should not be regarded as a substantial difference: there is no reason to expect that the AGM theory does not generalise to a theory of multiple change operations in a way similar to that given in section 18.

But (R) and (B) constitute significant departures from the AGM theory. I have already argued in section 17 why it is preferable, both from a methodological and a practical point of view, to develop a theory of theory change with theories as generated from (finite) bases as the principal units of investigation. To these considerations in favour of (B), another one has now to be added.

In section 16 I have argued that contraction operations should satisfy a *filtering condition*:

(F) If some sentence B is in a theory "just because" A is a member of that theory, then B should not be in the theory after A has been removed.

Now, intuitively it is clear what is meant by the phrase 'just because'. In practice we generate theories from small bases. Each sentence in the base should be thought of as having been included on its own merit: for a sentence to be included in the base of a theory, it does not depend on the presence of any other sentence in the base (this was the basic idea underlying the notion of a virtually irredundant base). But certain sentences in the theory, generated from the base in question do depend crucially on the presence of certain other sentences in the base; in this sense some sentences are in a theory just because certain others are. Note that for this explication of the filtering condition to yield any meaningful results, we need to assume that a theory is generated from a set of sentences which is, in the intuitive sense explained in section 17, virtually irredundant. The AGM theory, however, operates at a level of abstraction at which sight is lost of the base from which a theory has been generated. When the AGM theory is then applied to theories which in fact are generated from virtually (or near virtually) irredundant bases, the filtering condition will not be satisfied. Couched in the terminology favoured here, we may also say: contraction, according to the AGM theory, operates on theories as if they were generated from superredundant bases; the filtering condition is then satisfied under the condition that superredundant bases are always virtually irredundant – at which point we loose any intuitive grip on the notion of virtual irredundancy.[23]

That sufficiently large databases are likely to contain inconsistent pieces of information has become a fact of life. That even our most carefully constructed theories can turn out to be inconsistent is likewise well known, and the "end-users" of such theories frequently worry very little about it. Most people have a preference for consistent theories, however, and the rule of the game, when encountering an inconsistent corpus of information, is to restore consistency as soon as we can. But that may take some time. To take a very simple example, consider a database filled with addresses and the kind of information typically collected in a census. Retrieving information about N.N. we find that N.N. maintains one household and lives both in Sydney and Melbourne. Obviously, the information is inconsistent. I suppose, the rational course of action would be to retain all information for the time being, and to send a letter to both addresses, asking N.N. to supply the correct information. In many cases the inconsistency of data about particular subjects is even likely to go unnoticed. But the occasional "hidden" inconsistency should not affect the usefulness of the database as far as the correct information contained in it is concerned.

[23] It should be needless to say that AGM have never suggested that a superredundant base could be virtually irredundant. The fact of the matter is that the filtering condition simply plays no role in their theory. AGM place their emphasis on conserving information when contracting a theory – on pain of conserving "dangerous litter".

The moral of such examples is that we make frequent use of and maintain theories which are inconsistent with respect to some pieces of information. Such theories serve us well nevertheless, because we do not think of local inconsistencies as contaminating the consistent parts of a theory. And we may adjust inconsistent theories in the light of new evidence while carrying local inconsistencies over into the adjusted theory, whenever the adjustment is made in response to evidence unrelated to the inconsistencies in question.

The AGM theory fails in adequately representing both the static as well as the dynamic aspects of inconsistent theories. For, the AGM theory is firmly based on a classical notion of theoryhood; thus, inconsistent theories are trivial, comprising all well formed sentences of the language in which they are formulated. This does not mean, however, that, according to the AGM theory, inconsistent theories can not be properly contracted or revised (although they certainly can not be properly expanded). Both partial meet and safe contraction – and their associated revision operations – satisfy

(-4G) if $\nvdash A$ then $A \notin T - A$, and

(*5G) if $\nvdash \sim A$ then T^*A is consistent ,

even when T is inconsistent and, hence, trivial. But, according to the AGM theory, restoring consistency by means of a contraction, is an all-or-nothing matter. We cannot contract T so as to make it consistent with respect to A without making T *absolutely* consistent. The reason is, of course, that – in virtue of *ex falso quodlibet* – every maximal subset of T from which A is classically not derivable must be free from contradiction; and any inconsistent pair of sentences is a minimal subset of T classically implying A. Thus the adequacy condition that we should be able to make *local* adjustments in inconsistent theories is violated. By contrast, our departure (R) from the AGM theory, allows us to handle the statics and dynamics of inconsistent theories in a satisfactory satisfactory way.

21. The modal logic of theory change

The formal aspects of theory change is a relatively new field of inquiry. Despite the impressive work of Alchourron, Makinson, and Gärdenfors, there are many as yet unexplored avenues for future research.

Such unexplored avenues testify to the vitality of the subject. We have investigated in this chapter some of the consequences for theory change of closing theories under logics weaker than classical logic. One of the most striking such consequences is the need to distinguish between two different kinds of contraction operation, lest no satisfactory account of revisions will be forthcoming. Section 19 ended with some notes on the notion of a mind-opening contraction. But, obviously, the relation between minimal and mind-opening contraction needs further investigation. I shall conclude the present contribution to a new subject by yet another section in which the balance between results obtained in the past and suggestions for future research tips heavily towards the latter.

