To appear in *South African Computer Journal*

Reasoning about Changing Information

Patrick Blackburn* Jan Jaspars† Maarten de Rijke‡

*Computerlinguistik, Universität des Saarlandes, Postfach 1150, D-66041 Saarbrücken, Germany
†CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
‡Dept. of Computer Science, University of Warwick, Coventry CV4 7AL, England

Abstract

The purpose of these notes is two-fold: (i) to give a reasonably self-contained introduction to a particular approach to theory change, known as the Alchourrón-Gärdenfors-Makinson (AGM) approach, and to discuss some of the alternatives, and extensions that have been proposed to it over the past few years; (ii) to relate the AGM approach to other 'information-oriented' branches of logic, including intuitionistic logic, non-monotonic reasoning, verisimilitude, and modal and dynamic logic.

Keywords: *Theory change, knowledge representation, modal logic, dynamic logic.*
Computing Review Categories: *F.4.0, F.4.1, I.2.3, I.2.4.*

1 Introduction

In these notes we study a number of approaches to *theory change* alias *belief revision* alias *belief change* alias *theory revision*. This enterprise is about coping with changing information, either because new facts have become known, or because the world has changed. If the newly obtained information is consistent with our old theory, there's no problem: we can simply extend our theory with the new piece of information. However, complex decisions need to be made when the new information conflicts with our old theory. For example, it may be that the new information casts a doubt on parts of our old theory — in that case we will probably want to get rid of the doubtful parts. But it may also be that the newly obtained information is in outright contradiction with our old theory. Assuming that we want to keep our theory consistent, this will force us to make adjustments. But how? Here's an example; assume that the following are part of our theory:

Bert is a post-doc in logic. (1)

Bert lives in Amsterdam. (2)

Amsterdam is located in the Netherlands. (3)

All Dutch post-docs in logic are unemployed. (4)

From our theory we can derive the following:

Bert is unemployed. (5)

Assume next that *as a matter fact* Bert happens to have a job, say at CWI. This means that we want to extend our theory with the fact $\neg(5)$. But then inconsistency strikes. So if we want to keep our theory consistent, we have to perform some kind of change, and give up some of the beliefs in our original theory. As we went through considerable effort to arrive at our theory in the first place, we don't want to give up the whole of it. But then, which of the *reasons* for the inconsistency do we have to give up? Also, which of the *consequences* of the old theory do we want to keep? For example, the following is a direct consequence of (4):

> All Dutch post-docs in logic who aren't Bert are unemployed.

Should we keep this (slightly weaker) generalization or not? This is not easy to decide. The complicating factor is that our theory is more than just a collection of atomic facts: there are complex logical dependencies between the elements of our theory, and logical considerations alone are not going to tell us which beliefs to give up.

Actual operations of theory change tend to be rather non-trivial functions whose definition may involve various orderings and relations on theories and sentences; usually, the additional structure reflects the importance of certain information. A number of general laws have been proposed to describe the behaviour of such operations; some of these are discussed below.

Semantically, one may view acts of belief revision as moves in an information space. The states of this space are some sort of information carriers, and a sentence ϕ can be part of a theory associated with an information state; in this case ϕ represents a static piece of information — it simply describes a belief engaged in that state. However, if ϕ does not belong to the theory associated with a given state, we may view it as an instruction telling us to move to a state whose theory does include ϕ; see Figure 1. Various formal languages for reasoning about such structures will be presented below.

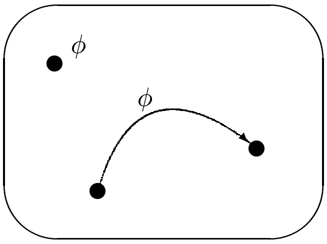


Figure 1. Information space

Similar ideas of viewing the meaning of sentences as instructions, actions or transitions have emerged in many current theories of natural language semantics, logic and artificial intelligence. Such approaches are usually called *dynamic* approaches; it turns out that there are fundamental links between various information-based approaches in logic, including theory change.

A Brief Overview

Part I

The *Alchourrón-Gärdenfors-Makinson* (or AGM) tradition proposes a set of postulates that are meant to govern the ways in which intelligent agents (human or mechanical) cope with theory change. In Part I of the notes we will discuss this theory, as well as some refinements, extensions and alternative proposals.

In addition to the AGM postulates we discuss various proposals for giving explicit definitions of operations of theory change that satisfy these laws.

Part II

In Part II of the notes we change tack. We consider various information-oriented formalisms in logic, and try to link these to AGM theory. The formalisms considered are non-monotonic logic, verisimilitude, and descriptive approaches based on modal and dynamic logic.

I–The AGM Theory

2 Preliminaries

Before we introduce the Alchourrón-Gärdenfors-Makinson theory of belief change, we review a number of underlying basic assumptions, both concerning the methodology and concerning the conceptual and logical apparatus.

Kinds of Theory Change

We will assume the beliefs are represented by sentences of some formal language, and that beliefs ϕ can only be

accepted, or
rejected (accept $\neg\phi$), or
neither accepted nor rejected.

At first sight there only seem to be two basic kinds of theory change, namely

- to insert (or accept) information, and
- to delete information (that is, to switch from acceptance to rejection, or to 'neither acceptance nor rejection').

What we are interested in is how, and under which circumstances, these basic actions are performed. The proposals for handling theory change found in the literature can be divided into two kinds: a *direct* mode, and an *indirect* mode.

In the direct mode one simply inserts or deletes information without bothering about the consistency requirement. Such simple operations are accompanied by a complex, usually para-consistent or defeasible inference engine to determine which conclusions can actually be drawn from the theory. Thus, in the direct mode the complexity of theory change is hidden in the inference engine. Truth maintenance systems form an important example of the direct mode (see [6]).

In the indirect mode one tries to perform theory changes subject to (some or all of) the following methodological assumptions:

Consistency. The beliefs in a theory should be kept consistent whenever possible. This assumption may well be the dominating motive for the whole enterprise of 'theory change.' It is certainly what distinguishes theory change from from such fields as para-consistent logic, where one is also interested in handling conflicting information without, however, necessarily deleting reasons for conflict, by changing the inference engine.

Closure. If the theory implies a belief ϕ, then ϕ should be in the theory. This is an obvious idealization, but for the time being we will adhere to it.

The Consistency and Closure assumptions concern the static aspects of theory change: the things that are actually being changed. The following two assumptions concern the way in which theory change takes place.

Minimality. The amount of information lost in a belief change should be kept minimal. The idea is that information doesn't come for free and unnecessary losses are therefore to be avoided.

First Things Last. If there is a measure according to which some beliefs are considered to be more important than others, one should give up the least important ones first.

Functionality. For every theory K and every sentence ϕ, there is a *unique* theory representing the removal or addition of ϕ from or to K. In other words: theory change is a *function* from theories and formulas to theories.

If one tries to play the game following the above constraints, the theory change operations themselves become highly non-trivial, but in return one can use standard logics as the underlying inference engine.

The AGM theory which we will discuss below is the most prominent example of belief revision in the indirect mode. In the AGM approach three main kinds of belief change are considered:

Expansion. A new sentence ϕ consistent with the old theory K is added to K. The belief system that results from expanding K by ϕ is denoted with $K+\phi$. Expansions of a theory K result in theories that are at least as large as K.

To illustrate the use of expansions, consider a situation in which the theory K is simply (the set of logical consequences of) the sentence 'Bert wrote five job applications last month.' Now suppose that a job opening at the University of Warwick was announced; this new fact could simply be added to K.

Revision. A new sentence ϕ is *in*consistent with the theory K; ϕ is added to K, but to preserve consistency, some old sentences from K are deleted. We use $K*\phi$ to denote the result of revising K by ϕ. Revisions of K may lead to theories that are neither extensions of K nor subsets of K.

Recall the earlier example involving Bert and the unemployed post-docs; there the incoming information (Bert has a job) conflicted with a sentence that was derivable from the database (Bert is unemployed). To incorporate the new information we have to revise the old theory with the new information.

Contraction. Somewhat intermediate between expansion and revision is contraction. Some sentence is retracted from K without new information being added. To ensure logical closure further sentences from K may have to be given up. We write $K-\phi$ to denote the result of contracting K with ϕ. Contractions of K lead to subsets of K.

To understand contractions it is best to think of thought experiments. Gärdenfors [10] considers the story of Oscar who wonders what would have become of his life if he hadn't married his wife; would he have had the drinking problem that he's developing?

Now for the crucial question: How should we define expansion, contraction, and revision operators, given a theory K? Before we can attempt to answer this, we need to review a number of assumptions about our background logic.

Logical Bits and Pieces

We will assume that our background logic **L** is classical propositional logic, and that we have the usual boolean connectives $\neg$, $\vee$, $\wedge$, $\rightarrow$, $\leftrightarrow$ available, as well as the constants $-$ and $\top$. We write $K \vdash_{\mathbf{L}} \phi$ if there is a proof (axiomatic, or otherwise) of ϕ from K in **L**. Usually we will work with a consequence operator Cn instead of the turnstile $\vdash$:

$$\psi \in \mathrm{Cn}(K) \text{ iff } K \vdash_{\mathbf{L}} \psi.$$

Note that the consequence operator Cn satisfies the following conditions:

Cut. If $\phi \in \mathrm{Cn}(K)$, $\psi \in \mathrm{Cn}(K \cup \{\phi\})$, then $\psi \in \mathrm{Cn}(K)$.

Deduction. If $\psi \in \mathrm{Cn}(H \cup \{\phi\})$ then $(\phi \rightarrow \psi) \in \mathrm{Cn}(H)$.

Compactness. If $\phi \in \mathrm{Cn}(H)$, then $\phi \in \mathrm{Cn}(H_0)$ for some finite $H_0 \subseteq H$.

We will sometimes need to exploit further properties of Cn such as monotonicity (if $H \subseteq K$, then $\mathrm{Cn}(H) \subseteq \mathrm{Cn}(K)$) or reflexivity ($H \subseteq \mathrm{Cn}(H)$).

We will model belief states by means of sets of **L**-sentences. That is, a *belief set* is a set K of **L**-sentences that is closed under logical consequence (i.e., $\mathrm{Cn}(K) \subseteq K$). We use $K_\perp$ to denote the inconsistent belief set; that is, the set of all formulas.

The advantage of modeling belief states by means of belief sets is that this approach handles facts, constraints and derivation rules in a uniform way. It is also a convenient way of modeling partial information. The disadvantages are that working with logically closed sets is an idealization that will cause problems when it comes to implementation, since in general such sets will be infinite. Moreover, belief sets don't take into account the fact that belief systems are structured: some beliefs do not have an independent standing but arise as consequences of more basic beliefs — when we perform revisions or contractions we act on some *finite* base for the belief set. Formally, we say that a set of sentences H is a *belief base* for a belief set K if $H \subseteq K$ is finite and $\mathrm{Cn}(H) = K$. Below we will see how theory change can be made to work on belief bases instead of belief sets; this will lead us to consider so-called 'base revisions.'

Logic Suffices to Define Expansions

Defining explicit operations of theory change that implement the earlier kinds of theory change turns out to be virtually trivial in the case of expansions.

Definition 2.1 Let some collection of propositional variables be given, let **L** be classical propositional logic over this language, and let K be a set of sentences that is **L**-consistent and closed under **L**-consequences. Let ϕ be a sentence in the language of **L**. We define $K+\phi$ (the expansion of K by ϕ) as follows:

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

That is: $K+\phi$ is the logical closure of K together with ϕ.

Unfortunately, it is not possible to give similar simple and explicit definitions of revisions and contractions. The problem was already hinted at following the 'post-doc from Amsterdam' example in Section 1: when trying to accommodate new information there is no purely logical reason for choosing to delete one piece of information rather than another. One important underlying issue is that theories are not just collections of atomic facts, but collections of facts from

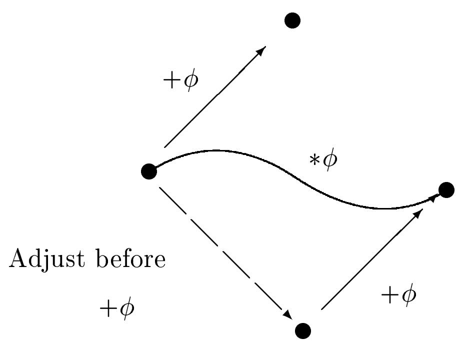


Figure 2. A global perspective on revision

which other facts can be derived. It's the logical dependencies between facts and derived facts that causes the difficulties.

From here we can proceed in two directions:

1. Present explicit constructions of revision and contraction functions.
2. Come up with standards for revision and contraction functions against which such constructions can be tested.

Below we will start with the second item, namely with the so-called AGM postulates for revision and contraction. After that we will come up with example constructions that satisfy those postulates.

3 AGM Postulates for Theory Change

In this section we will work with belief sets, and our goal will be to formulate the AGM postulates for rational revision and contraction functions defined over such sets.

Postulates for Revision

In Figure 2 we give a pictorial representation of the possible structure of a revision function. The underlying intuition is that a revision with ϕ consists of an adjustment to prepare for ϕ, followed by an expansion with ϕ. Clearly, the adjustment function is the vague part of this composition. A naive candidate which comes to mind first is to take the adjustment to be the converse operation of the expansion function $+$ (that is, simply removing a formula). But this idea doesn't work, for such an operation most often leads to no adjustment at all. To see how this can happen, consider the following example:

$$\mathrm{Cn}(\mathrm{Cn}(\{p, p \to q\}) \setminus \{q\}) = \mathrm{Cn}(\{p, p \to q\}). \qquad (6)$$

If we really wish to remove q from $\mathrm{Cn}(\{p, p \to q\})$, we also need to remove everything from this belief set from which we can infer q.

Another idea is to define the adjustment function to be some maximal belief set contained in the result of performing a retraction as just sketched. The difficulty here is that there can be more than one such a maximal belief subset, and then revision would no longer be a function. Fortunately, there are ways to make some satisfactory new belief set out of these maximal parts, as we shall later see. In the following subsections we will give a set of postulates for revisions and contractions which restrict the searchspace for good definitions of this adjustment function.

We first turn to the postulates for revision. For the time being, we take a revision function $*$ to be a function from belief sets and sentences to belief sets.

The Basic Postulates

The first two postulates for revision correspond to the two methodological remarks made earlier.

($K*1$) For any sentence ϕ, and any belief set K, $K * \phi$ is a belief set. (Closure)

This postulate says that outputs of revisions are indeed belief sets. The second one says that the input sentence ϕ is accepted in $K * \phi$.

($K*2$) $\phi \in K * \phi$ (Success)

This means that incoming information is given absolute priority over the original theory.

If the incoming information ϕ is consistent with our old theory K, we can simply expand K with ϕ, and close it under logical consequence. It is only when the incoming information ϕ contradicts what is known in the theory K (i.e., if $\neg\phi \in K$) that we have to resort to revision.

($K*3$) $(K * \phi) \subseteq (K + \phi)$ (Expansion 1)

($K*4$) If $\neg\phi \notin K$, then $(K + \phi) \subseteq (K * \phi)$ (Expansion 2)

The ($K*3$)-principle completes the square structure of Figure 2. That is, revising by ϕ will never get you more information than expanding with ϕ, and if ϕ is not rejected by the original theory K, then expanding and revising with ϕ yield the same result. This can be paraphrased as 'in the consistent case, logic suffices to guide belief revision.' It also implements the minimality idea that we should change as little as possible when we perform a revision.

The purpose of a revision is to produce a new *consistent* belief set. Thus, $K * \phi$ should be consistent whenever ϕ is:

($K*5$) $K * \phi = K_\perp$ iff $\vdash_{\mathbf{L}} \neg\phi$ (Consistency Preservation)

Furthermore, only the content of the input should matter, not its actual formulation. That is: logically equivalent sentences should yield identical revisions.

($K*6$) $\vdash_{\mathbf{L}} \phi \leftrightarrow \psi$ implies $K * \phi = K * \psi$. (Extensionality)

Postulates ($K*1$)–($K*6$) are the basic postulates governing the revision operator $*$. Before we go on to add two further postulates, we consider some elementary consequences of ($K*1$)–($K*6$). First of all, assuming that our background logic $\mathbf{L}$ is consistent, we do *not* have commutativity of revisions:

$$(K * \phi) * \psi = (K * \psi) * \phi. \qquad (7)$$

For, to arrive at a contradiction, assume that we do have commutativity. Take ψ to be $\neg\phi$ in (7). Then, by $(K{*}2)$, $\neg\phi \in (K * \phi) * \neg\phi$ and $\phi \in (K * \neg\phi) * \phi$, and hence

$$\phi, \neg\phi \in (K * \phi) * \neg\phi = (K * \neg\phi) * \phi = K_\perp.$$

By $(K{*}5)$ it follows that $\vdash \neg\neg\phi$ and $\vdash \neg\phi$. In other words, $\vdash -$.

To spell out the relation between consistency and revision even further, we assume that our consequence operator Cn is monotonic. Then $(K{*}4)$ is equivalent to

$$\text{if } \phi \text{ is consistent with } K\text{, then } K \subseteq K * \phi. \qquad (8)$$

To prove that $(K{*}4)$ implies (8), assume $(K{*}4)$ and the joint consistency of ϕ and K, that is: $\neg\phi \notin K$. By $(K{*}4)$ it follows that $K + \phi \subseteq K * \phi$. So then

$$K \subseteq \mathrm{Cn}(K) \subseteq \mathrm{Cn}(K + \phi) \subseteq K + \phi \subseteq K * \phi,$$

as required. For the converse implication, assume (8) and assume also that $\neg\phi \notin K$. Then, by $(K{*}2)$ we have $K \cup \{\phi\} \subseteq (K * \phi) \cup \{\phi\} = K * \phi$. So

$$K + \phi = \mathrm{Cn}(K \cup \{\phi\}) \subseteq \mathrm{Cn}(K * \phi) = K * \phi,$$

by monotonicity and $(K{*}1)$.

Two More Postulates for Revision

The two additional postulates for theory revision that are usually considered in conjunction with the basic postulates $(K{*}1)$–$(K{*}6)$ concern composite belief revisions that involve conjunctions $\phi \wedge \psi$. If a theory K is to be changed minimally so as to include two sentences ϕ and ψ, such a change should be possible by first revising K with respect to ϕ and then expanding $K * \phi$ with ψ provided that ψ does not contradict the information accepted by $K * \phi$.

$(K{*}7)$	$K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$	(Conjunction 1)
$(K{*}8)$	If $\neg\psi \notin K * \phi$, then $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$	(Conjunction 2)

Observe that when $\neg\psi \in K * \phi$, then $(K * \phi) + \psi = K_\perp$.

We now discuss some consequences of the full set of postulates $(K{*}1)$–$(K{*}8)$. First of all, some of the basic postulates become derived ones in the presence of $(K{*}7)$, $(K{*}8)$. For example, assuming that revisions with a tautology are trivial, that is: $K * \top = K$, postulates $(K{*}3)$ and $(K{*}4)$ turn out to be special cases of $(K{*}7)$ and $(K{*}8)$, respectively. To see this, take ϕ to be $\top$ in $(K{*}7)$. Then, in the presence of $(K{*}6)$ and $(K{*}7)$ we get

$$K * \psi = K * (\top \wedge \psi) \subseteq (K * \top) + \psi = K + \psi,$$

which proves $(K{*}3)$. We leave it to the reader to deduce $(K{*}4)$ from $(K{*}8)$.

$(K{*}7)$ and $(K{*}8)$ are very powerful postulates. To see just how powerful, we derive a result about the interaction of revision and disjunction. As a first step, consider the following:

$$K * \phi = K * \psi \text{ iff } \psi \in K * \phi \text{ and } \phi \in K * \psi. \qquad (9)$$

The right-to-left implication in (9) says that if ψ is accepted in $K * \phi$, then the change needed to include ψ in K is no greater than the change needed to include ϕ. Observe that $(K{*}6)$ is a consequence of (9).

Claim 3.1 *The AGM postulates for revision imply* (9).

Proof. Given the AGM postulates for revision it's trivial to prove the left to right implication in (9). For the other direction, assume $\phi \in K * \psi$, $\psi \in K * \phi$. Then

$$\begin{array}{rcl} K * \phi & = & (K * \phi) + \psi, \text{ by definition of } + \\ & = & K * (\phi \wedge \psi), \text{ by } (K{*}7) \text{ and } (K{*}8) \\ & = & K * (\psi \wedge \phi), \text{ by } (K{*}6) \\ & = & (K * \psi) + \phi \\ & = & K * \psi. \quad \dashv \end{array}$$

Next, consider the following statement:

$$(K * \phi) \cap (K * \psi) \subseteq K * (\phi \vee \psi). \qquad (10)$$

Claim 3.2 *Given* $(K{*}1)$–$(K{*}6)$*,* $(K{*}7)$ *is equivalent to* (10).

To obtain the desired characterization of the interaction between revision and disjunction we need one more intermediate result. Consider the following:

$$\neg\psi \notin K * (\phi \vee \psi) \text{ implies } K * (\phi \vee \psi) \subseteq K * \psi. \quad (11)$$

In words: if ψ is consistent with the revision of K by $\phi \vee \psi$, then an additional revision by ψ will not destroy any information.

Claim 3.3 *Given* $(K{*}1)$–$(K{*}6)$*,* $(K{*}8)$ *is equivalent to* (11).

Putting the above claims together, we arrive at the desired characterization of revision and disjunction:

Theorem 3.4 *Given the eight AGM postulates for revision we have*

$$K * (\phi \vee \psi) \quad = \quad \left\{ \begin{array}{ll} K * \phi, & \textit{or} \\ K * \psi, & \textit{or} \\ (K * \phi) \cap (K * \psi). & \end{array} \right. \quad (12)$$

In fact, in the presence of the basic postulates $(K{*}1)$–$(K{*}6)$*, the conjunction of* $(K{*}7)$ *and* $(K{*}8)$ *is equivalent to* (12).

Proof. To prove the left to right implication, we distinguish a number of cases. Assume first that $\neg\psi \in K * (\phi \vee \psi)$; then $\phi \in K * (\phi \vee \psi)$, and so $K * (\phi \vee \psi) = K * \phi$, by (9). Likewise, if $\neg\phi \in K * (\phi \vee \psi)$, then $K * (\phi \vee \psi) = K * \psi$. The third possibility is that $\neg\phi, \neg\psi \notin K * (\phi \vee \psi)$. Then, by (11), $K * (\phi \vee \psi) \subseteq K * \phi \cap K * \psi$. The inverse inclusion follows from (10).

For the right to left implication, observe that (12) implies (10), which in turn implies $(K{*}7)$. To derive $(K{*}8)$, observe that from (12) and $(K{*}6)$ we get that

$K * \phi$ equals one of $K * (\phi \wedge \psi)$, $K * (\phi \wedge \neg\psi)$, or $K * (\phi \wedge \psi) \cap K * (\phi \wedge \neg\psi)$. If $K * \phi = K * (\phi \wedge \psi)$, then

$$(K * \phi) + \psi = (K * (\phi \wedge \psi)) + \psi = K * (\phi \wedge \psi).$$

If $K * \phi = K * (\phi \wedge \neg\psi)$, then

$$(K * \phi) + \psi = (K * (\phi \wedge \neg\psi)) + \psi = K_\perp.$$

And if $K * \phi = K * (\phi \wedge \psi) \cap K * (\phi \wedge \neg\psi)$, then

$$\begin{aligned}
&(K * \phi) + \psi \\
&= (K * (\phi \wedge \psi) \cap K * (\phi \wedge \neg\psi)) + \psi \\
&= (K * (\phi \wedge \psi)) + \psi \cap (K * (\phi \wedge \neg\psi)) + \psi \\
&= K * (\phi \wedge \psi) \cap K_\perp \\
&= K * (\phi \wedge \psi).
\end{aligned}$$

In all three cases, we find that if $\neg\psi \notin K * \phi$, then $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$. $\dashv$

Postulates for Contraction

We will now introduce the AGM postulates for contraction; recall that the contraction of a theory K by a sentence ϕ is meant to result in a theory that no longer contains ϕ, whenever this is possible.

To begin with, we require logical closure for contraction:

$(K{-}1)$ $K - \phi$ is a belief set for every sentence ϕ and belief set K. (Closure)

Another easy postulate states that contraction is a reduction:

$(K{-}2)$ $K - \phi \subseteq K$. (Inclusion)

Contraction does not change anything when the formula to be contracted is not contained in the current belief set:

$(K{-}3)$ If $\phi \notin K$, then $K - \phi = K$ (Vacuity)

Furthermore, a contraction with a formula ϕ is always successful whenever ϕ is not tautological:

$(K{-}4)$ If $\nvdash_{\mathbf{L}} \phi$, then $\phi \notin K - \phi$. (Success)

Note that the converse is always valid by definition of a belief set.

These first four 'easy' postulates already fix some interaction between expansion and contraction:

$$(K - \phi) + \phi \subseteq K, \text{ whenever } \phi \in K. \qquad (13)$$

The following postulate says that the inclusion relation in (13) can be replaced by equality:

$(K{-}5)$ $K \subseteq (K - \phi) + \phi$, whenever $\phi \in K$. (Recovery)

In other words, if one believes that ϕ then consecutively contracting and adding it again will bring the believer in the same state of belief. The vacuity condition $(K{-}3)$ goes some way towards expressing the minimality idea, but the idea is only fully expressed by the recovery postulate $(K{-}5)$: enough must be left of the original theory K so as to enable us to restore it after a contraction.

As for revision, we postulate that contractions of equivalent formulas lead to the same belief sets.

$(K{-}6)$ If $\vdash_{\mathbf{L}} \phi \leftrightarrow \psi$ then $K - \phi = K - \psi$. (Extensionality)

Intuitively, one may also expect that contractions of stronger formulas lead to smaller belief sets: if $\vdash_{\mathbf{L}} \phi \rightarrow \psi$, then $K - \psi \subseteq K - \phi$. However, such a strengthening of $(K{-}6)$ is not valid in general.

Example 3.5 Assume that Jan believes that both Maarten and Bert (a logician) are Dutch, and that Jan also believes that if Bert is a Dutch logician then he is unemployed.[1] Clearly, Jan also believes that Bert is unemployed. If Jan has to give up his belief about Maarten being Dutch (p), then Jan will still believe that Bert is unemployed. However, if Jan removes his belief that both Maarten and Bert are Dutch ($p \wedge q$) then Jan will also give up his belief about the unemployed status of Bert. Jan is no longer certain about Bert being Dutch, which he used to infer that Bert is unemployed (r).[2] Formally, let $K = \mathrm{Cn}(\{p, q, q \rightarrow r\})$. Clearly, $r \in K$, and we also expect $r \in K - p$, but $r \notin K - (p \wedge r)$.

For similar reasons the principle of monotonicity does not hold for contraction:

$(K{-}M)$ If $H \subseteq K$, then $H - \phi \subseteq K - \phi$. (Monotonicity)

This failure of monotonicity is a serious structural weakness of the contraction operation. For a start, it indicates that providing suitable definitions for contraction is much harder than for expansion. The following postulate on contraction of conjunctions compensates the non-monotonicity a little:

$(K{-}7)$ $(K - \phi \cap K - \psi) \subseteq K - (\phi \wedge \psi)$. (Conjunction 1)

The intuitive invalidity of $K - \phi \subseteq K - (\phi \wedge \psi)$ has been illustrated by the Maarten-Bert example above. $(K{-}7)$ says that it at least holds for the largest belief set which is contained by the belief sets which results from contracting both conjuncts separately.

Obviously the converse of $(K{-}7)$ cannot be valid. We are only allowed to conclude that $K - (\phi \wedge \psi)$ is a subset of $K - \phi$ if ϕ is not contained in $K - (\phi \wedge \psi)$. This principle is the eighth postulate of contraction.

$(K{-}8)$ If $\phi \notin K - (\phi \wedge \psi)$ then $K - (\phi \wedge \psi) \subseteq K - \phi$. (Conjunction 2)

Clearly, $(K{-}8)$ implies

$$K - (\phi \wedge \psi) \subseteq K - \phi \text{ or } K - (\phi \wedge \psi) \subseteq K - \psi. \qquad (14)$$

In fact, the somewhat intransparent last two postulates are equivalent with three disjunctive cases which make the intuition about their nature much clearer.

Theorem 3.6 *Under the assumption that the postulates* $(K{-}1)$–$(K{-}6)$ *hold for the contraction function* $-$, *the additional principles* $(K{-}7)$ *and* $(K{-}8)$ *hold*

[1] Assume that these are the 'only' beliefs of Jan.

[2] What Jan still believes is that Maarten or Bert is Dutch.

iff

$$K-(\phi\wedge\psi) \quad = \quad \begin{cases} K-\phi, & or \\ K-\psi, & or \\ (K-\phi)\cap(K-\psi). \end{cases} \tag{15}$$

Proof. We only prove the left-to-right implication. So, suppose $(K-1)$–$(K-8)$ hold. Let

$$K-(\phi\wedge\psi)\neq K-\phi \text{ and } K-(\phi\wedge\psi)\neq K-\psi. \tag{16}$$

We need to show that $K-(\phi\wedge\psi)=K-\phi\cap K-\psi$. By $(K-7)$ we only need to prove that

$$K-(\phi\wedge\psi)\subseteq K-\phi \text{ and } K-(\phi\wedge\psi)\subseteq K-\psi. \tag{17}$$

Suppose that the latter conjunct does not hold. Then by (14) the former must hold. Because of this and (16) we may infer that there exists χ such that

$$\chi\in K-\phi \text{ and } \chi\notin K-(\phi\wedge\psi). \tag{18}$$

By $K-(\phi\wedge\psi)\not\subseteq K-\psi$ and $(K-8)$ we have $\psi\in K-(\phi\wedge\psi)$, and also $\psi\in K$ $(K-2)$. This means $(\psi\to\chi)\notin K-(\phi\wedge\psi)$. In contrast, $(\psi\to\chi)\in K-\phi$ $(\chi\vdash\psi\to\chi)$.

$(K-5)$ and $\psi\in K$ implies that $(K-\psi)+\psi=K$. In other words, $\mathrm{Cn}((K-\psi)\cup\{\psi\})=K$. Because $\chi\in K$, we know that $K-\psi\cup\{\psi\}\vdash_{\mathbf{L}}\chi$, and hence, $K-\psi\vdash_{\mathbf{L}}\psi\to\chi$. $(K-1)$ tells us that $(\psi\to\chi)\in K-\psi$, and by (18) $(\psi\to\chi)\in K-\phi\cap K-\psi$. Now, $(K-7)$ entails $(\psi\to\chi)\in K-(\phi\wedge\psi)$, and because ψ is also contained in the latter, we infer $\chi\in K-(\phi\wedge\psi)$ which contradicts our assumption. ⊣

Theorem 15 is not only a partial compensation for the monotonicity failure, but it also sums up how close we get to the so-called *fullness condition.* This condition is a very strong (in fact, too strong) criterion on the effect of contraction.

$(K-F)$ If $\phi,\psi\in K$ and $(\phi\vee\psi)\in K-\chi$, then either $\phi\in K-\chi$ or $\psi\in K-\chi$. (Fullness)

Example 3.7 The Jan-Maarten-Bert-setting immediately entails a counterexample. Say, Jan believes that Maarten and Bert are Dutch $(p\wedge q)$. After a contraction of this information, you want Jan still to believe that Maarten or Bert is Dutch $(p\vee q)$. Jan should give up his belief that Maarten is Dutch (p) and also that Bert is Dutch (q).

Formally, $p\vee q\in\mathrm{Cn}(\{p\wedge q\})-(p\wedge q)$, but $p\notin\mathrm{Cn}(\{p\wedge q\})-(p\wedge q)$ and $q\notin\mathrm{Cn}(\{p\wedge q\})-(p\wedge q)$.

Back and Forth between Revisions and Contractions

Consider Figure 2, and, especially, the dashed 'adjustment actions' shown there. This is a 'vague' part of the story we have been telling so far about revisions. Recall that preparing a belief set for revising with a proposition ϕ comes down to a removal of all things in the original belief set K which contradict ϕ. Clearly, the main culprit for a possible conflict is the presence of $\neg\phi$. So, assuming that a contraction function $-$ is available, we can implement the adjustment by contracting with $\neg\phi$, and moving to $K-\neg\phi$. This boils down to the following definition of revision on the basis of contraction.

$$K*\phi=(K-\neg\phi)+\phi. \tag{Levi Identity}$$

Which of the revision postulates come out true, whenever the given contraction function validates the postulates of the previous subsection? In fact, rather smooth correspondence results can be given.