In this final section I shall suggest an alternative way of presenting a theory of theory change. Obviously, what we are dealing with in this theory are *processes* – transformations of one theory into another, changed, one. In recent years formalisations of reasoning about processes (dynamic logics) have attracted much attention among logicians and computer scientists. Thus, a natural idea is to explore the prospects for a special kind of dynamic logic: a modal logic with a set of theory change operators.

21.1. Dynamic logic

Dynamic logic generalises standard modal logic – with modal operators $\Box$ and $\Diamond$ – to encompass a denumerable infinity of modal operators $[\alpha_1],\langle\alpha_1\rangle, \ldots, [\alpha_n],\langle\alpha_n\rangle$. Intuitively, α stands for a process (type) and an expression of the form $[\alpha]A$ is intended to express that A is the case every time the process α has been completed. (Similarly, read $\langle\alpha\rangle A$: A may be the case after completion of α.) Dynamic logic was "invented" in the late seventies as a formal tool for reasoning about computer programs. As a logic of programs, α is more narrowly interpreted as a computer program, and A stands for a total state of the machine (or machine type) on which the program can be run. For our adaptation of the framework of dynamic logic to the needs of logic of theory change, we shall think of α as a change program (an expansion, contraction, or revision program) with a sentence supplied as its argument. In Artificial Intelligence such programs are known as *update programs*.[24]

[24] For good introductions to dynamic logic and several of its variants, see Harel (1979) and (1984), Segerberg (1980), or Parikh (1981). The application of dynamic logic to theory change is not a *completely* unexplored area. Manchanda has designed an

21.2. Propositional dynamic logic with expansion and minimal contraction programs

We start by extending our basic propositional language **La** to a language with unary program connectives **LaD**. To keep things at a reasonable level of simplicity, we shall draw on the fact that we have already defined the set $\mathbf{Wff}_{\mathbf{La}}$ of all well-formed formulae of **La**. Then the set **Wff** of well-formed formulae of **LaD** can be built up in six steps as follows.

I. The only *primitive program functor* is the contraction functor $-$.

II. 0 is a *primitive program expression*, and for each $A \in \mathbf{Wff}_{\mathbf{La}}$, $-A$ and $+A$ are primitive program expressions. The set of all primitive program expressions will be denoted by **Pr**.

III. **Pop** is the set $\{;\, , \sqcup\}$ of *operations on programs*.

IV. **Prog**, the set of all *programs* (variables: $\alpha,\beta,\gamma,...$), is the smallest set such that

(a) $\mathbf{Pr} \subseteq \mathbf{Prog}$, and

(b) if $\alpha \in \mathbf{Prog}$ and $\beta \in \mathbf{Prog}$, then $;\alpha\beta \in \mathbf{Prog}$ and $\sqcup\alpha\beta \in \mathbf{Prog}$.
(Instead of $;\alpha\beta$ and $\sqcup\alpha\beta$, we shall write $(\alpha;\beta)$ and $(\alpha\sqcup\beta)$ respectively, omitting brackets where no ambiguity of scope can arise.)

V. **Con** is the smallest set such that all the familiar *connectives* of **La** are in **Con** together with a unary connective (program modality) for each program expression, i.e.

(a) $\mathbf{Con}_{\mathbf{La}} \subseteq \mathbf{Con}$, and

(b) if $\alpha \in \mathbf{Prog}$, then $[\alpha] \in \mathbf{Con}$.

---

extension of PROLOG using dynamic logic – tailored to the purposes of database updating – as a basis. This new programming language, DYNAMIC PROLOG, improves on PROLOG's *assert* and *retract* functions. Manchanda's approach, however, has a very different flavour from the one chosen here; it has also severe limitations, since updates are only allowed with respect to literals (i.e. atomic formulae or their negations). (See Manchanda (1988), containing also references to earlier work by D.S. Warren.)

VI. The set of all *well-formed formulae* of **LaD**, **Wff**, is the smallest set such that

(a) $\mathbf{Wff}_{\mathbf{La}} \subseteq \mathbf{Wff}$,

(b) if $A \in \mathbf{Wff}$ and $B \in \mathbf{Wff}$, then $\sim A \in \mathbf{Wff}$, $[\alpha]A \in \mathbf{Wff}$, $A \vee B \in \mathbf{Wff}$, $A \& B \in \mathbf{Wff}$, and $A \rightarrow B \in \mathbf{Wff}$.

We think of programs as a sequence of commands converting one set of data (the input) to another (the output). Input and output need not be distinct. By a set of data, we mean a list of statements (sentences). Thus, the formula $[\alpha]A$ is meant to express that every time the program α is run, taking some contextually fixed data file as input, the statement A will be among the data retrievable from the output. More succinctly, we shall say that $[\alpha]A$ expresses that A holds after every run of α. The two compounding operations on programs are to be understood as follows. $\alpha \sqcup \beta$ stands for a *choice program*: run either α or β! Thus, $[\alpha \sqcup \beta]A$ expresses that A holds after each run of either α or β. $\alpha;\beta$ stands for the *program sequence*: run first α, then run β! Thus, $[\alpha;\beta]A$ states that A holds after running first α and then β.

The identity program 0 simply copies input into output – nothing changes. If we were only interested in what holds according to a given set of data, we could do without the identity program and just let A express that A holds (i.e. is an item in the file). But we are also interested in what does not hold and we shall not grant us the assumption that sets of data are always complete and consistent with respect to the language in which they are cast. Hence, to say – with respect to a set of data – that $\sim A$ holds, is not equivalent to saying that A is absent from the data set in question. We express the latter fact by means of the formula $\sim[0]A$. Thus, the addition of the identity program 0 makes an essential contribution to the expressive power of our language **LaD**.

For an axiomatisation of a logic of expansions and minimal contractions, we start off with a set of axioms and rules (including Modus Ponens and Adjunction) for our favourite logic **L**. To these we add a set of basic postulates common to all dynamic logics.

D1. $$\frac{A \rightarrow B}{[\alpha]A \rightarrow [\alpha]B}$$

D2. $$[\alpha]A \& [\alpha]B \rightarrow [\alpha](A \& B)$$

D3. $$[\alpha \sqcup \beta]A \leftrightarrow [\alpha]A \& [\beta]A$$

D4. $[\alpha;\beta]A \leftrightarrow [\alpha][\beta]A$

Thus, with respect to the operator $[\alpha]$, these postulates determine a conjunctively regular modal system (a C-modal system) based on L. In standard dynamic logic, it is usual to add to these schemas distribution of $[\alpha]$ over implication,

D5. $[\alpha](A \rightarrow B) \rightarrow .[\alpha]A \rightarrow [\alpha]B$,

and the rule of necessitation,

D6. $\dfrac{A}{[\alpha]A}$.

Since we are here interested only in the barest outline of a logic of theory change, we shall treat D5 and D6 as optional extras.

We extend this list of postulates for arbitrary programs by a set of axioms expressing some basic properties of expansion and contraction programs. First, four basic postulates for minimal contraction programs:

C1. $[\alpha;-A]B \rightarrow [\alpha]B$,

C2. $\sim[\alpha;-A]A$,

C3. $\dfrac{A \leftrightarrow A'}{[\alpha;-A]B \leftrightarrow [\alpha;-A']B}$,

C4. $\sim[\alpha]A \rightarrow .[\alpha]B \rightarrow [\alpha;-A]B$.

C1 to C4 are translations into **LaD** of the postulates (-2) (inclusion), (-4) (success), (-5) (preservation), and (-3) (vacuity).

The postulates for expansion programs are as follows.

E1. $([\alpha]B \rightarrow [\beta]B) \rightarrow ([\alpha;+A]B \rightarrow [\beta;+A]B)$,

E2. $[\alpha;+A]A$,

E3. $\dfrac{A \leftrightarrow A'}{[\alpha;+A]B \leftrightarrow [\alpha;+A']B}$.

E1 expresses that expansions are monotonic in the sense that if $T \subseteq T'$, then $T+A \subseteq T'+A$. E2 is a success condition for expansions, and E3 requires that the expansion of a theory by some sentence A is determined by no other considerations but the logical strength of A. In virtue of the general D-postulates, the results of contractions and expansions are closed under adjunction and implication according to our chosen logic; hence, they are theories in the sense in which we have used the term throughout this chapter.

The identity program 0 is governed by the two schemas

I1. $[0;\alpha]A \leftrightarrow [\alpha;0]A$

I2. $[0;\alpha]A \leftrightarrow [\alpha]A$.

The basic logic of expansions and minimal contractions based on **L**, **L.TC**, is defined as the smallest system that results from adding to the axioms and rules of **L** the above D-, C-, E-, and I-postulates. Before proceeding to the semantics of **L.TC**, we note that choice contraction programs are naturally represented in the language of **L.TC** by means of the compounding contraction programs by means of the operation $\cup$. There should however be no temptation to associate meet contraction programs with the ;-operation. The operation ; is not commutative – it essentially compounds programs sequentially and not concurrently.

The system **L.TC** may be modelled by generalising appropriately the notion of an R*(L)S-model to that of an R*(L)D-model

$$M = \langle \mathbf{0}, \mathbf{K}, R, D, *, V \rangle.$$

An R*(L)D-model is just like an R*(L)S-model, except that the single binary accessibility relation S is replaced by a function D, assigning to every program α a binary relation $S_\alpha \subseteq 2^{\mathbf{K}}$. The valuation clauses are as usual for the nonmodal connectives. For program connectives we have the schema

([α]) $a \vDash [\alpha]A$ iff $(\forall x)(S_\alpha ax \supset x \vDash A)$ $[\forall \alpha \in \mathbf{Prog}]$.