Theorem 3.8 *If* $(K-1)$–$(K-6)$ *hold for a contraction function* $-$*, then the revision function which is defined by the Levi identity satisfies* $(K*1)$–$(K*6)$.

If one assumes the basic postulates for contraction, then the Levi identity entails that $(K-7)$ implies $(K*7)$ and $(K-8)$ implies $(K*8)$.

Theorem 3.9 *If* $(K-1)$–$(K-6)$ *hold for a contraction function* $-$*, then the revision function which is defined by the Levi identity and satisfies* $(K-7)$*, also fulfills* $(K*7)$*. Likewise, if* $-$ *satisfies* $(K-8)$ *then the revision function defined by the Levi identity satisfies* $(K*8)$.

Going the other way around, we can also try to give a definition of contraction in terms of revision. The so-called Harper identity tells us that contraction with a formula ϕ should be equal to the result of removing everything from the initial belief set which would not remain if we revised it with the negation of ϕ.

$$K-\phi=K\cap(K*\neg\phi). \tag{Harper Identity}$$

Combining the Levi and Harper identity yields:

$$\begin{aligned} K-\phi &= K\cap((K-\phi)+\neg\phi) \\ K*\phi &= (K\cap(K*\phi))+\phi \end{aligned}$$

The AGM postulates for $-$ and $*$ are strong enough to deduce these equalities (without using the Levi and Harper identities), and this fact gives additional support to the Harper and Levi identities; it also seems to hint at a kind of conservativity of the two identities over the AGM postulates.

Theorem 3.10 *If* $(K*1)$–$(K*6)$ *hold for a revision function* $*$*, then the contraction which is defined by the Harper identity satisfies* $(K-1)$–$(K-6)$.

Furthermore, given the Harper identity, we have the following implications:

$$\begin{aligned} (K*1)\text{–}(K*6) \textit{ and } (K*7) &\implies (K-7) \\ (K*1)\text{–}(K*6) \textit{ and } (K*8) &\implies (K-8) \end{aligned}$$

Notes

The postulates for revision and contraction discussed here were originally proposed in the early 1980s in work by Alchourrón, Gärdenfors and Makinson; see

[1, 10]. Further discussions on the connections between postulates for revision and postulates for contraction may be found in [25].

4 Syntax-Based Contraction Functions

In Section 3 we considered laws that any reasonable candidate revision or contraction function should satisfy according to the AGM theory. In the present section we will give explicit constructions of functions satisfying some or all of those postulates; the functions to be defined below are syntax-based in that they are defined in terms of sets of formulas. We will concentrate on constructing contraction functions — using the Levi identity suitable revision functions can be defined on top of those.

We focus on two influential proposals: one using so-called meet functions, and another using so-called epistemic entrenchment relations. In the next section we will see an example of a model-based definition of revision functions.

Meet Functions

To motivate the introduction of contraction functions based on meet functions, observe the following. Let a belief set K and a formula ϕ be given. To compute $K-\phi$, the result of contracting K with ϕ, we are interested in the *maximal* parts of K that don't imply ϕ (this is because of the principle that we should change as little as possible of our old theory). As there may be several such parts, we should somehow make a selection, and take their intersection — this is the main idea behind meet functions.

We define meet functions in a number of steps, starting with so-called remainders.

Definition 4.1 Let K be a belief set, H a set of formulas, and ϕ a formula. H is called a *maximal subset* of K that *fails to imply* ϕ if the following conditions are met:

- $H \subseteq K$
- $H \nvdash \phi$
- for all G such that $H \subsetneq G \subseteq K$, $G \vdash \phi$.

We will sometimes use *remainders of* K *after removing* ϕ to refer to maximal subsets that fail to imply ϕ. We use $K - \phi$ to denote the set or remainders of K after removing ϕ.

Lemma 4.2 *Let H be a remainder of K after removing ϕ. Then* $\mathrm{Cn}(H) \subseteq H$*, that is: H is a belief set.*

Proof. Let H be a remainder of K after removing ϕ. Assume $H \vdash \psi$, but $\psi \notin H$. We will derive a contradiction. First, observe that $H \cup \{\psi\} \nvdash \phi$, for otherwise $H \vdash \phi$. So $H \cup \{\psi\}$ is a proper superset of H that fails to imply ϕ — but this contradicts the maximality of H. $\dashv$

Example 4.3 Let $K = \mathrm{Cn}(\{p, q\})$, and $\phi = p \wedge q$. What are the elements of $K - \phi$? That is, what are the maximal subsets of $\mathrm{Cn}(\{p, q\})$ that fail to imply $p \wedge q$? A first guess may that $\mathrm{Cn}(\{p\})$ and $\mathrm{Cn}(\{q\})$ are in $K - \phi$, for if you have to contract with $p \wedge q$, then it suffices to contract with at least one of p and q, and keep the other. But this is not the case: if the language contains more proposition letters than just p and q, the elements of $K - \phi$ are far larger than either $\mathrm{Cn}(\{p\})$ or $\mathrm{Cn}(\{q\})$:

Claim. *Let $H \in K - \phi$. Then, for every formula ψ, either $(\phi \vee \psi) \in H$ or $(\phi \vee \neg\psi) \in H$.*

Proof. Observe first that as $\phi \in K$, we have $(\phi \vee \psi) \in K$ and $(\phi \vee \neg\psi) \in K$. Assume $(\phi \vee \psi) \notin H$, $(\phi \vee \neg\psi) \notin H$. As $H \in K - \phi$, it follows that

$$\begin{aligned} H \cup \{\phi \vee \psi\} &\vdash \phi \\ H \cup \{\phi \vee \neg\psi\} &\vdash \phi. \end{aligned}$$

So $H \vdash (\psi \to \phi) \wedge (\neg\psi \to \phi)$. From this and propositional logic it follows that $H \vdash \phi$, which is a contradiction as $H \in K - \phi$. $\dashv$

From the above claim it follows that, in languages with more proposition letters than just p and q, $\mathrm{Cn}(\{p\})$ and $\mathrm{Cn}(\{q\})$ are too small to be in $K - (p \wedge q)$. For if r is any proposition letter different from p (and q), we have that $(p \wedge q) \vee r$, $(p \wedge q) \vee \neg r \notin \mathrm{Cn}(\{p\})$, as neither formula is a consequence of p.

Observe that our first guess is incomplete even in languages with just the two proposition letters p and q. We leave it to the reader to check that in this case, the elements of $K - (p \wedge q)$ are $\mathrm{Cn}(\{p\})$, $\mathrm{Cn}(\{q\})$, as well as $\mathrm{Cn}(\{p \leftrightarrow q\})$.

Next we need to define selection functions.

Definition 4.4 A *selection function* for a theory K is any function

$$s_K : \mathcal{P}(\mathcal{P}(K)) \to \mathcal{P}(\mathcal{P}(K))$$

such that $\emptyset \subsetneq s_K(\mathcal{X}) \subseteq \mathcal{X}$ for all $\mathcal{X} \neq \emptyset$, and $s_K(\emptyset) = \{K\}$.

We are ready now to define contraction functions based on selection functions. The general scheme is the following; if $K - \phi$ is non-empty, then we define

$$K - \phi = \bigcap s_K(K - \phi) \qquad (19)$$

for some selection function s_K.

The literature contains several restrictions on selection functions; below we will discuss three important ones.

Maxi-Choice Meet Functions

The first restriction on selection functions that we consider is that $|s(\mathcal{X})| = 1$, that is: s always makes a unique selection by returning a singleton.

Definition 4.5 Let s be a selection function for K that only returns singletons. Then the contraction function $-$ defined by (19) is called a *maxi-choice contraction function.*

Let us examine some properties of maxi-choice contraction functions.

Lemma 4.6 *Any maxi-choice contraction function $-$ satisfies the six basic postulates for contraction, that is* $(K{-}1)$–$(K{-}6)$.

Proof. We only show that the fifth postulate $(K{-}5)$ is valid, leaving the others to the reader. We have to establish $K \subseteq (K - \phi) + \phi$.

Let $\psi \in K$ be arbitrary. Then $(\phi \to \psi) \in K$. Let $K' \in K - \phi$, and assume next that $(\phi \to \psi) \notin K'$. Then $(\phi \to \psi) \to \phi \in K'$ (why?). As $((\phi \to \psi) \to \phi) \to \phi$ is a tautology, it belongs to K', so we get $\phi \in K'$ — which contradicts $K' \in K - \phi$. So, $(\phi \to \psi) \in K'$. But then $(\phi \to \psi) \in (K - \phi) \subseteq (K - \phi) + \phi$, so $\psi \in (K - \phi) + \phi$. ⊣

Maxi-choice selection functions also satisfy the following fullness condition:

$(K{-}F2)$ If $\phi, \psi \in K$ and $\psi \notin K - \phi$, then $(\psi \to \phi) \in K - \phi$. (Fullness 2)

In general the postulates $(K{-}7)$ and $(K{-}8)$ don't hold for maxi-choice contraction functions. In fact, one can show that maxi-choice selection functions are characterized by the six basic postulates together with the second fullness condition $(K{-}F2)$:

Theorem 4.7 (AGM) *A contraction function satisfies $(K{-}1)$–$(K{-}6)$ and $(K{-}F2)$ iff it can be defined as a maxi-choice contraction function as in Definition 4.5.*

Are maxi-choice functions any good? Given that they always satisfy the second fullness condition, it does seem that they are 'too large.' This can be made more precise as follows. Call a belief set K *maximal* if for every formula ψ, either $\psi \in K$ or $\neg\psi \in K$.

As a lemma we have that if $\phi \in K$, and $-$ is a maxi-choice contraction function, then, for every formula ψ, either $(\phi \vee \psi) \in K - \phi$ or $(\phi \vee \neg\psi) \in K - \phi$ (see Example 4.3).

Theorem 4.8 *Let $*$ be a revision function that is defined from a maxi-choice function through the Levi Identity. Then, for any ϕ such that $\neg\phi \in K$, the revision of K with ϕ, $K * \phi$ is maximal.*

Proof. Let $-$ be a maxi-choice contraction function. By the previous remarks we have that $(\neg\phi \vee \psi) \in K - \neg\phi$ or $(\neg\phi \vee \neg\psi) \in K - \neg\phi$ for any formula ψ. In other words, either $(\phi \to \psi) \in K - \neg\phi$, or $(\phi \to \neg\psi) \in K - \neg\phi$. By the Levi Identity $\phi \in K * \phi = (K - \neg\phi) + \phi$. Hence, $\psi \in K * \phi$, or $\neg\psi \in K * \neg\psi$. ⊣

Full Meet Functions

As maxi-choice contraction functions return results that are too large, let us consider contraction function based on selection functions that make far larger selections, namely selection functions that return as much as possible — for then their intersection (and hence the value of the contraction function) will be as small as possible. So, we consider selection functions s such that $s(\mathcal{X}) = \mathcal{X}$ whenever $\mathcal{X} \neq \emptyset$. Let us call such selection functions *identity* selection functions.

Definition 4.9 A contraction function $-$ is called a *full meet contraction function* if it is defined as in (19) using an identity selection function.

To spell out the above definition: $-$ is a full meet contraction function iff it is defined as follows:

$$K - \phi = \begin{cases} \bigcap(K - \phi), & \text{whenever } K - \phi \neq \emptyset \\ K, & \text{otherwise.} \end{cases}$$

As the selection function on which a full meet contraction function is based selects a large collection of elements from $K - \phi$, the result of a full meet contraction will be small.

Theorem 4.10 *Let $-$ be a full meet contraction function and $\phi \in K$. Then $\psi \in K - \phi$ iff ($\psi \in K$ and $\neg\phi \vdash \psi$). In other words: $K - \phi = K \cap \mathrm{Cn}(\neg\phi)$.*

Furthermore, if $*$ is a revision function defined by the Levi identity using a full meet contraction function, then $\neg\phi \in K$ implies $K * \phi = \mathrm{Cn}(\{\phi\})$.

By way of example, if $-$ is a full meet contraction function, then, by the above result, $\mathrm{Cn}(\{p \wedge q\}) - (p \wedge q) = \mathrm{Cn}(\{p \wedge q\} \cap \mathrm{Cn}(\{\neg p \vee \neg q\}) = \mathrm{Cn}(\{\top\})$. This really shows that the result of a full meet contraction can be too small (cf. Example 4.3). One would expect $(p \vee q) \in \mathrm{Cn}(\{p \wedge q\}) - (p \wedge q)$ (cf. Example 3.7).

Proof. (of Theorem 4.10) We only prove the first half of the theorem, leaving the second half to the reader. We distinguish two cases. Assume first that $\vdash \phi$. Then $\neg\phi \vdash \psi$ and $K - \phi = K$, so $\psi \in K - \phi$ iff $\psi \in K$, and we're done.

Assume next that $\nvdash \phi$. Left us first prove the right-to-left implication: assume $\psi \in K$ and $\neg\phi \vdash \psi$. Suppose for contradiction that $\psi \notin K - \phi$. So then there exists $K' \in K - \phi$ such that $\psi \notin K'$. As $\psi \in K$, we have by the maximality of K' that $K' \cup \{\psi\} \vdash \phi$. But on the other hand, from $\neg\phi \vdash \psi$ it follows that $\neg\psi \vdash \phi$, and hence $K' \cup \{\neg\psi\} \vdash \phi$. Putting the two statements together, we find that $K' \vdash \phi$, which contradicts $K' \in K - \phi$.

As for the left-to-right implication, assume $\psi \notin K$ or $\neg\phi \nvdash \psi$. We need to show that $\psi \notin K - \phi$, that is: for some $K' \in K - \phi$, $\psi \notin K'$. Now, if $\psi \notin K$, then certainly $\psi \notin K - \phi$. If $\neg\phi \nvdash \psi$, then $\neg\psi \nvdash \phi$, and so $\phi \vee \neg\psi \nvdash \phi$. So there exists a $K' \in K - \phi$ that contains this formula $\phi \vee \neg\psi$. As $(\phi \vee \neg\psi) \wedge \psi \vdash \phi$ it follows that $\psi \notin K'$. ⊣

In addition to the six basic postulates $(K{-}1)$–$(K{-}6)$ for contraction, full meet contraction functions also satisfy the following *intersection* condition:

$(K{-}I)$ For all ϕ, $\psi \in K$, $K - (\phi \wedge \psi) = (K - \phi) \cap (K - \psi)$.

The restriction to ϕ, $\psi \in K$ in the intersection condition $(K{-}I)$ is necessary: if we were to drop the restriction, $(K{-}I)$ would no longer hold for full meet contraction.

Observe that $(K{-}7)$ is an immediate consequence of $(K{-}I)$: it's simply the right-to-left inclusion; recall also that we argued in Section 3 that the left-to-right inclusion in $(K{-}I)$ is intuitively invalid.

Theorem 4.11 *A contraction function satisfies the basic postulates* $(K-1)$–$(K-6)$ *as well as* $(K-I)$ *iff it can be defined as a full meet contraction function.*

Partial Meet Functions

Now that maxi-choice contraction functions proved to be too large, and full meet contractions too small, we will try out a third option in between those two choices, namely partial meet functions.

Definition 4.12 Let K be a belief set. A contraction function $-$ is a *partial meet contraction function* over K if there is a selection function s_K such that

$$K - \phi = \bigcap s_K(K - \phi).$$

Thus, instead of taking just one element in $K - \phi$ (as with maxi-choice contraction functions) or taking all elements in $K - \phi$ (as with full meet contraction functions), we take an arbitrary selection of elements in $K - \phi$ and define its intersection to be the result of a contraction.[3]

Theorem 4.13 *A contraction function* $-$ *can be defined as a partial meet contraction function iff it satisfies* $(K-1)$–$(K-6)$.

Proof. The left-to-right direction is left to the reader. Here's a sketch for the right-to-left direction. Let $-$ be a contraction function satisfying $(K-1)$–$(K-6)$. Define a *canonical selection function* s_K by putting

$$s_K(K - \phi) = \begin{cases} \{K' \in K - \phi \mid K - \phi \subseteq K'\}, & \text{if } K - \phi \neq \emptyset \\ \{K\}, & \text{otherwise.} \end{cases}$$

We need to show that

(1) s_K is well-defined (that is: $K - \phi = K - \psi$ implies $s_K(K - \phi) = s_K(K - \psi)$)
(2) $s_K(K - \phi) = \{K\}$ whenever $K - \phi = \emptyset$
(3) $s_K(K - \phi) \subseteq K - \phi$ if $K - \phi \neq \emptyset$
(4) $K - \phi = \bigcap s_K(K - \phi)$.

To establish these claims, use the postulates and properties of $-$. ⊣

Now that we have a contraction function that satisfies the six basic postulates for contraction, the obvious next question is: how do we get the seventh and eighth postulate? It turns out that we need to make one more addition to our apparatus, namely we have to add a preference relation $\ll_K$ over subsets of the theory K that is being contracted. Intuitively, $\ll_K$ orders the parts of K according to their relative importance for the theory as a whole. So, if $X \ll_K Y$ then Y is 'preferred to' or 'at least as good as' X from the point of view of K. Such preference relations will help us make more refined selection functions.

We will say that a function s is a *selection function generated by a relation* just in case

1. s is a selection function
2. there is a reflexive relation $\ll_K$ on the collection $\bigcup\{K - \phi \mid \not\vdash \phi\}$ such that

$$s(\mathcal{X}) = \{X \in \mathcal{X} \mid Y \ll_K X \text{ for all } Y \in \mathcal{X}\}.$$

So, a selection function generated by a relation selects those subsets that are most preferred or dominating according to the relation on which it is based.

To define a contraction function using the above machinery, let K be a belief set. Then, a contraction function $-$ is called a *(transitively) relational partial meet contraction function* if there exists a selection function s that is generated by a (transitive) relation on subsets of K such that for all ϕ,

$$K - \phi = \bigcap s(K - \phi).$$

In words, a transitively relational partial meet contraction function returns the intersection of the most preferred elements of $K - \phi$.

Theorem 4.14 *Let* $-$ *be a contraction function, and let* K *be a belief set. Then* $-$ *is a transitively relational partial meet contraction function over* K *iff it satisfies the contraction postulates* $(K-1)$–$(K-8)$ *for* K.

Proof. The left-to-right direction is left to the reader. For the converse, let K be a belief set, and define a binary relation

$$\ll_K \subseteq \bigcup\{K - \phi \mid \not\vdash \phi\} \times \bigcup\{K - \phi \mid \not\vdash \phi\}$$

by

$X \ll_K Y$

$$\text{iff} \begin{cases} \text{either } X = Y = K \\ \text{or } \quad 1. \quad X, Y \in K - \phi, \text{ for some } \phi, \text{ and} \\ \qquad\; 2. \quad K - \phi \subseteq Y, \text{ for some } \phi, \text{ and} \\ \qquad\; 3. \quad \text{for all } \psi\text{: if } X, Y \in K - \psi \text{ and} \\ \qquad\qquad\; K - \psi \subseteq X \text{ then } K - \psi \subseteq Y. \end{cases}$$

It can be shown that given the eight postulates for contraction, the above relation $\ll_K$ gives rise to a transitively relational partial meet contraction function that coincides with $-$. ⊣

By the above theorem we finally have an explicit construction of a contraction function that satisfies all of the AGM postulates for contraction. We should point out, however, that there are various shortcomings to partial meet contraction functions as a model for contraction. For a start, the computational costs of such functions are high because of the need to determine maximal subsets of a belief set. Also, in the final stage of the construction we had to employ a preference relation on subsets of belief sets — why not take such relations seriously and give them a first-class status?

Epistemic Entrenchment

In this subsection we employ a preference relation on individual formulas to construct contraction functions.

[3] For instance, if $K = \mathrm{Cn}(\{p \wedge q\})$ then $s_K(K \perp (p \wedge q)) = \{\mathrm{Cn}(\{p\}), \mathrm{Cn}(\{q\})\}$ seems a reasonable candidate selection function (see also Example 4.3). In this case we obtain $(p \vee q) \in \mathrm{Cn}(\{p \wedge q\}) \perp (p \wedge q)$ (see also Example 3.7).

The idea is that 'while all sentences in a belief set must count as fully accepted, some are more accepted than others' (Fuhrmann [8]).

Example 4.15 Bert has a sister, Freddie. She is thinking about her next career move. In her theory about the job market and job opportunities the rule that (nearly) all Dutch post-docs in logic are unemployed is probably of greater value than the fact that some post-docs in logic (like Bert) happen to have a job that even happens to be reasonably well-paid.

The epistemic entrenchment of a sentence is its informational value within the belief set. For example, *lawlike* sentences tend to have a greater informational value than mere observations. There may be many sources for ranking the sentences in a theory, ranging from information theory to the philosophy of science — we won't pursue these issues here.

How are we to use a preference relation on formulas to define a contractions or revisions? If we have to give up sentences from a belief set K, then we give up those sentences that are least important to us according to the preference relation:

$\phi \leq_K \psi$ iff ψ is at least as important or entrenched as ϕ from the point of view of K.

Definition 4.16 Let K be a belief set. A binary relation on formulas $\leq_K$ is an *epistemic entrenchment* relation if it satisfies the postulates $(EE1)$–$(EE5)$ below.

$(EE1)$ If $\phi \leq_K \psi$ and $\psi \leq_K \chi$, then $\phi \leq_K \chi$ (Transitivity)

$(EE2)$ If $\phi \vdash \psi$, then $\phi \leq_K \psi$ (Dominance)

The second postulate for entrenchment says that if either ϕ or ψ has to be given up, then giving up ϕ is a smaller change than giving up ψ, for if ψ is given up, so should ϕ because of $\phi \vdash \psi$.

$(EE3)$ $\phi \leq_K (\phi \wedge \psi)$ or $\psi \leq_K (\phi \wedge \psi)$.

One slightly inaccurate reading of $(EE3)$ is: if you have to give up the conjunction $\phi \wedge \psi$, then it suffices to give up just one of ϕ and ψ.

If a sentence is not contained in K, then it isn't entrenched in K. Hence it should be minimal in the entrenchment relation.

$(EE4)$ $\phi \notin K$ iff for all ψ: $\phi \leq_K \psi$ (Minimality)

The other way around: the sentences that are most entrenched are the logical laws.

$(EE5)$ If $\phi \leq_K \psi$ for all ϕ, then $\vdash \psi$. (Maximality)

Although the entrenchment postulates may seem quite natural, they have a number of counter-intuitive consequences, the most striking being the following.

Proposition 4.17 *Let $\leq$ be a relation between formulas that satisfies the first three entrenchment postulates $(EE1)$–$(EE3)$. Then, for all formulas ϕ and ψ: $\phi \leq \psi$ or $\psi \leq \phi$.*

Proof. By $(EE2)$ $\phi \wedge \psi \leq \phi$ and $\phi \wedge \psi \leq \psi$. By $(EE3)$ $\phi \leq \phi \wedge \psi$ or $\psi \leq \phi \wedge \psi$. Putting this together and using $(EE1)$, we find $\phi \leq \psi$ or $\psi \leq \phi$. ⊣

As a consequence, epistemic entrenchment relations decide on the relative order of any two sentences, even though they may be totally unrelated, such as 'the moon circles around the earth' and 'most post-docs in logic are unemployed.'

We are ready now to look at the connection between contraction functions and entrenchment relations.

From entrenchment to contraction. Let $\leq_K$ be an epistemic entrenchment relation. We define a contraction $-$ as follows:

$\psi \in K - \phi$ iff

$$\psi \in K \text{ and } \begin{cases} \phi <_K (\phi \vee \psi), & \text{or} \\ \phi \notin K, & \text{or} \\ \top \leq_K \phi. & \end{cases} \tag{20}$$

The intuition here is that if ϕ is not provable, then $K - \phi$ consists of members of K that are more entrenched than ϕ. However, for technical reasons one has to use $\phi <_K (\phi \vee \psi)$ instead of $\phi <_K \psi$.

From contraction to entrenchment. Let $-$ be a contraction function. An obvious first guess at a definition of entrenchment in terms of $-$ would be to put $\phi \leq_K \psi$ iff $\phi \notin K - (\phi \wedge \psi)$. In other words: ψ is preferred to ϕ if whenever we have to choose between the two (i.e., against $\phi \wedge \psi$), we choose against ϕ. However, without some restrictions, this will lead to conflicts: if ψ is a tautology, then certainly $\psi \in K - \psi$, so the above definition would yield $\neg(\psi \leq_K \psi)$; on the other hand, entrenchment relations are required to be reflexive — a contradiction.

We therefore define a relation $\leq_K$ on formulas as follows:

$$\phi \leq_K \psi \quad \text{iff} \quad \begin{cases} \vdash \psi, & \text{or} \\ \phi \notin K - (\phi \wedge \psi), & \text{if} \nvdash \psi. \end{cases} \tag{21}$$

The above two constructions are neatly linked up by the following representation results.

Theorem 4.18 *1. If $\leq_K$ is an epistemic entrenchment relation satisfying $(EE1)$–$(EE5)$, then the contraction function defined by (20) satisfies the AGM postulates for contraction $(K-1)$–$(K-8)$.*

2. If $-$ is a contraction function satisfying $(K-1)$–$(K-8)$, then (21) defines an epistemic entrenchment relation satisfying $(EE1)$–$(EE5)$.

In a picture:

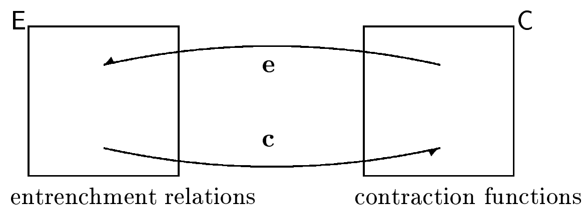


Figure 3. Back and forth between entrenchment relations and contraction functions

In fact, the result is that if **c** denotes the function taking entrenchment relations to contraction functions, and **e** is the function taking contraction functions to entrenchment relations, then

$$\mathbf{e} \circ \mathbf{c} = \mathrm{id}_{\mathsf{E}} \quad \text{and} \quad \mathbf{c} \circ \mathbf{e} = \mathrm{id}_{\mathsf{C}}. \qquad (22)$$

That is: the composition of **e** and **c** is the identity on the class of entrenchment relations, and the composition of **c** and **e** is the identity on the class of contraction functions.

Notes

In the following section we will see an example of a *model-based* definition of revision functions (in the setting of revisions for belief *bases* instead of belief sets). Besides these, the meet function constructions, and the epistemic entrenchment relations there are a number of alternative ways of constructing explicit contraction and revision functions.

The method of *safe contraction* due to Alchourrón and Makinson is the mirror image of the meet functions approach. Rather than selecting certain maximal sets that fail to imply the formula to be contracted, safe contractions prune the minimal subsets of a theory that imply the sentence to be contracted.

There are various *quantitative approaches* to contraction and revision. One influential one involves so-called Shackle measures whose qualitative part can be shown to coincide with epistemic entrenchment relations. An alternative quantitative approach is advocated by Spohn who uses ordinal conditional functions to represent revisions.

In Part II of these notes we will examine constructions of contraction and revision operations based on non-monotonic logic, verisimilitude, and modal and dynamic logic.

5 Variations and Extensions

In this section we will look at two variations of the AGM framework discussed in the previous sections. First, we move to knowledge *bases* that need not be deductively closed instead of belief sets as the basic objects of theory change. Another change to the basic format occurs when we move to multiple contractions, where the new information to be incorporated in our theory is a *set* of formulas rather than a single formula.

Base Revisions

When we consider computer-based knowledge bases, we need to fix a formalism and a *finite* representation of our knowledge base. Katsuno and Mendelzon [18] propose a format in which a knowledge base is represented by a single propositional formula ψ, and in which a *revision* is represented as a connective $\circ$: $\psi \circ \mu$ denotes the revision of ψ by μ.

Katsuno and Mendelzon propose the following collection of postulates to govern revisions on (finite) knowledge bases:

(R1) $\psi \circ \mu$ implies μ
(R2) If $\psi \wedge \mu$ is satisfiable, then $(\psi \circ \mu) \leftrightarrow (\psi \wedge \mu)$
(R3) If μ is satisfiable, then $\psi \circ \mu$ is also satisfiable
(R4) If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$, then $(\psi_1 \circ \mu_1) \leftrightarrow (\psi_2 \circ \mu_2)$

New knowledge (μ) is retained in the updated knowledge base (R1). The actual presentation of the information is irrelevant (R4). The obvious path is taken when there's no conflict (R2), and revision doesn't introduce unwarranted inconsistency (R3).

To relate the Katsuno-Mendelzon postulates for revision to the AGM postulates for revision, we restrict attention to belief sets that are generated by a single formula: $K = \{\phi \mid \psi \vdash \phi\}$. Then, it is easy to show that $(R1)$–$(R4)$ are equivalent to $(K{*}2)$–$(K{*}6)$.

What about counterparts to $(K{*}7)$ and $(K{*}8)$? It turns out that (for finitely generated belief sets) these are equivalent to the following Katsuno-Mendelzon postulates:

(R5) $(\psi \circ \mu) \wedge \phi$ implies $\psi \circ (\mu \wedge \phi)$
(R6) If $(\psi \circ \mu) \wedge \phi$ is satisfiable, then $\psi \circ (\mu \wedge \phi)$ implies $(\psi \circ \mu) \wedge \phi$.

What do these postulates mean? Consider the collection of all models of the knowledge base ψ, $\mathsf{Mod}(\psi)$. Suppose that there is some metric for measuring the 'distance' between $\mathsf{Mod}(\psi)$ and any interpretation I for the language of ψ. In line with the minimal change requirements of earlier sections, we want the models of $\psi \circ \mu$ to be the models of μ that are closest to $\mathsf{Mod}(\psi)$.

Rule (R5) says that closeness is well-behaved: if we pick an interpretation I which is closest to $\mathsf{Mod}(\psi)$ in a certain set, namely in $\mathsf{Mod}(\mu)$, and I also belongs to a smaller set, $\mathsf{Mod}(\mu \wedge \phi)$, then I must also be closest to $\mathsf{Mod}(\psi)$ within the smaller set $\mathsf{Mod}(\mu \wedge \phi)$. A counterexample to (R6) would be an interpretation I that is closer to the knowledge base than J within a certain set, while J is closer than I within some other set.

To make sense of these ideas about closeness, consider the following. Let $\mathcal{I}$ be the set of all interpretations of our language, and consider a function that assigns to each propositional formula ψ a pre-order $\leq_\psi$ over $\mathcal{I}$. This assignment is called *faithful* if the following hold:

1. If I, $I' \in \mathsf{Mod}(\psi)$, then $I <_\psi I'$ does not hold
2. If $I \in \mathsf{Mod}(\psi)$ and $I' \notin \mathsf{Mod}(\psi)$, then $I <_\psi I'$ holds
3. If $\vdash \phi \leftrightarrow \phi$, then $\leq_\psi = \leq_\phi$.

Let M be a subset of $\mathcal{I}$. An interpretation I is called *minimal* in M with respect to $\leq_\psi$ if $I \in \mathsf{M}$ and there is no $I' \in \mathsf{M}$ such that $I' <_\psi I$. Define

$$\mathrm{Min}(\mathsf{M}, \leq_\psi) = \{I \mid \ I \text{ is minimal in } \mathsf{M} \text{ w.r.t. } \leq_\psi\}.$$

Intuitively, $I' \leq_\psi I$ means that I' is closer to $\mathsf{Mod}(\psi)$ than I; and $\mathrm{Min}(\mathsf{M}, \leq_\psi)$ picks out the interpretations in M that are closest to $\mathsf{Mod}(\psi)$.

Theorem 5.1 *A revision operator* $\circ$ *satisfies* (R1)–(R6) *iff there exists a faithful assignment that maps each knowledge base to a total pre-order* $\leq_\psi$ *such that* $\mathsf{Mod}(\psi \circ \mu) = \mathrm{Min}(\mathsf{Mod}(\mu), \leq_\psi)$.