To ensure that the heredity condition is satisfied and that the modal axioms are true in all models, some conditions on the binary accessibility relations have to be imposed. (Again we write $S_\alpha a$ for $\{b : S_\alpha ab\}$ and $0a$ for $a \in \mathbf{0}$.)

(d0) If $a \rhd b$, then $S_\alpha b \subseteq S_\alpha a$.

This will secure the heredity lemma. Like the schemas RM and □C in the class of all R*(L)S-models, D1 and D2 are valid in all R*(L)D-models. By contrast, D3 and D4 will only be true in R*(L)D-models satisfying certain conditions. We first define:

Def. $S_\alpha | S_\beta ab$ iff $(\exists x)(S_\alpha ax \;\&\; S_\beta xb)$.

Then (d3) is the modelling condition corresponding to D3 and (d4) corresponds to D4.

(d3) $S_{\alpha\cup\beta}ab$ iff $S_\alpha ab$ or $S_\beta ab$.

(d4) $S_{\alpha;\beta}ab$ iff $S_\alpha | S_\beta ab$.

The modelling conditions for the contraction postulates C1 to C4, the expansion postulates E1 to E3, and the two identity postulates are as follows.

For any formula A and points a, b in a model:

(c1) $S_\alpha a \subseteq S_{\alpha;-A}\, a$

(c2) if $0a$, then $S_{\alpha;-A}\, a^* \not\subseteq |A|$

(c3) if $|A| = |A'|$, then $S_{\alpha;-A} = S_{\alpha;-A'}$

(c4) if $S_\alpha a^* \not\subseteq |A|$, then $S_{\alpha;-A}\, a \subseteq S_\alpha a$

(e1) if $S_\alpha a \subseteq |A|$ implies $S_\beta b \subseteq |A|$,
then $S_{\alpha;+A}\, a \subseteq |A|$ implies $S_{\beta;+A}\, b \subseteq |A|$

(e2) $S_{\alpha;+A}\, a \subseteq |A|$

(e3) if $|A| = |A'|$, then $S_{\alpha;+A} = S_{\alpha;+A'}$

(i1) $S_{0;\alpha} = S_{\alpha;0}$

(i2) $S_{0;\alpha} = S_\alpha$.

(Note that (c2) to (c4) and (e1) to (e3) are conditions on models rather than frames: the conditions are stated with reference to a function V evaluating formulae at points.)

Truth (in a model) and validity (in a class of models) are defined as usual and we make the obvious conjecture that the logic determined by the proof-theoretic postulates above (together with a suitable nonmodal background logic) is complete with respect to the class of all R*(L)D-models satisfying the corresponding semantic conditions.[25]

[25] "Obvious" is of course not to be understood in the sense of "obviously true". As usual, the soundness part of the conjecture is readily verified. Note that our dynamic logic with contraction programs lacks program connectives formed by means of the do-an-arbitrary-number-of-times program *. The presence of the *-program prevents a straightforward completeness proof by means of the canonical models method (see Segerberg (1982)). But without *, there is no reason to suspect that the completeness

The system **L.TC** provides only a *basis* for an intuitively satisfactory logic of expansions and minimal contractions. Although we conjecture that **L.TC** is complete with respect to the class of R*(L)D-models that satisfy the above conditions, the system remains seriously incomplete with respect to the intended interpretations of the program functors $-$ and $+$. Although it is not true of programs generally that running the same program more than once in a sequence results in the same output as running it only once, this is certainly an important property of expansion and contraction programs: expanding once by A is the same as expanding twice by A, and contracting once by A is the same as contracting twice by A. Thus, if program letters are to range over update programs only – that is, programs composed of expansion and contraction programs only – then all instances of the idempotence schema

D7. $[\alpha;\alpha]A \leftrightarrow [\alpha]A$ ought to be theorems.

The axiomatic characterisation of expansion programs in **L.TC** is also rather austere. Since we have

(+4) $T \subseteq T{+}A$ and

(+5) if $A \in T$, then $T{+}A \subseteq T$,

we should expect all instances of

E4. $[\alpha]B \rightarrow [\alpha;{+}A]B$ and

E5. $[\alpha]B \rightarrow .[\alpha;{+}A]B \rightarrow [\alpha]B$

to be theorems of a logic of expansions. (The conjunctive form of E5, $[\alpha]B \,\&\, [\alpha;{+}A]B \rightarrow [\alpha]B$ is of course a theorem of **L.TC**.) But, where **L** is a relevant logic, small R*(L)D-countermodels to both schemas are easily constructed. And E4 fails even for classical TC-logics. Hence, given the soundness of the **TC**-logics with respect to classes of R*(L)D-models, none of the relevant **TC**-logics contain either E4 or E5, and classical **TC**-logics do not contain E4. The system **L.TC** provides thus only the skeleton of a logic of theory change. However, we can conclude this section with a theorem to the effect that the skeleton is substantial enough to derive some central properties of revisions defined from expansions and minimal (not mind-opening!) contractions.

---

arguments of chapter II cannot simply be adapted to yield completeness for our logic of expansions and contractions.

THEOREM 21.1.

Let revision programs be defined as composites of expansion and minimal contraction programs by means of the following version of the Levi identity:

Def.* $[\alpha;{*A}]B := [\alpha;{-{\sim}A};{+A}]B$.

Then all instances of the following schemas are derivable in **L.TC**:

R1. $[\alpha;{*A}]A$;

R2. $[\alpha;{*A}]B \rightarrow [\alpha;{+A}]B$;

R3. $\dfrac{A \leftrightarrow A'}{[\alpha;{*A}]B \rightarrow [\alpha;{*A'}]B}$;

R4. ${\sim}[\alpha]{\sim}A \ \&\ [\alpha;{+A}]B \rightarrow [\alpha;{*A}]B$.

Proof. R1 is immediate from E2. For R3 combine C3 and E3.

Ad R2. By C1:

(1) $[\alpha;{-{\sim}A}]B \rightarrow [\alpha]B$.

By E1:

(2) $[\alpha;{-{\sim}A}]B \rightarrow [\alpha]B \rightarrow .[\alpha;{-{\sim}A};{+A}]B \rightarrow [\alpha;{+A}]B$.

whence from (1) by MP:

(3) $[\alpha;{-{\sim}A};{+A}]B \rightarrow [\alpha;{+A}]B$, i.e., by definition,

(4) $[\alpha;{*A}]B \rightarrow [\alpha;{+A}]B$.

Ad R4. By C4:

(1) ${\sim}[\alpha]{\sim}A \rightarrow .[\alpha]B \rightarrow [\alpha;{-{\sim}A}]B$.

By E1:

(2) $[\alpha]B \rightarrow [\alpha;{-{\sim}A}]B \rightarrow .[\alpha;{+A}]B \rightarrow [\alpha;{-{\sim}A};{+A}]B$.

Hence, by rule transitivity from (1) and (2):

(3) ${\sim}[\alpha]{\sim}A \rightarrow .[\alpha;{+A}]B \rightarrow [\alpha;{-{\sim}A};{+A}]B$, i.e., by definition,

(4) ${\sim}[\alpha]{\sim}A \rightarrow .[\alpha;{+A}]B \rightarrow [\alpha;{*A}]B$. ■

References

C.E. Alchourron, "Logic of norms and logic of normative propositions," *Logique et Analyse*, vol. 12, pp. 242-268, 1969.

C.E. Alchourron, "The intuitive background of normative legal discourse and its formalisation," *Journal of Philosophical Logic*, vol. 1, pp. 447-463, 1972.

C.E. Alchourron and D. Makinson, "Hierarchies of regulations and their logic," in *New Studies in Deontic Logic*, ed. Hilpinen, pp. 125-148, Reidel, Dordrecht, 1981.

C.E. Alchourron and D. Makinson, "On the logic of theory change: contraction functions and their associated revision functions," *Theoria*, vol. 48, pp. 14-37, 1982.

C.E. Alchourron, P. Gärdenfors, and D. Makinson, "On the logic of theory change: partial meet contraction and revision functions," *Journal of Symbolic Logic*, vol. 50, pp. 510-530, 1985.

C.E. Alchourron and D. Makinson, "On the logic of theory change: safe contraction," *Studia Logica*, vol. 44, pp. 405-422, 1985.

C.E. Alchourron and D. Makinson, "Maps between some different kinds of contraction function: the finite case," *Studia Logica*, vol. 45, pp. 187-198, 1986.

A.R. Anderson and N.D. Belnap, *Entailment. The logic of relevance and necessity,* 1, Princeton University Press, Princeton and London, 1975.

O. Becker, "Zur Logik der Modalitäten," *Jahrbuch für Philosophie und Phänomenologische Forschung*, vol. 11, pp. 497-548, Niemeyer Verlag, Halle, 1930.

J. van Benthem, *Essays in Logical Semantics,* Studies in Linguistics and Philosophy, 29, Reidel, Dordrecht, 1986.

R.T. Brady and R. Routley, "The non-triviality of extensional dialectical set theory," in *Paraconsistent Logics*, ed. Priest and Routley, Philosophia, Munich, 198+.

R.T. Brady, "The non-triviality of dialectical set theory," in *Paraconsistent Logics*, ed. Priest and Routley, Philosophia, Munich, 198+.

R.T. Brady, "Depth relevance of some paraconsistent logics," *Studia Logica*, vol. 43, pp. 63-73, 1984.

R.T. Brady, "Natural deduction systems for quantified relevant logics," *Logique et Analyse*, vol. 28, pp. 355-377, 1985.

C. Brink, "Power structures and logic," *Quaestiones Mathematicae*, vol. 9, pp. 69-94, 1986.

C. Brink and J. Heidema, "A verisimilar ordering of theories phrased in a propositional language," *Technical Report 4/87*, Automated Reasoning Project, ANU, 1987.

R. Bull and K. Segerberg, "Basic modal logic," in *Handbook of Philosophical Logic*, ed. Gabbay and Guenthner, vol. 2, pp. 1-88, Reidel, Dordrecht, 1984.