Example 5.2 (Dalal's revision) Dalal measures the distance $dist(I, J)$ between two interpretations I and J by counting the number of proposition letters on which they disagree. The distance between $\mathsf{Mod}(\psi)$ and I is defined as

$$dist(\mathsf{Mod}(\psi), I) = \min_{J \in \mathsf{Mod}(\psi)} dist(J, I).$$

Now define a faithful assignment of a total pre-order $\leq_\psi$ by putting $I \leq_\psi J$ iff

$$dist(\mathsf{Mod}(\psi), I) \leq dist(\mathsf{Mod}(\psi), J).$$

One can then define a revision operator $\circ_D$ as in Theorem 5.1.

Assume that our language has just 4 proposition letters p, q, r, s, and that interpretations are represented as boolean vectors of length 4. Consider the following interpretations

$$\begin{array}{lcllcl} I_1 & = & (1,1,1,1), & I_2 & = & (0,0,0,0) \\ J_1 & = & (0,0,1,1), & J_2 & = & (1,0,0,0), \\ J_3 & = & (0,0,1,0). & & & \end{array}$$

Let $\psi = Th(I_1, I_2)$, $\phi_1 = Th(J_1, J_2, J_3)$, $\phi_2 = Th(J_1, J_2)$, and $\phi_3 = Th(J_1, J_3)$. Here $Th(\mathsf{M})$ is a formula whose set of models is exactly M. Then

$$\begin{array}{lcl} \psi \circ_D \phi_1 & \leftrightarrow & Th(J_2, J_3) \\ \psi \circ_D \phi_2 & \leftrightarrow & Th(J_2) \\ \psi \circ_D \phi_3 & \leftrightarrow & Th(J_3). \end{array}$$

Example 5.3 (Borgida's revision) Borgida's revision operator $\circ_B$ is based on comparing sets of proposition letters on which a model of ψ and a model of μ differ. For two interpretations I, J, we write $\mathit{Diff}(I, J)$ to denote the set of proposition letters whose interpretation is different in I and J. $\mathit{Diff}(I, \mu)$ is

$$\bigcup_{J \in \mathsf{Mod}(\mu)} \mathit{Diff}(I, J).$$

Borgida's revision is defined as follows. If μ is inconsistent with ψ then an interpretation J is a model of $\psi \circ_B \mu$ iff J is a model of μ and there is some model I of ψ such that $\mathit{Diff}(I, J)$ is a minimal element of $\mathit{Diff}(I, u)$. Otherwise, if μ is consistent with ψ, then $\psi \circ_B \mu$ is defined as $\psi \wedge \mu$.

Consider the previous example, and add the following two interpretations

$$J_4 = (1,1,0,0), J_5 = (1,1,1,0).$$

Let $\phi_4 := Th(J_2, J_4)$, $\phi_5 := Th(J_4, J_5)$, and $\phi_6 := Th(J_2, J_4, J_5)$.

Now, suppose for a contradiction that there is a faithful assignment of a pre-order $\leq_\psi$ that captures the $\circ_B$ operator. Then $J_2 \not< J_4$ because $(\psi \circ_B \phi_4) \leftrightarrow Th(J_2, J_4)$. Further, $J_5 \not< J_4$ follows from $(\psi \circ_B \phi_5) \leftrightarrow Th(J_4, J_5)$. On the other hand, either $J_2 <_\psi J_4$ or $J_5 <_\psi J_4$ follows from $(\psi \circ_B \phi_6) \leftrightarrow Th(J_2, J_5)$ — a contradiction.

Multiple Contractions

In this subsection we work with infinite belief sets again, just as in the original AGM paradigm, but we change one of the other parameters. We will discuss *multiple contractions*, i.e., contractions by a set of formulas rather than by a single formula. It will turn out that there are two major variants: one where all sentences must be removed from the belief set, and one where it suffices that at least one of them is removed.

Before plunging into formal details, let us briefly motivate multiple contractions. Suppose that we want to give up two different points of view, ϕ and ψ. The result of the removal should be a belief set that implies neither ϕ nor ψ, and is otherwise as similar to the old theory as is possible. The following may seem to have the same effects as removing the set $\{\phi, \psi\}$ from our theory:

1. contracting by $\phi \vee \psi$
2. intersecting the results of contracting by ϕ and contracting by ψ
3. first contract by ϕ, then by ψ (or the other way around)
4. contracting by $\phi \wedge \psi$

None of these options is quite satisfactory as an explanation of removing $\{\phi, \psi\}$. In the fourth option it suffices to just give up one of the two formulas; in the third option the problem is that the order in which single-sentence contractions are performed matters; as is shown in [14], the result of first contracting by ϕ and then by ψ need not be the same as the result of first contracting by ψ and then by ϕ. As to the first option, if we inadvertently admit $\neg\phi$ to our theory while ϕ is already part of our theory, then a common strategy is to 'suspend belief in ϕ'; that is, to remove both ϕ and $\neg\phi$ from the theory; if the first option were used, then this would be impossible: we can't remove tautologies from belief sets. Finally, the second option seems to have some credibility as an explanation of a multiple contraction with ϕ and ψ.

Fuhrmann and Hansson [9] distinguish two forms of multiple contraction: *package* contractions in which all information needs to be removed, and *choice* contractions in which it suffices to remove just some of the information. Changes in legal systems typically involve simultaneous changes in many parts of the legal code, and as such they constitute a nice example of package contractions.

To help understand the formal properties of these two forms of multiple contraction, it is useful to have two notions of derivability:

$X \vdash Y$**:** $Y \subseteq \mathrm{Cn}(X)$
$X \Vdash Y$**:** $Y \cap \mathrm{Cn}(X) \neq \emptyset$

We will write $K - [A]$ to denote the package contraction of K by the set A, and $K - \langle A \rangle$ to denote the choice contraction of K by A.

Let us consider how the basic AGM postulates for contraction $(K{-}1)$–$(K{-}6)$ have to be amended for multiple contractions.

Closure. For any belief set K and set of sentences A, both $K - [A]$ and $K - \langle A \rangle$ are deductively closed.

Success. This postulate comes in two flavors:
- Choice: if $\not\vdash A$ then $A \not\subseteq K - \langle A \rangle$.
- Package: if $\not\Vdash A$ then $A \cap (K - [A]) = \emptyset$.

Inclusion. Both $K - [A]$ and $K - \langle A \rangle$ are subsets of K.

Vacuity. This postulate too comes in two flavors:
- Choice: if $A \not\subseteq K$ then $K = K - \langle A \rangle$
- Package: if $A \cap K = \emptyset$ then $K = K - [A]$.

Extensionality. Again, a postulate in two flavors:
- Choice: if $\mathrm{Cn}(A) = \mathrm{Cn}(B)$ then $K - \langle A \rangle = K - \langle B \rangle$
- The obvious counterpart for package contraction (if $\mathrm{Cn}(A) = \mathrm{Cn}(B)$ then $K - [A] = K - [B]$) is not an acceptable principle: take $K = \mathrm{Cn}(\{p\})$, then $K - [\{p \wedge q\}] = K$ by Vacuity, but $K - [\{p \wedge q, p\}] \neq K$ by Success, yet $\mathrm{Cn}(\{p \wedge q\}) = \mathrm{Cn}(\{p \wedge q, p\})$. Instead, Fuhrmann and Hansson propose a postulate called
  Package Uniformity: If $A \equiv_K B$ then $K - [A] = K - [B]$, where $A \equiv_K B$ holds if 'A and B are equivalent-modulo-K,' that is to say: $A \equiv_K B$ if $\forall X \subseteq K\,(X \Vdash A \Leftrightarrow X \Vdash B)$.

The *Recovery* postulate $(K \subseteq (K - \phi) + \phi)$ is replaced by a *Relevance postulate* that is meant to capture the idea of *minimal* change in the same way as the Recovery postulate.

Relevance. If $\psi \in K \setminus (K - \phi)$, then there exists a set K' such that
1. $K - \phi \subseteq K' \subseteq K$
2. $K' \not\vdash \phi$
3. $K', \psi \vdash \phi$.

Given the other basic AGM postulates $(K{-}1)$, …, $(K{-}4)$ and $(K{-}6)$, the Recovery postulate is equivalent to the Relevance postulate.

Now, choice and package versions of the Relevance postulate are easily arrived at:

Choice Relevance. Take the relevance postulate for ordinary contraction and replace $K - \phi$ by $K - \langle A \rangle$, and ϕ by A.

Package Relevance. Take the relevance postulate for ordinary contraction and replace $K - \phi$ by $K - [A]$, ϕ by A, and $\vdash$ by $\Vdash$.

With these postulates to our disposal we can formulate representation results in the spirit of the previous section. To adapt the machinery of meet functions to the present setting we use two definitions of remainders:

Choice Remainders: $X \in K - A$ if
- $X \subseteq K$
- $X \not\vdash A$
- $\forall Y\,(X \subsetneq Y \subseteq K \Rightarrow Y \vdash A)$.

Package Remainders: $X \in K - A$ if
- $X \subseteq K$
- $X \not\Vdash A$
- $\forall Y\,(X \subsetneq Y \subseteq K \Rightarrow Y \Vdash A)$.

On top of this one can define choice and package partial meet functions as intersections of selections of choice and package remainders, respectively.

Theorem 5.4 *1. A binary operation on sets of formulas can be defined as a choice partial meet function if it satisfies the postulates for choice contraction.*

2. A binary operation on sets of formulas can be defined as a package partial meet function if it satisfies the postulates for package contraction.

Notes

The discussion of base revisions is taken from Katsuno and Mendelzon [18]. The material on multiple contractions is based on Fuhrmann and Hansson [9]. Further changes to some of the parameters in the basic AGM framework include:
1. Relaxing the requirement that revisions be *functional* [24].
2. Change-recording theory change, as opposed to knowledge-adding theory change [19].
3. Iterated theory change [22, 20].
4. Multi-agent theory change [15].

II–Alternative Approaches

In the second half of these notes we discuss alternative approaches to dealing with changing information; we start with non-monotonic logic and verisimilitude, and then turn to descriptive approaches based on modal and dynamic logic.

6 Incomplete Information

Non-monotonic logic is concerned with inferring information from given sentences in ways that don't satisfy the monotonicity property

$$\text{(Monotonicity)} \quad \frac{\models \phi \to \psi \qquad \psi \models \chi}{\phi \models \chi}.$$

In other words, in non-monotonic logic the addition of premises may lead to *fewer* conclusions. Here's an example. Consider the following set of premises.

Most post-docs have nice jobs
Post-docs in logic are unemployed
Mike is a post-doc

In the absence of further information, you would probably conclude that Mike has a nice job. However, if you were then to find out that Mike is actually a postdoc in logic, you would cancel your inference, but you wouldn't give up any of the premises.

In this section we consider the Kraus-Lehmann-Magidor (KLM) framework for non-monotonic reasoning. At first sight it may seem strange that revision and non-monotonic reasoning have anything to do with each other; after all, theory change deals with the *dynamics* of belief sets and with the way we adapt our state of information in the face of new data. Non-monotonic logic, on the other hand, studies *jumps to conclusions* from (often) incomplete data. Yet, we will be able to establish connections between revision and non-monotonic inference, and we will do so in two ways: syntactically, by translating postulates into inference rules and conversely, and by using the model theory developed within the KLM framework.

Many calculi perform non-monotonic inferences (circumscription, default logic, auto-epistemic logic, negation-as-failure, inheritance systems, ...). The first systematic study of non-monotonic *consequence relations* is due to Dov Gabbay. Independently, Makinson and Shoham have proposed model theories for non-monotonic inference. KLM present a unified framework in which both approaches are linked together via representation results.

The KLM framework uses an ordinary propositional language, with the usual boolean operations ($\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$). In addition there's a collection $\mathcal{I}$ of interpretations for this propositional language; this may be seen as the collection of all worlds considered possible. The language and interpretations are related by a satisfaction relation $\models$ which is assumed to behave normally for $\neg$, $\wedge$, ...

If ϕ, ψ are formulas, then $\phi \mid\sim \psi$ is called a *conditional assertion*; the intuitive reading is 'if ϕ, then normally ψ,' or 'ψ is a plausible consequence of ϕ.' We will now give a collection of axioms and inference rules that defines $\mid\sim$; after that it will be matched with a suitable model theory.

KLM: Inference Rules for Preferential Reasoning

KLM propose the following inference rules for preferential reasoning: Reflexivity, Left Logical Equivalence, Right Weakening, Cut, Cautious Monotonicity, and Or. The resulting system is called **P**.

(Reflexivity) $\phi \mid\sim \phi$

(Left Logical Equivalence) $\dfrac{\models \phi \leftrightarrow \psi \qquad \phi \mid\sim \chi}{\psi \mid\sim \chi}$

In words, logically equivalent formulas have the same plausible consequences.

The following says that one should accept as plausible consequences everything that is logically implied by things that are plausible consequences.

(Right Weakening) $\dfrac{\models \phi \rightarrow \psi \qquad \chi \mid\sim \phi}{\chi \mid\sim \psi}$

Observe that from Reflexivity and Right Weakening we get: if $\phi \models \psi$, then $\phi \mid\sim \psi$.

(Cut) $\dfrac{\phi \wedge \psi \mid\sim \chi \qquad \phi \mid\sim \psi}{\phi \mid\sim \chi}$

That is, to obtain a plausible conclusion, one can first add a hypothesis and then deduce (plausibly) the added hypothesis. To replace the full monotonicity principle from classical logic, the following 'cautious' version is introduced:

(Cautious Monotonicity) $\dfrac{\phi \mid\sim \psi \qquad \phi \mid\sim \chi}{\phi \wedge \psi \mid\sim \chi}$.

If ϕ is reason enough to believe ψ and also to believe χ, then ϕ and ψ should be reason enough to believe χ, since ϕ was enough anyway, and on this basis, ψ was expected. The final rule of preferential reasoning is

(Or) $\dfrac{\phi \mid\sim \chi \qquad \psi \mid\sim \chi}{\phi \vee \psi \mid\sim \chi}$

In words, any formula that is, separately, a plausible consequence of two different formulas, should also be a plausible consequence of their disjunction.

The following are derived rules in the system **P**.

(Equivalence) $\dfrac{\phi \mid\sim \psi \quad \psi \mid\sim \phi \quad \phi \mid\sim \chi}{\psi \mid\sim \chi}$

(And) $\dfrac{\phi \mid\sim \psi \quad \phi \mid\sim \chi}{\phi \mid\sim \psi \wedge \chi}$

(MPC) $\dfrac{\phi \mid\sim \psi \rightarrow \chi \quad \phi \mid\sim \psi}{\phi \mid\sim \chi}$

MPC is short for Modus Ponens in the Consequent; And says that the conjunction of two plausible consequences should again be plausible, and Equivalence expresses that formulas that are plausibly equivalent have the same plausible consequences.

None of the rules below should be part of a logic that claims to be non-monotonic, as they all imply Monotonicity:

(EDT) $\dfrac{\phi \mid\sim \psi \rightarrow \chi}{\phi \wedge \psi \mid\sim \chi}$

(Transitivity) $\dfrac{\phi \mid\sim \psi \quad \psi \mid\sim \chi}{\phi \mid\sim \chi}$

(Contraposition) $\dfrac{\phi \mid\sim \psi}{\neg\psi \mid\sim \neg\phi}$

EDT stands for the Easy half of the Deduction Theorem.

Lemma 6.1 *Given the rules of the system* **P***, Monotonicity, EDT and Transitivity are equivalent. Moreover, Contraposition implies Monotonicity.*

The following Hard half of the Deduction Theorem (HDT) is a consequence of **P**:

(HDT) $\dfrac{\phi \wedge \psi \mid\sim \chi}{\phi \mid\sim \psi \rightarrow \chi}$.

The idea is that deductions performed under strong assumptions may be useful even if the assumptions are not known facts.

KLM: Preferential Models

The goal of this subsection is to define models for **P**, and to show that each model gives rise to a consequence relation satisfying the rules of **P**, and, conversely, that each consequence relation satisfying those rules is defined by a model.

The models consist of a set of states (representing the possible states of affairs), and a binary relation on those states that represents the preferences a reasoner may have between states: $s \prec t$ if s is more preferred (satisfies more of our default assumptions) than t.