E. Bulygin and C.E. Alchourron, "Unvollständigkeit, Widersprüchlichkeit und Unbestimmtheit der Normenordnungen," in *Deontische Logik und Semantik*, ed. Conte, Hilpinen and von Wright, Linguistische Forschungen, vol. 15, pp. 20-32, Athenaion, Wiesbaden, 1977.

B.F. Chellas, *Modal Logic,* Cambridge University Press, Cambridge, 1980.

A. Church, "The weak theory of implication," in *Kontrolliertes Denken*, ed. Angsil, pp. 22-37, Alber, Munich, 1951.

A. Church, *Introduction to Mathematical Logic,* Princeton Univ. Press, Princeton, 1956.

B.J. Copeland, "On when a semantics is not a semantics,". *Journal of Philosophical Logic*, vol. 7, pp. 399-413, 1979.

B.J. Copeland, "Pure semantics and applied semantics," *Topoi*, vol. 2, pp. 197-204, 1983.

B.J. Copeland, "What is a semantics for classical negation?," *Mind*, vol. 95, pp. 478-490, 1986.

H.B. Curry, "The inconsistency of certain formal systems," *Journal of Symbolic Logic*, vol. 7, pp. 115-117, 1942.

J.M. Dunn, "Relevance logic and entailment," in *Handbook of Philosophical Logic*, ed. Gabbay and Guenthner, vol. 3, pp. 117-224, Reidel, Dordrecht, 1986.

B. Ellis, *Rational Belief Systems,* AQP Library of Philosophy, Blackwell, Oxford, 1979.

D. Foellesdal and R. Hilpinen, "Deontic logic: an introduction," in *Deontic Logic: Introductory and Systematic Readings*, ed. Hilpinen, pp. 1-35, Reidel, Dordrecht, 1971.

B. van Fraassen, "The logic of conditional obligation," *Journal of Philosophical Logic*, vol. 1, pp. 417-438, 1972.

A. Fuhrmann, "Resticted necessity and conditionals in relevant logics," typescript, Canberra, 198+. (To be submitted for publication.)

A. Fuhrmann, "Reflective modalities and theory change," *Technical Report 10/87*, Automated Reasoning Project, ANU, 1987. (Forthcoming in *Synthese*.)

P. Gärdenfors, "Rules for rational changes of belief," in *<320311>: Philosophical Essays Dedicated to Lennart Aqvist on His Fiftieth Bithday*, ed. Pauli, pp. 88-101, Dept of Philosophy, University of Uppsala, Uppsala, 1982.

P. Gärdenfors, "Epistemic importance and minimal changes of belief," *Australasian Journal of Philosophy*, vol. 62, pp. 136-157, 1984.

P. Gärdenfors, "Epistemic importance and the logic of theory change," in *Foundations of Logic and Linguistics*, ed. Weingartner, pp. 345-367, Plenum Press, New York, 1985.

P. Gärdenfors, *Knowledge, in Flux. Modeling the Dynamics of Epistemic States*, Bradford / MIT, Cambridge (Massachusetts), 1988.

D.M. Gabbay, *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*, Reidel, Dordrecht, 1976.

B. Hansson, "An analysis of some deontic logics," in *Deontic Logic*, ed. Hilpinen, pp. 121-147, Reidel, Dordrecht, 1971.

D. Harel, *First-Order Dynamic Logic*, Lecture Notes in Computer Science, 68, Springer, New York, 1979.

D. Harel, "Dynamic logic," in *Handbook of Philosophical Logic*, ed. Gabbay and Guenthner, vol. 2, pp. 497-604, Reidel, Dordrecht, 1984.

R. Hilpinen, ed., *Deontic Logic. Introductory and Systematic Readings*, Reidel, Dordrecht, 1971.

R. Hilpinen, ed., *New Studies in Deontic Logic*, Reidel, Dordrecht, 1981.

J. Hintikka, *Knowledge and Belief. An Introduction to the Logic of the Two Notions*, Cornell Univ. Press, Ithaca, 1962.

F. Jackson, *On Conditionals*, Blackwell, Oxford, 1987.

F. Jackson, "Understanding the logic of obligation," *Proceedings of the Aristotelian Society*, pp. 255-270, 1988.

J. Jackson and R. Pargetter, "Oughts, options, and actualism," *Philosophical Review*, vol. 95, pp. 233-255, 1986.

S. Jaskowski, "Propositional calculus for contradictory deductive systems," *Studia Logica*, vol. 24, pp. 143-160, 1969. (Originally published in Polish in 1948.)

H. Kelsen, *Reine Rechtslehre*, Springer, Vienna, 1960 (2nd edition).

S. Kripke, "Semantical analysis of modal logic I. Normal modal propositional calculi," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 9, pp. 67-96, 1963.

E.J. Lemmon, "New foundations for Lewis modal systems," *Journal of Symbolic Logic*, vol. 22, pp. 176-186, 1957.

E.J. Lemmon, "Deontic logic and the logic of imperatives," *Logique et Analyse*, vol. 8, pp. 39-71, 1965.