We need the following definitions. Given a binary relation $\prec$ and a domain S with $\prec \subseteq S \times S$, a state $t \in S'$ is *$\prec$-minimal* in $S' \subseteq S$ if for all $s \in S'$ we have $s \not\prec t$. We call t a *$\prec$-minimum of* S' if for all $s \in S'$, $s \neq t$ implies $t \prec s$. Furthermore, a subset $S' \subseteq S$ is *smooth* if every t in S' is either $\prec$-minimal in S' or there exists an $s \in S'$ with $s \prec t$ that is $\prec$-minimal in S'.

Definition 6.2 A *preferential model* M is a tuple $(S, l, \prec)$ where S is a set of states; $l : S \to \mathcal{I}$ assigns an interpretation to each state, and $\prec$ is a strict partial order (i.e., irreflexive and transitive) satisfying the following *smoothness condition*:

> for every formula ϕ, $[\![\phi]\!] = \{s \mid l(s) \models \phi\}$ is smooth.

Preferential models $M = (S, l, \prec)$ define a consequence relation $\mid\!\sim_M$ as follows:

$$\phi \mid\!\sim_M \psi \text{ iff } \mathrm{Min}([\![\phi]\!], \prec) \subseteq [\![\psi]\!].$$

Here $\mathrm{Min}(\mathsf{M}, \prec)$ is the collection of $\prec$-minimal models in M (compare Section 5!). Thus, ϕ plausibly implies ψ if ψ is true in all most preferred ϕ states.

Lemma 6.3 (Soundness) *Every preferential model M gives rise to a consequence relation $\mid\!\sim_M$ that satisfies all the rules of the system* $\mathbf{P}$.

Our next aim is to prove the converse of the above lemma. To this end the following comes in useful. Define $\phi\mathcal{O}\psi$ iff $\phi \vee \psi \mid\!\sim \psi$; intuitively, $\phi\mathcal{O}\psi$ expresses that ϕ is strong enough to (plausibly) imply ψ. Observe that $\mathcal{O}$ is reflexive (by Reflexivity and Left Logical Equivalence) and transitive (by Left Logical Equivalence and Right Weakening).

Call an interpretation I *normal* for ϕ if for any formula ψ such that $\phi \mid\!\sim \psi$ we have $I \models \psi$.

Lemma 6.4
1. *If $\phi\mathcal{O}\psi$ and I is a normal interpretation for ϕ that satisfies ψ, then I is normal for ψ as well.*
2. *If $\phi\mathcal{O}\psi\mathcal{O}\chi$, I is normal for ϕ and satisfies χ, then I is normal for ψ.*

We now define the preferential model that we need for the main representation result. Given a preferential consequence relation $\mid\!\sim$ we define $M = (S, l, \prec)$ by

- $S = \{(I, \phi) \mid I \text{ is a normal interpretation for } \phi\}$
- $l((I, \phi)) = I$
- $(I, \phi) \prec (J, \psi)$ iff ($\phi\mathcal{O}\psi$ and $I \not\models \psi$).

As Johannes Heidema pointed out, what the latter condition seems to say is that ψ should be fairly weak compared to ϕ ($\phi\mathcal{O}\psi$), but that it shouldn't be too weak — there should be interpretations that refute it (the idea being that almost every interpretation is a model for very weak formulas).

To get our representation result we need to show that M is in fact a preferential model (i.e., $\prec$ is a strict partial order, and the smoothness condition is satisfied), and that $\mid\!\sim$ coincides with $\mid\!\sim_M$. The following lemma is useful.

Lemma 6.5 *Assume that the background logic is compact. Let $\mid\!\sim$ satisfy all the rules of* $\mathbf{P}$. *Then, for any two formulas ϕ and ψ we have the following equivalence: all normal interpretations for ϕ satisfy ψ iff $\phi \mid\!\sim \psi$.*

Lemma 6.6 *M is a preferential model.*

Proof. We first show that $\prec$ is a strict partial order. It is certainly irreflexive, as $(I, \phi) \prec (I, \phi)$ would imply $I \not\models \phi$ (by definition), and $I \models \phi$ (by normality and $\phi \mid\!\sim \phi$). To prove transitivity, assume that $(I_0, \phi_0) \prec (I_1, \phi_1) \prec (I_2, \phi_2)$. Then $\phi_0\mathcal{O}\phi_1\mathcal{O}\phi_2$. As $\mathcal{O}$ is transitive, this gives $\phi_0\mathcal{O}\phi_2$. Further, I_0 is normal for ϕ_0 but $I \not\models \phi_1$. Hence by Lemma 6.4, $I \not\models \phi_2$.

Next comes smoothness. We use the following characterization of minimal sets of the form $[\![\phi]\!]$: (I, ψ) is minimal in $[\![\phi]\!]$ iff $I \models \phi$ and $\psi\mathcal{O}\phi$. Assuming this characterization, smoothness is easily established. For suppose that $(I, \psi) \in [\![\phi]\!]$. Then $I \models \phi$. If $\psi\mathcal{O}\phi$, then, by the characterization, (I, ψ) is minimal in $[\![\phi]\!]$, and we're done. On the other hand, if $\psi\mathcal{O}\phi$ does *not* hold, then $\phi \vee \psi \not\mid\!\sim \psi$. So by Lemma 6.5, there exists an interpretation J for $\phi \vee \psi$ with $J \not\models \psi$. As $(\phi \vee \psi)\mathcal{O}\psi$ we get $(J, \phi \vee \psi) \prec (I, \psi)$. From $J \models \phi \vee \psi$ and $J \not\models \psi$ we get $J \models \phi$. Since $(\phi \vee \psi)\mathcal{O}\phi$, our characterization gives us that $(J, \phi \vee \psi)$ is minimal in $[\![\phi]\!]$. ⊣

Lemma 6.7 *The two consequence relations $\mid\!\sim$ and $\mid\!\sim_M$ coincide.*

Proof. We first show that $\phi \mid\!\sim \psi$ implies $\phi \mid\!\sim_M \psi$. Assume $\phi \mid\!\sim \psi$. We need to show $\mathrm{Min}([\![\phi]\!], \prec) \subseteq [\![\psi]\!]$. Suppose (I, χ) is minimal in $[\![\phi]\!]$. Then I is normal for χ and satisfies ϕ. So by the characterization of minimality mentioned above, $\chi\mathcal{O}\phi$. Therefore, by Lemma 6.4 I is normal for ϕ, so $I \models \psi$.

Now, for the converse implication, observe that given any normal interpretation I for ϕ, (I, ϕ) is minimal in $[\![\phi]\!]$. If $\phi \mid\!\sim_M \psi$, ψ is satisfied by all normal interpretations of ϕ, and so by Lemma 6.5, $\phi \mid\!\sim \psi$. ⊣

Theorem 6.8 *A consequence relation is a preferential consequence relation (satisfying the rules of* $\mathbf{P}$*) iff it can be defined by a preferential model.*

Connecting Theory Change and Non-monotonic Logic

We now explore two ways of connecting revision and non-monotonic logic: a syntactic one based on translating inference rules and postulates, and a semantic one that uses plausible consequence to define the revision operator.

Postulates and Inference Rules

The syntactic links between revision and non-monotonic logic start from the following observation. Let K be a (deductively closed) belief set, and ϕ a formula. We can view the revision of K by ϕ as an inference from ϕ, using K as the background information on the basis of which inferences are made. Thus the set of plausible consequences of ϕ (given K) is precisely $K * \phi$: $\phi \mid\!\sim_K \psi$ iff $\psi \in K * \phi$.

Let us check right away that $\mid\!\sim_K$ as defined above is indeed a non-monotonic inference relation: there are ϕ, ψ, χ such that $\phi \vdash \psi$, $\psi \mid\!\sim_K \chi$, but $\phi \not\mid\!\sim \chi$. Take $\phi = p \wedge q$ and $K = \mathrm{Cn}(\{\neg p \vee \neg q\})$. Then $K * \phi = K * (p \wedge q) \ni (p \wedge q)$ (by Success), but $\not\ni p \wedge \neg q$ (by Consistency). Clearly $\phi \vdash p$. Yet $K * p = \mathrm{Cn}(\{\neg p \vee \neg q\} \cup \{p\}) = \mathrm{Cn}(\{p \wedge \neg q\})$. So, $\phi \vdash p$, $p \mid\!\sim p \wedge \neg q$, but $p \wedge q \not\mid\!\sim p \wedge \neg q$.

From $*$ to $\mid\!\sim$. The above suggests the following translation scheme for expressing postulates for theory change in non-monotonic logic:

$$\psi \in K * \phi \;\Rightarrow\; \phi \mid\!\sim_K \psi.$$

This scheme does not tell us how to handle expressions of the form $K + \phi$, but we can use the following trick here: if K is consistent (and closed under Cn) then $K * \top = \mathrm{Cn}(K \cup \{\top\}) = \mathrm{Cn}(K) = K$. Thus, using the fact that $\psi \in K + \phi$ iff $(\phi \to \psi) \in K$, we rewrite $\psi \in (K + \phi)$ to $(\phi \to \psi) \in K * \top$, which translates into $\top \mid\!\sim \phi \to \psi$.

Let us now translate the eight AGM postulates for revision.

(K*1) $\mathrm{Cn}(K * \phi) = K * \phi$.
Translation: $\{\psi \mid \phi \mid\!\sim_K \psi\} = \mathrm{Cn}\{\psi \mid \phi \mid\!\sim_K \psi\}$. The right to left inclusion of the translation may be proved using And and MPC.

(K*2) $\phi \in K * \phi$.
Translation: $\phi \mid\!\sim_K \phi$. This is obviously valid in **P**.

(K*3) $K * \phi \subseteq K + \phi$.
Translation: $\phi \mid\!\sim \psi$ implies $\top \mid\!\sim (\phi \to \psi)$, and this is a special case of HDT (a derived rule in **P**).

(K*4) If $\neg\phi \notin K$, then $K + \phi \subseteq K * \phi$.
Rewrite to: if $\neg\phi \notin K$ and $\psi \in \mathrm{Cn}(K \cup \{\phi\})$, then $\psi \in K * \phi$.
Translation: If $\top \not\mid\!\sim_K \neg\phi$ and $\top \mid\!\sim_K (\phi \to \psi)$, then $\phi \mid\!\sim_K \psi$. This principle is not valid on all preferential models; simply take M such that $S = \{0, 1, 2\}$, $\prec = \{(1, 2)\}$ and assign interpretations to 0, 1, and 2 so that $l(0), l(2) \models p$, $l(0) \models q$. Then $\not\mid\!\sim_M \neg p$, but $\top \vdash_M (p \to q)$, yet 2 (a minimal p-world) has $\neg q$, and so $p \not\mid\!\sim_M q$.

(K*5) $K * \phi = K_\perp$ only if $\vdash \neg\phi$.
Translation: if $- \in \{\psi \mid \phi \mid\!\sim_K \psi\}$ then $- \in \mathrm{Cn}(\{\phi\})$. Again this is not a valid principle of non-monotonic reasoning; giving a countermodel is left to the reader.

(K*6) If $\vdash \phi \leftrightarrow \psi$, then $K * \phi = K * \psi$.
Translation: $\vdash \phi \leftrightarrow \psi$ implies $\phi \mid\!\sim \chi$ iff $\psi \mid\!\sim \chi$; this is the rule of Left Logical Equivalence.

(K*7) $K * (\phi \wedge \psi) \subseteq (K * \psi) + \phi$.
Translation: $\psi \wedge \phi \mid\!\sim \chi$ implies $\psi \mid\!\sim (\phi \to \chi)$. The latter is the derived rule HDT.

(K*8) If $\neg\phi \notin K * \psi$, then $\mathrm{Cn}(K * \psi \cup \{\phi\}) \subseteq K * (\psi \wedge \phi)$.
Translation: $\psi \not\mid\!\sim_K \neg\phi$ and $\psi \mid\!\sim (\phi \to \chi)$, then $\psi \wedge \phi \mid\!\sim_K \chi$. The latter principle is not valid on all preferential models; cf. the discussion following (K*4).

From $\mid\!\sim$ to $*$. Let us now work in the opposite direction, and translate the inference rules of the system **P** into statements about revision. We have already seen that Reflexivity and Left Logical Equivalence can be obtained from the revision postulates.

(Right Weakening) If $\models \phi \to \psi$ and $\chi \mid\!\sim \phi$, then $\chi \mid\!\sim \psi$.
Translation: if $\models \phi \to \psi$ and $\phi \in K * \chi$ then $\psi \in K * \chi$. Given that belief sets are deductively closed, this is obviously valid.

(Cut) If $\phi \wedge \psi \mid\!\sim \chi$ and $\phi \mid\!\sim \psi$, then $\phi \mid\!\sim \chi$.
Translation: if $\psi \in K * \phi$ and $\chi \in K * (\phi \wedge \psi)$, then $\chi \in K * \phi$. This may be derived from the postulates (use (K*7)).

(Cautious Monotonicity) If $\phi \mid\!\sim \psi$ and $\phi \mid\!\sim \chi$, then $\phi \wedge \psi \mid\!\sim \chi$.
Translation: if $\psi \in K * \phi$ and $\chi \in K * \phi$, then $\chi \in K * (\phi \wedge \psi)$. This may be derived from the postulates (use (K*8)).

(Or) If $\phi \mid\!\sim \chi$ and $\psi \mid\!\sim \chi$, then $\phi \vee \psi \mid\!\sim \chi$.
Translation: if $\chi \in K * \phi$ and $\chi \in K * \psi$, then $\chi \in K * (\phi \vee \psi)$. To see that this is a valid principle, use the the following characterization of revisions by disjunctions: $K * (\phi \vee \psi)$ is one of $K * \phi$, $K * \psi$, or $(K * \phi) \cap (K * \psi)$ (Theorem 3.4).

The upshot of the above connections between revision and non-monotonic logic is the following. Every inference rule for preferential reasoning is a valid principle for revision, but the converse does not hold: some principles for revision are not valid for preferential reasoning (notably, (K*4), (K*5)). Thus, revision is governed by *more* principles than preferential reasoning.

A Model-Theoretic Approach

There is a close similarity between the definition of plausible inference in a preferential model ($\mid\!\sim_M$) and the model-based definition of revision presented in Section 5. To see how exactly, note first that our translation scheme ($*$ to $\mid\!\sim$) suggests the following definition of revision using preferential models.

$$K *^{\prec} \phi := \{\psi \mid \mathrm{Min}(\mathsf{Mod}(\phi), \prec) \subseteq \mathsf{Mod}(\psi)\}. \quad (23)$$

Note, however, that there is no dependence on the belief set K in the above.

If we restrict ourselves to finitely generated theories, the above definition can be compared to the model-based definition of revision given in Section 5. There each formula/theory χ was equipped with a total pre-order $\leq_\chi$ on interpretations in such a way that all models for χ are strictly less than non-models for χ, but no model for χ is strictly less than another model for χ, and equivalent formulas are associated with the same pre-order. This machinery gave rise to the following definition of revision:

$$K *^{\leq} \phi := \{\psi \mid \mathrm{Min}(\mathsf{Mod}(\phi), \leq_\chi) \subseteq \mathsf{Mod}(\psi)\}, \quad (24)$$

where χ is assumed to generate K.

Clearly, definition (24) is more restrictive than (23): there are more conditions on the relation $\leq_\chi$, which is, moreover, dependent on the theory being revised. This is reflected in the following. We showed, in Section 5, that revisions defined by (24) satisfy the AGM postulates for revision. On the other hand, not all of the AGM postulates for revision are satisfied by definition (23); proving this is left to the reader. But this is to be expected in the light of our earlier results of intertranslating postulates for $*$ and rules for $\mid\!\sim$: we found that $*$ was the more restricted notion of the two that was governed by more laws than $\mid\!\sim$.

Ryan and Schobbens [32] propose adapting the set-up of preference relations and preferential models in the following manner to arrive at a better match.

As before, let $\mathcal{I}$ denote a collection of interpretations for the language, and let $\mathcal{T}$ be the set of deductively closed sets of sentences over our background language. An *RS preference relation* $\sqsubseteq$ is a ternary relation $\sqsubseteq \subseteq \mathcal{I} \times \mathcal{T} \times \mathcal{I}$ such that, for all $K \in \mathcal{T}$, $\sqsubseteq_K$ is reflexive and transitive.