E.J. Lemmon, *An Introduction to Modal Logic,* American Philosophical Quarterly Monographs, 11, Blackwell, Oxford, 1977.

W. Lenzen, *Glauben, Wissen, Wahrscheinlichkeit. Systeme der epistemischen Logik,* Library of Exact Philosophy, 12, Springer, New York, 1980.

I. Levi, "Subjunctives, dispositions and chances," *Synthese*, vol. 34, pp. 423-455, 1977.

D. Lewis, *Counterfactuals,* Blackwell, Oxford, 1973.

D. Lewis, "Logic for equivocators," *Nous*, vol. 16, pp. 431-441, 1982.

M. Lockwood, (ed.), *Moral Dilemmas in Modern Medicine,* Oxford Univ. Press, Oxford, 1985.

D. Makinson, "On some completeness theorems in modal logic," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 12, pp. 379-384, 1966.

D. Makinson, "How to give it up: a survey of some formal aspects of the logic of theory change," *Synthese*, vol. 62, pp. 347-363, 1985.

D. Makinson, "On the status of recovery," *Journal of Philosophical Logic*, vol. 15, 1987.

S. Manchanda, "A dynamic logic programming language for relational updates," *Technical Report 2/88*, Dept of Computer Science, University of Arizona, Tucson, 1988.

T. Mautner, "The Absent-Minded Legislator," *Logique et Analyse*, vol. 14, pp. 105-110, 1971.

T. Mautner, "Flaws in Laws," *Philosophical Review*, vol. 82, pp. 83-98, 1973.

D. McDermott and J. Doyle, "Non-monotonic logic I," *Artificial Intelligence*, vol. 13, pp. 41-72, 1980.

D. McDermott, "Non-monotonic logic II," *Journal of the ACM*, vol. 29, pp. 33-57, 1982.

R.K. Meyer, *Topics in Modal and Many-Valued Logics,* University of Pittsburgh, Ph.D. dissertation, 1966.

R. Montague, "Pragmatics," in *Contemporary Philosophy: A Survey*, ed. Klibansky, vol. 1, pp. 102-122, La Nuova Editrice, Florence, 1968.

R.C. Moore, *Semantical considerations on non-monotonic logic*, SRI International, Technical Note 284, Menlo Park, 1983.

R. Parikh, "Propositional dynamic logics of programs: a survey," in *Logic of Programs*, ed. Engeler, Lecture Notes in Computer Science, vol. 125, pp. 102-144, Springer, New York, 1981.

G. Priest and R. Routley (eds.), *Paraconsistent Logics*, Philosophia, Munich, 198+ (forthcoming).

A.N. Prior, *Formal Logic*, Oxford Univ. Press, Oxford, 1955.

W.V.O. Quine, *Methods of Logic*, Holt, Rinehart, and Winston, New York, 1950.

W. Rautenberg, *Klassische und nichtklassische Aussagenlogik*, Logik und Grundlagen der Mathematik, 22, Vieweg, Braunschweig, 1979.

S.L. Read, *Relevant Logic. A philosophical examination of the basis of inference*, Blackwell, Oxford, 1988.

N. Rescher and R. Brandom, *The Logic of Inconsistency*, AQP Library of Philosophy, Blackwell, Oxford, 1980.

R. Routley and R.K. Meyer, "The semantics of entailment II," *Journal of Philosophical Logic*, vol. 1, pp. 53-73, 1972(a).

R. Routley and R.K. Meyer, "The semantics of entailment III," *Journal of Philosophical Logic*, vol. 1, pp. 192-208, 1972(b).

R. Routley and R.K. Meyer, "The semantics of entailment," in *Truth, Syntax and Modality*, ed. Leblanc, pp. 194-243, North-Holland, Amsterdam, 1973.

R. Routley, "The choice of logical foundations: non-classical choices and the ultralogical choice," *Studia Logica*, vol. 39, pp. 77-98, 1980.

R. Routley, R.K. Meyer, V. Plumwood, and R.T. Brady, *Relevant Logics and Their Rivals*, vol. 1, Ridgeview, Atascadero, 1982.

R. Routley and V. Plumwood, "Moral dilemmas and the logic of deontic notions," *Discussion Papers in Environmental Philosophy 6*, Departments of Philosophy, ANU, Canberra, 1984.

P.K. Schotch and R.E. Jennings, "Non-Kripkean deontic logic," in *New Studies in Deontic Logic*, ed. Hilpinen, pp. 149-164, Reidel, Dordrecht, 1981.

D.S. Scott, "Advice on modal logic," in *Philosophical Problems in Logic*, ed. Lambert, pp. 143-173, Reidel, Dordrecht, 1970.

D.S. Scott, "On engendering an illusion of understanding," *Journal of Philosophy*, vol. 68, pp. 787-807, 1971.