A set of interpretations M is called *closed* if

$$\mathsf{Mod}(\mathrm{Th}(\mathsf{M})) = \mathsf{M}.$$

The set $\downarrow_{\sqsubseteq_K}(\mathsf{M})$ is defined as $\{J \mid \exists I \in \mathsf{M}\; J \sqsubseteq_K I\}$, and likewise for $\uparrow_{\sqsubseteq_K}(\mathsf{M})$.

The following properties of RS preference relations will be used below. An RS preference relation is

1. *sound* if for any satisfiable K, I is $\sqsubseteq_K$-minimal in $\mathcal{T}$ iff $I \models K$;
2. *stoppered* if for all sets of formula X and interpretations $I \in \mathsf{Mod}(X)$, there exists $J \sqsubseteq_K I$ with $J \in \mathrm{Min}(\mathsf{Mod}(X), \sqsubseteq_K)$;
3. *abstract* if $\mathrm{Th}(I) = \mathrm{Th}(J)$ implies $I \sqsubseteq_K J \sqsubseteq_K I$;
4. *preserves closed sets* if for all K and closed sets of interpretations M, the sets $\mathrm{Min}(\mathsf{M}, \sqsubseteq_K)$ and $\downarrow_{\sqsubseteq_K}(\mathsf{M})$ and $\uparrow_{\sqsubseteq_K}(\mathsf{M})$ are closed.

Conditions 1 and 3 are similar to (part of the) faithfulness conditions defined in Section 5. Condition 2 is similar to the smoothness condition used in the definition of preferential models; it tells us that any theory has minimal models. Preservation of closedness is, indeed, about preservation of closedness under certain operations on sets of interpretations.

Given the above definitions, we can define revisions in the obvious way:

$$K *^{\sqsubseteq} \phi := \mathrm{Th}(\mathrm{Min}(\mathsf{Mod}(\phi), \sqsubseteq_K)). \quad (25)$$

Ryan and Schobbens introduce a number of rather technical conditions on $\sqsubseteq$ to guarantee that $*^{\sqsubseteq}$ satisfies the AGM postulates. Our next concern here is to relate the above to a relation on theories called verisimilitude — this will be the topic of the following section.

Notes

The presentation of preferential reasoning and preferential models in this section is based on [21]. The translations taking postulates for revision to nonmonotonic inference rules, and conversely, were first proposed in [27]. The final parts of this section are based on [32]. A comparison of a variety of approaches to using *minimality* in logic is presented in [26].

7 Verisimilitude

Verisimilitude is about measuring how close theories are to the truth; that is, about measuring which theories are better approximations of the complete theory of 'everything' than others. Instead of closeness to 'The Truth,' we will be more modest, and look at closeness to 'the available evidence,' i.e., approximations of a given theory K which is taken to represent the evidence.

A *verisimilitude relation* is a ternary relation between theories:

$$A \Subset_K B \text{ if } A \text{ is as close to } K \text{ as } B \text{ is.}$$

We will assume that $\Subset$ is reflexive and transitive.

The ternary relation allows us to select, from a given collection $\mathcal{T}$ of theories, one which is closest to a given theory K in an obvious way: A is *closest* to K if it is $\Subset_K$-minimal in $\mathcal{T}$:

$$A \in \mathrm{Min}(\mathcal{T}, \Subset_K).$$

Observe that the above allows for different incomparable theories to be closest to K.

Historical Comments

The first formal definition of verisimilitude is due to Karl Popper:

$$A \Subset_K^{(P)} B \text{ iff } B \cap K \subseteq A \text{ and } A \setminus K \subseteq B.$$

If we assume that K is maximal (i.e., decides on any given formula), then the above definition says that A is closer to K than B if A has all the true sentences that B has, and A has no more false sentences in it than B has. Without the assumption of maximality, the latter condition is not very intuitive.

An alternative definition is due to Miller and Kuipers:

$$A \Subset_K^{(M)} B \text{ iff } \mathsf{Mod}(B) \cap \mathsf{Mod}(K) \subseteq \mathsf{Mod}(A) \text{ and } \mathsf{Mod}(A) \setminus \mathsf{Mod}(K) \subseteq \mathsf{Mod}(B).$$

The intuition here is that any model for B that might be the true situation must also be a model for A, and any model for A that can't be the true situation must also be a model for B.

The above two proposals have a number of undesirable properties:

1. $A \neq B$ and $A \Subset_K^{(P)} B$ implies $A \subseteq K$
2. If K is maximal and $\mathsf{Mod}(K) \cap \mathsf{Mod}(A) = \emptyset$, then $A \Subset_K^{(M)} B$.

In words, item 1 says that $\Subset_K^{(P)}$ can't strictly order 'false' theories (i.e., theories that contain at least one statement not in K)... but this was exactly the purpose of verisimilitude. Item 2 says that the contradictory theory (with no models) is an improvement on any theory that shares no models with K.

Back and Forth between $\sqsubseteq$ and $\Subset$

We will now describe ways of obtaining a verisimilitude relation from a preference relation, and the other way around. We start with the latter case. Given a preference relation, how can we use it to define a verisimilitude relation? We use an idea from computer science called the power-ordering or the Egli-Milner ordering. What this does is the following. Given a relation R on a set $\mathcal{X}$, one can 'lift' R to a relation on the power set $\mathcal{P}(\mathcal{X})$ of $\mathcal{X}$ as follows.

$$XR^+Y \text{ iff } \forall x \in X \exists y \in Y\, xRy \wedge \forall y \in Y \exists x \in X\, xRy.$$

Thus, Y reaches up higher than X, and X reaches down lower than Y.

Brink and Heidema propose using the powering idea in the following manner to derive a verisimilitude relation from a preference relation. By identifying theories A, B with sets of interpretations (namely $\mathsf{Mod}(A)$ and $\mathsf{Mod}(B)$), $\sqsubseteq$ can be lifted as follows

$$\begin{aligned} A \Subset_K^{\sqsubseteq} B \text{ iff } & \forall I \in \mathsf{Mod}(A) \exists J \in \mathsf{Mod}(B)\, I \sqsubseteq_K J \text{ and} \\ & \forall J \in \mathsf{Mod}(B) \exists I \in \mathsf{Mod}(A)\, I \sqsubseteq_K J. \end{aligned}$$

The intuition here is that A is as close to K as B if every model of A is as close to K as some model of B, and every model of B is a far from K as some model of A.

Going from $\Subset$ to $\subseteq$ is easier. Every interpretation I gives us a theory, namely $\mathrm{Th}(I)$. If we are comparing theories for closeness to K, then we're comparing interpretations too:

$$I \sqsubseteq_K^{\Subset} B \text{ iff } \mathrm{Th}(I) \Subset_K \mathrm{Th}(J).$$

Proposition 7.1 *If $\sqsubseteq$ is a preference relation, then $\Subset^{\sqsubseteq}$ is a verisimilitude relation. If $\Subset$ is a verisimilitude relation, then $\sqsubseteq^{\Subset}$ is a preference relation.*

Using a verisimilitude relation, one can define a revision operation as follows. Given a belief set K and a formula ϕ, consider the theories that contain ϕ, and select among them those that are closest to K:

$$K *^{\Subset} \phi := \bigcap \mathrm{Min}(\mathrm{Ctg}(\phi), \Subset_K),$$

where $\mathrm{Ctg}(\phi) = \{A \in \mathcal{T} \mid \phi \in A\}$. The similarity between the definitions of $*^{\Subset}$ and $*^{\sqsubseteq}$ becomes very clear once we unfold the definitions:

$$\begin{aligned} K *^{\Subset} \phi &= \bigcap \mathrm{Min}(\mathrm{Ctg}(\phi), \Subset_K) \\ &= \{\psi \mid \mathrm{Min}(\mathrm{Ctg}(\phi), \Subset_K) \subseteq \mathrm{Ctg}(\psi)\} \\ K *^{\sqsubseteq} \phi &= \mathrm{Th}(\mathrm{Min}(\mathsf{mod}(\phi), \sqsubseteq_K)) \\ &= \{\psi \mid \mathrm{Min}(\mathsf{Mod}(\phi), \sqsubseteq_K) \subseteq \mathsf{Mod}(\psi)\}. \end{aligned}$$

Our next aim is to make a number of round-trips: start from a preference relation, use it to define a verisimilitude relation, and use that to define a preference relation again. What is the connection between the first and the second preference relation? And similarly, what happens if we make a round-trip starting from a verisimilitude relation? And what is the relation between the revision operators defined from $\sqsubseteq$ and $\Subset$?

Before considering these questions, we state a result that nicely relates the 'best' theories to the 'best' interpretations:

Lemma 7.2 *If $\sqsubseteq$ is stoppered and preserves closed sets, then*

$$\begin{aligned} & B \in \mathrm{Min}(\mathrm{Ctg}(A), \Subset_K^{\sqsubseteq}) \textit{ iff} \\ & \quad \mathsf{Mod}(B) \subseteq \mathrm{Min}(\mathsf{Mod}(A), \sqsubseteq_K). \end{aligned}$$

So, the best theories are those with the most preferred models.

If we start from a preference relation, we can expect the round-trip to produce the same preference relation for the following reasons. A preference relation $\sqsubseteq$ orders *total* models, but $\Subset$ contains much more structure in that it also compares partial and *incomplete* information. Thus moving from $\sqsubseteq$ to $\Subset$ introduces a lot of additional structure, but when we move on from $\Subset$ to $\sqsubseteq$ we forget about this additional structure again.

Proposition 7.3 *If $\sqsubseteq$ is abstract, then $\sqsubseteq^{\Subset^{\sqsubseteq}} = \sqsubseteq$.*

By the same intuitions, we should expect a safe round-trip starting from a verisimilitude relation $\Subset$ if we only consider maximal theories.

Proposition 7.4 *If A, B are maximal and consistent, then $A \Subset_K^{\sqsubseteq^{\Subset}} B$ iff $A \Subset B$.*

To get a safe round-trip for arbitrary theories, we need to impose strict conditions on $\Subset$; see [32] for details.

Finally, we compare some of the different revision operators generated in this section.

Proposition 7.5 *If $\sqsubseteq$ is a stoppered preference relation that preserves closed sets, then $*^{\sqsubseteq} = *^{\Subset^{\sqsubseteq}}$.*

The above proposition can be proved using fairly weak assumptions; this has to do, again with the fact that moving from a preference relation to a verisimilitude relation does not destroy information. Moving in the opposite direction, we do loose information, and, hence, we need to impose stronger conditions to arrive at $*^{\Subset} = *^{\sqsubseteq^{\Subset}}$. The conditions involved are 'proof-generated' and too involved to be explained here.

Notes

The presentation in this section is based on [32] and [4]; see also [5].

8 A Descriptive Approach

In this section we develop a descriptive approach to theory change. By this we mean the following. As explained in Section 2, theory change may be viewed as a process, where new information induces a transition from one information state to another. The sort of questions we are interested in include: Which transitions are possible? Is there a specific structure to the transitions? Which formulas are accepted in an information state after a transition induced by some formula ϕ? In short, the same kind of questions as one finds in descriptive approaches to the semantics of programs (see, for example, [28]).

There are several reasons why a descriptive approach to theory change may be valuable. First, it is much easier to compare and evaluate different proposals for dynamic operations in a technical way. Second, by using fairly 'neutral' tools, we can start meta-theoretical investigations of the phenomena being described using these tools, and thus get a precise, mathematical understanding of the complexities and requirements of proposals in theory change. Third, proposals which are defined in a standard logical style can be modified or extended fairly easily. And last but not least, a technical treatment of theory change may bring issues to light which are ignored by philosophical debate on postulates.

We will use models and languages from a number of logical disciplines, namely intuitionistic logic, modal logic, and dynamic logic.

Intuitionistic Logic

One branch of logic has had an information oriented flavor from the very start: intuitionistic logic. The underlying idea is that formulas of intuitionistic logic describe the way an idealized mathematician acquires new (mathematical) knowledge. So, we will take the truth of a formula ϕ in an information state to mean that our mathematician knows ϕ or that she has acquired ϕ.

The *language* of intuitionistic logic is simply the language of ordinary propositional logic:

$$\phi ::= p \mid - \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \rightarrow \phi.$$

This language is interpreted on so-called information models.

Definition 8.1 An *information frame* is a pair $M = (W, \trianglelefteq)$ with W a non-void set of *information states* and $\trianglelefteq$ a pre-order over W, that is: a reflexive, transitive relation, which is called an *information order*. Intuitively, a move along $\trianglelefteq$ is a move to a more informative state.

Let $\mathcal{L}$ be some language, with a set of atomic symbols *Prop*. An $\mathcal{L}$-*information model* M is a triple $(W, \trianglelefteq, V)$ with $(W, \trianglelefteq)$ an information structure, and $V : \mathit{Prop} \rightarrow \mathcal{P}(W)$. The function V will be called an $\mathcal{L}$-*valuation*.

A model for intuitionistic logic is an information model that satisfies the following *persistence condition*:

$$\text{If } w \trianglelefteq v \text{ and } w \in V(p), \text{ then } v \in V(q).$$

That is: one can never lose (atomic) information by moving up along the information order.

Next we define what it means for a formula 'ϕ to be '*true* at a state w in a model M' (in symbols: $M, w \models \phi$):

$$\begin{array}{rcl}
M, w \models - & & \text{never} \\
M, w \models p & \text{iff} & w \in V(p) \\
M, w \models \phi \wedge \psi & \text{iff} & M, w \models \phi \text{ and } M, w \models \psi \\
M, w \models \phi \vee \psi & \text{iff} & M, w \models \phi \text{ or } M, w \models \psi \\
M, w \models \phi \rightarrow \psi & \text{iff} & \text{for all } v \text{ such that } w \trianglelefteq v \text{ and} \\
 & & M, v \models \phi \text{ we have } M, v \models \psi \\
M, w \models \neg\phi & \text{iff} & \text{for no } v \text{ such that } w \trianglelefteq v \\
 & & \text{we have } M, v \models \phi
\end{array}$$

The idea of the clause for implications is that $\phi \rightarrow \psi$ is true of one's current information state if, whenever the information grows so as to include ϕ, it should also include ψ. And the clause for negation says that $\neg\phi$ is accepted if there's no way of extending our information so as to include ϕ.

Intuitionistic Logic as a Theory of Information

The only parts where the above truth definition deviates from the one for ordinary propositional logic, is in the clauses for $\rightarrow$ and $\neg$: they exploit the information order. As a result, certain familiar principles from classical logic are not valid in intuitionistic logic. The formula $p \vee \neg p$, for example, is not valid.

This is in full agreement with the idea of intuitionistic logic as logic that the describes the cognitive moves of a mathematician as she pursue new results: it would be unrealistic to demand that $p \vee \neg p$ be always true; this would mean that our mathematician is always in a position where she either knows p or knows $\neg p$.

It may be shown that *all* intuitionistic formulas are persistent, not just the atomic ones. As a consequence, we can really only talk about expansions in intuitionistic logic. To be able to specify contractions or revisions we need to be able to move backwards along the information order as well.

Modal and Temporal Logic

We will now consider a number of classical logics for reasoning about information models, the first one of which is *modal logic*. To be able to exploit the information order present in our models, the syntax of modal logic contains two unary operators $\Diamond$ ('diamond') and $\Box$ ('box'):

$$\phi ::= p \mid - \mid \neg\phi \mid \phi \wedge \phi \mid \Diamond\phi \mid \Box\phi.$$

We interpret modal formulas on information models as defined in Definition 8.1 by using the usual classical clauses for the boolean connectives, and the following clauses for the modal operators.

$$\begin{array}{lll} M, w \models \Diamond\phi & \text{iff} & \text{there exists } v \text{ with } w \trianglelefteq v \text{ and} \\ & & M, v \models \phi \\ M, w \models \Box\phi & \text{iff} & \text{for all } v, \text{ if } w \trianglelefteq v \text{ then } M, v \models \phi. \end{array}$$

A formula ϕ is *true on a model* if it is true in all states in the model; ϕ is called *valid* on an information frame $F = (W, \trianglelefteq)$ if it is true on all information models based on F. As an example, both $\Box p \to p$ and $\Box\phi \to \Box\Box\phi$ are valid on all information frames. The formula $\Diamond p \to \Box\phi$ is not valid on all information frames.