K. Segerberg, *An Essay in Classical Modal Logic,* Philosophical Studies, 13, Uppsala University Press, Uppsala, 1971.

K. Segerberg, "Applying modal logic," *Studia Logica*, vol. 39, pp. 275-295, 1980.

K. Segerberg, "A completeness theorem in the modal logic of programs," in *Universal Algebra And Applications*, Banach Center Publications, vol. 9, pp. 31 - 46, Polish Scientific Publishers (PWN), Warsaw, 1982.

J.K. Slaney, *Computers and Relevant Logics,* Australian National University, Ph.D. dissertion, 1980.

J.K. Slaney, "Reduced models for relevant logics without WI," *Notre Dame Journal of Formal Logic*, vol. 28, pp. 395-407, 1987.

R.C. Stalnaker, "A theory of conditionals," in *Studies in Logical Theory*, American Philosophical Quarterly Monographs, vol. 2, Blackwell, Oxford, 1968.

R.C. Stalnaker, *Inquiry,* Bradford / MIT, Cambridge (Massachusetts), 1984.

A. Tarski, "Fundamental concepts of the methodology of the deductive sciences," in *Logic, Semantics, Metamathematics*, pp. 60-109, Clarendon Press, Oxford, 1956. (Originally published in German in 1930.)

G.H. von Wright, "Deontic logic," *Mind*, vol. 60, pp. 1-15, 1951.

G.H. von Wright, *Norm and Action,* RKP, London, 1963.

G.H. von Wright, "On the logic of norms and actions," in *New Studies in Deontic Logic*, ed. Hilpinen, pp. 3-36, Reidel, Dordrecht, 1981.

ERRATA

Page, line(s)	For	Read
15, 6	D_m	D_n
18, 6	Rab'c	R'ab'c
18, 18	$D_i \rightarrow .B_i \rightarrow C_i$	$D_i \rightarrow .B_i \rightarrow C_i \in \mathbf{L}$
18, 24	$(D_1 \vee D_2) \rightarrow (B_1 \rightarrow C_1) \& (B_2 \rightarrow C_2) \in \mathbf{L}$	$(D_1 \vee D_2) \rightarrow (B_1 \rightarrow C_1) \vee (B_2 \rightarrow C_2) \in \mathbf{L}$
18, 26	$(B_1 \rightarrow C_1) \& (B_2 \rightarrow C_2) \rightarrow .B_1 \& B_2 \rightarrow C_1 \vee C_2 \in \mathbf{L}$	$(B_1 \rightarrow C_1) \vee (B_2 \rightarrow C_2) \rightarrow .B_1 \& B_2 \rightarrow C_1 \vee C_2 \in \mathbf{L}$
25, 6	logics logics	logics
26, 18	of 0, 1, 2 as	of 0, 2, 1 as
27, 16	severly	severely
28, 29	R*-model satisfies	R*-model M satisfies
30, 2	$b^* \nvDash B$	$b^* \nvDash B$
32, 24	original original	original
38, 19	if $\vdash_{\mathbf{BMC}}$ then	if $\vdash_{\mathbf{BMC}}$ A then
40, 18	some	any
41, 6-7	there exists (...) $A \notin b$.	for any formula A such that $A \notin b'$, there exists a superset b of b' such that b is a prime **L**-theory and $A \notin b$.
47, 16	model	modal
47, 18	contend	content
48, 30	$a \triangleright b$	0b
49, 16	follow	follows
50, 35-36	A thus name	A name
51, fn. 6	a eight	an eight
63, 10	rn	nn
63, 10	ri	ni
66, 24		*insert* (*) *before formula*
68, 29	$A \rightarrow B$	$A \rightarrow B \in T$
72, 36	applied".	applied".[3]
73, 14	level	level of
76, 2	belief	believes
82, 19	loosing	losing
86, 22	in	into
88, 23	commmitted	committed
99, 22	has	have
101, 5	loose	lose
106, 9	an harmless	a harmless
109, 37	is	if
114, 6	no subset	no proper subset
117, 19	derivations of C	derivations of A
118, fn.	(1977)	(1987)
124, 10	$A \notin T \Rightarrow T \subseteq T - A$	if $A \notin T$, then $T \subseteq T - A$

/...

Page, line(s)	For	Read
124, 12	Cn(A)=Cn(B) ⇒ T-A=T-B	if Cn(A)=Cn(B), then T-A=T-B
137, 11	=	⇐
139, 27	=	⇐
151, 15	R(A&B)⊆R(A) or R(A&B)⊆R(A)	R(A&B)⊆R(A) or R(A&B)⊆R(B)
152, 1	list on p.234	list on pp. 149f.
165, 18	then p is	then ~p is
165, 21	is not theorem	is not a theorem
167, 4	section	sections
168, 14	$t_t/\bar{A}$	$t/\bar{A}$
170, 14	To-B	T⊖B
171, fn.	Ga□denfors	Gärdenfors
172, 28, 31		*omit* B∈
173, 21	Gärdenfors	Gärdenfors'
180, 30	satisfactory satisfactory	satisfactory
181, 31	of logic	of a logic
182, 9	is (...) functor -	are the contraction functor - and the expansion functor +
191, 5	Bith-	Birth-
191, 33	J. Jackson	F. Jackson
192, 25	15	16