Knowledge and Information in Modal Logic.

Now that we have briefly introduced modal logic, how can we use it to reason about theory change? The first thing we need is a modal counterpart of what a *theory* is. Here the link with intuitionistic logic helps to guide the intuition: in intuitionistic logic information is represented by means of *persistent formulas*. What are the persistent formulas in modal logic? Clearly all boxed formulas (i.e., formulas of the form $\Box\phi$) are persistent, as are conjunctions and disjunctions of boxed formulas.

This suggests that we represent theories as sets of the form

$$\{\phi \mid w \models \Box\phi\},$$

where w is a state in a model. So, with each state we associate a theory, and a formula ϕ is an element of the theory of that state if $\Box\phi$ is true at the state. Then, a formula of the form $\Diamond\Box\phi$ can be given an expansion-like reading: it says that we can move up along the information order to a state where ϕ is in the theory. Further, one can define an expansion-like modal operator $[+\phi]\psi$ which should be read as 'ψ is in every theory resulting from an expansion with ϕ' by putting:

$$[+\phi]\psi := \Box(\Box\phi \to \Box\psi).$$

The above ideas can be traced back to a special branch of modal logic, called *epistemic logic* in which a formula of the form $\Box\phi$ is interpreted as 'it is known that ϕ.'

One obvious shortcoming of the modal language we have looked at so far, is that, just as with intuitionistic logic, we only seem to be able to express properties of expansions: there is no way we can move back along the information order to a state where a given formula is no longer part of the theory. To accommodate this, we will now extend our modal language.

Adding a Direction: Temporal Logic

In the language of temporal logic we are able to talk about moves back and forth along the information order, thus allowing us to model further operations of theory change besides expansions.

The language of temporal logic has both forward looking and backward looking operators. Instead of $\Diamond$ and $\Box$ we write $\langle\trianglelefteq\rangle$ for the forward looking diamond, and $[\trianglelefteq]$ for the forward looking box.

$$\phi ::= p \mid - \mid \neg\phi \mid \phi \wedge \phi \mid \langle\trianglelefteq\rangle\phi \mid [\trianglelefteq]\phi \mid \langle\trianglerighteq\rangle\phi \mid [\trianglerighteq]\phi.$$

Formulas of temporal logic are interpreted on information models in the following way:

$$\begin{array}{lll} M, w \models \langle\trianglelefteq\rangle\phi & \text{iff} & \text{there exists } v \text{ with } w \trianglelefteq v \text{ and} \\ & & M, v \models \phi \\ M, w \models [\trianglelefteq]\phi & \text{iff} & \text{for all } v \text{ with } w \trianglelefteq v,\ M, v \models \phi \\ M, w \models \langle\trianglerighteq\rangle\phi & \text{iff} & \text{there exists } v \text{ with } v \trianglelefteq w \text{ and} \\ & & M, v \models \phi \\ M, w \models [\trianglerighteq]\phi & \text{iff} & \text{for all } v \text{ with } v \trianglelefteq w,\ M, v \models \phi. \end{array}$$

Models for temporal logic are often given a temporal interpretation: instead of $\trianglelefteq$ one often writes $<$ or $\leq$, which is then read as 'later than' or 'not before'. Traditionally, in temporal logic one writes $F\phi$ (at some time in the future ϕ will hold) for $\langle\trianglelefteq\rangle\phi$; $G\phi$ (it is going to be the case that ϕ) for $[\trianglelefteq]\phi$; $P\phi$ (at some time in the past ϕ) for $\langle\trianglerighteq\rangle\phi$; and $H\phi$ (it has always been the case that ϕ) for $[\trianglerighteq]\phi$.

Knowledge and Information in Temporal Logic

Now that we have the means to talk about moves forward *and backwards* along the information order of our information models, let us try and use this to specify a contraction like operator. First of all, observe that a formula of the form

$$\langle\trianglerighteq\rangle\neg[\trianglelefteq]\phi$$

may be taken to describe the possibility of giving up ϕ from the current theory. Next, here's an implementation in our temporal language of a contraction operator $[-\phi]\psi$ ('ψ belongs to every theory resulting from contracting with ϕ'):

$$[-\phi]\psi := [\trianglerighteq](\neg[\trianglelefteq]\phi \to [\trianglelefteq]\psi).$$

Here's an alternative proposal for a modal contraction operator:

$$[-\phi]_2\psi :=$$

$$[\trianglerighteq]\Big(\neg[\trianglelefteq]\phi \to \langle\trianglelefteq\rangle(\neg[\trianglelefteq]\phi \wedge [\trianglelefteq]\psi)\Big).$$

The intuition here is the following. Instead of demanding that ψ is in *every* theory that results from giving up ϕ, we allow for a bit more flexibility. We require that after every way of giving up ϕ we can extend the resulting theory so as to arrive at a belief set that still doesn't contain ϕ and that will contain ψ.

Fuhrmann's Logic of Theory Change

Fuhrmann [7] proposed a modal approach to theory change in which each formula ϕ is associated with its own 'contract-with-ϕ'-relation, instead of having a global information order of which individual contractions are subsets. Formally, Fuhrmann's language is given by the following rule

$$\phi ::= p \mid - \mid \neg\phi \mid \phi \wedge \phi \mid [-\phi]\phi \mid \Box\phi.$$

The operator $[-\phi]\psi$ is read 'ψ holds after every contraction with ϕ,' and $\Box\phi$ as 'ϕ holds,' or 'ϕ is currently in the theory.'

The minimal logic in this language is axiomatized by taking all classical tautologies together with

- $[-\phi](\psi \to \chi) \to ([-\phi]\psi \to [-\phi]\chi)$
- $\Box\phi \leftrightarrow [-\top]\phi$
- from $\vdash \psi$ infer $\vdash [-\phi]\psi$
- from $\vdash \phi \leftrightarrow \psi$ infer $[-\phi]\chi \leftrightarrow [-\psi]\chi$
- from $\vdash \phi$, $\vdash \phi \to \psi$ infer $\vdash \psi$.

On top of these one may add axioms corresponding to the AGM postulates for contraction. These may be obtained by translating statements about set theoretic inclusion into implications, and statements of the form $K \dagger \phi$ ($\dagger \in \{+,-\}$) translate into $[\dagger\phi]\cdot$.

(F2) $[-\phi]\psi \to \Box\psi$
(F3) $\neg\Box\phi \wedge \Box\psi \to [-\phi]\psi$
(F4) $[-\phi]\phi \to [-\psi]\phi$
(F5) $\Box\psi \to [-\phi](\phi \to \psi)$
(F7) $[-\phi]\chi \wedge [\psi]\chi \to [-(\phi \wedge \psi)]\chi$
(F8) $\neg[-(\phi \wedge \psi)]\phi \wedge [-(\phi \wedge \psi)]\chi \to [-\phi]\chi$.

The translation of the fourth postulate (if $\not\vdash \phi$, then $\phi \notin K - \phi$) calls for some comments. Unlike the other postulates for contraction, it resists a direct translation into Fuhrmann's logic. What we have given as a translation is at least an approximation: if ϕ survives a contraction with ϕ, then it must survive any contraction, the idea being that only theorems of the logic survive 'self-contraction.'

Let us turn to the semantics for Fuhrmann's logic now. A *Fuhrmann* frame is a structure (W, P, C) with W a non-empty set of states, $P \subseteq \mathcal{P}(W)$, and C a family of binary relations on W, one for each element of P: $C = \{C_X \subseteq W^2 \mid X \in P\}$. The following closure conditions are imposed (here $C_X(w)$ denotes $\{v \mid C_X wv\}$):

- $W \in P$
- if $X \in P$, then $W \setminus X \in P$
- if $X, Y \in P$, then $(X \cup Y) \in P$
- if $X, Y \in P$, then $\{v \mid C_X(v) \subseteq Y\} \in P$.

A Fuhrmann frame is turned into a *Fuhrmann model* by adding a valuation that assigns elements of P to proposition letters. The boolean truth conditions are the usual ones, while $w \models [-\phi]\psi$ iff $C_{V(\phi)}(w) \subseteq V(\psi)$, and $w \models \Box\psi$ iff $C_W(w) \subseteq V(\psi)$.

On top of the basic Fuhrmann models one may impose conditions that correspond to the (translations of the) AGM postulates for contraction. For each of the (translated) postulates (Fn) there is a corresponding semantic condition (Cn) that is satisfied by a Fuhrmann frame iff the postulate is true in every model based on the frame.

(C2) $C_W \subseteq C_X$
(C3) if $C_W(w) \not\subseteq X$, then $C_X \subseteq C_W$
(C4) if $C_X(w) \subseteq X$, then $X = W$
(C5) if $C_X(w) \cap X \neq \emptyset$, then $C_X \subseteq C_W$
(C7) $C_{X\cap Y} \subseteq (C_X \cup C_Y)$
(C8) if $C_{X\cap Y}(w) \not\subseteq X$, then $C_X \subseteq C_{(X\cap Y)}$.

The result is that Fuhrmann's logic, extended with the axioms (F2)–(F8) is sound and complete for all Fuhrmann frames satisfying conditions (C2)–(C8).

The relations C_X ($X \in P$) are rather abstract tools — what is their relation to the more intuitive picture involving an information order $\trianglelefteq$ discussed before? Clearly, the relations C_X can be combined into an information order by putting

$$w \trianglelefteq v \quad \text{iff} \quad w = v \text{ or there are } w_0, w_1, \ldots, w_n \text{ and } X_1, \ldots, X_n \in P \text{ such that } w = w_0 C_{X_1} w_1 \cdots C_{X_n} w_n = v$$

This definition ensures that $\trianglelefteq$ is a pre-order. The precise connection between these (and other descriptive) ways of modeling theory change remains to be determined, though.

To conclude this section we return to the format of information models equipped with an information order, with contractions and revisions being modeled as subsets of moves along the information order.

Dynamic Modal Logic

Dynamic modal logic (DML) was designed as a general and expressive modal language for reasoning about as many of the currently available proposals for analyzing change and action in logic, computer science and linguistics. The main motivation was that a general language like the DML language would allow one to compare these proposals in a common setting so that differences and similarities would become visible, and so that techniques and results from one proposal could be transferred to another.

In the remainder of this section we briefly introduce the framework offered by DML, and then show how it may be used to talk about theory change.

DML has propositions to describe states and procedures to describe moves through information models. In addition it has systematic links taking propositions to procedures and vice versa, as sketched below:

$$\text{propositions (BA)} \;\underset{\text{projections}}{\overset{\text{modes}}{\rightleftarrows}}\; \text{procedures (RA)}$$

We use ϕ, ψ, ... to denote formulas, and α, β, ... to denote procedures. Formulas and procedures are produced by the following rules:

$$\begin{array}{lcl} \phi & ::= & p \mid - \mid \neg\phi \mid \phi \wedge \phi \mid \mathsf{dom}(\alpha) \mid \mathsf{ran}(\alpha) \mid \mathsf{fix}(\alpha) \\ \alpha & ::= & \mathsf{exp}(\phi) \mid \mathsf{con}(\phi) \mid -\alpha \mid \alpha\breve{} \mid \alpha;\alpha \mid \alpha \cup \alpha \mid \phi? \end{array}$$

The readings of the above formulas and procedures are as follows:

$\mathsf{dom}(\alpha)$ is a formula that takes a procedure α as its input; it is true at a state if from that state an α-procedure can be executed.

$\mathsf{ran}(\alpha)$ is a formula that takes a procedure α as its input; it is true at a state if that state can be reached by executing an α-procedure.

$\mathsf{fix}(\alpha)$ is a formula that takes a procedure α as its input; it is true at a state if it is a fixed point for α, that is, if an α-loop can be made at that state.

$\mathsf{exp}(\phi)$ is a procedure that takes a formula ϕ as its input; it denotes all steps along the information order that lead to a state where ϕ holds.

$\mathsf{con}(\phi)$ is a procedure that takes a formula ϕ as its input; it denotes all steps *backwards* along the information order that lead to a state where ϕ fails.

$\phi?$ is again a procedure that takes a formula ϕ as its input; it denotes the test-for-ϕ relation.

The operations on procedures are familiar ones from relation algebra; $-$ is complementation, $\breve{}$ is the converse operation; ; is composition, and $\cup$ is simply union.

With the above explanations we can say what it means for a formula to be true. We interpret formulas on information models that are extended with a device $[\![\cdot]\!]$ for handling procedures. So, in the remainder of this section information models are 4-tuples $M = (W, \trianglelefteq, V, [\![\cdot]\!])$ where the first three items are as before, and $[\![\cdot]\!]$ associates a binary relation on W with every procedure α.

The novel truth clauses are the following:

$$\begin{array}{rcl} M, w \models \mathsf{dom}(\alpha) & \text{iff} & \text{there exists } v \text{ with } (w, v) \in [\![\alpha]\!] \\ M, w \models \mathsf{ran}(\phi) & \text{iff} & \text{there exists } v \text{ with } (v, w) \in [\![\alpha]\!] \\ M, w \models \mathsf{fix}(\alpha) & \text{iff} & (w, w) \in [\![\alpha]\!]. \end{array}$$

The clauses for assigning meanings to procedures are

$$\begin{array}{rcl} [\![\mathsf{exp}(\phi)]\!] & = & \{(x,y) \mid x \trianglelefteq y \text{ and } M, y \models \phi\} \\ [\![\mathsf{con}(\phi)]\!] & = & \{(x,y) \mid y \trianglelefteq x \text{ and } M, y \not\models \phi\} \\ [\![-\alpha]\!] & = & (W \times W) \setminus [\![\alpha]\!] \\ [\![\alpha\breve{}]\!] & = & \{(x,y) \mid (y,x) \in [\![\alpha]\!]\} \\ [\![\alpha;\beta]\!] & = & [\![\alpha]\!];[\![\beta]\!] \\ & = & \{(x,y) \mid \exists z\, ((x,z) \in [\![\alpha]\!] \wedge (z,y) \in [\![\beta]\!])\} \\ [\![\alpha \cup \beta]\!] & = & [\![\alpha]\!] \cup [\![\beta]\!] \\ [\![\phi?]\!] & = & \{(x,y) \mid x = y \wedge M, y \models \phi\}. \end{array}$$

The familiar diamond and box operators from modal and temporal logic can be expressed in DML as follows:

$$\begin{array}{rcl} \langle\trianglelefteq\rangle\phi & \leftrightarrow & \mathsf{dom}(\mathsf{exp}(\phi)) \\ \langle\trianglerighteq\rangle\phi & \leftrightarrow & \mathsf{dom}(\mathsf{con}(\neg\phi)) \end{array}$$

So, the earlier modal and temporal logics can both be viewed as fragments of DML.

Before moving on, here are some useful abbreviations: $\delta := \top?$; so δ denotes the *diagonal* $\{(x,y) \mid x = y\}$, and $-\delta$ denotes the *diversity* relation $\{(x,y) \mid x \neq y\}$.

One important feature missing from the modal and temporal logics introduced earlier in this section was the ability to express *minimal* moves along the information order. With the DML machinery to our disposal we are finally able to express such minimal moves. A minimal move along the information order to a state w where ϕ holds is nothing but a move along $\trianglelefteq$ to w that cannot be decomposed into a move to a closer ϕ-state followed by a further step along $\trianglelefteq$ to w. Formally, we use μ-$\mathsf{exp}(\phi)$ to denote such minimal $\trianglelefteq$-moves to a ϕ-state:

$$\mu\text{-}\mathsf{exp}(\phi) \quad := \quad \mathsf{exp}(\phi) \cap -\Big(\mathsf{exp}(\phi); (-\delta \cap \mathsf{exp}(\top))\Big).$$

A minimal version of con can be defined similarly.

Modeling Theory Change

We will now introduce modal operators $[\dagger\phi]\psi$, where $\dagger$ is one of $+$, $-$, $*$ to denote that ψ is in every theory that results from expanding (contracting/revising) the current theory with ϕ.

As before, we use 'boxed formulas' to represent theories, and hence if $[\trianglelefteq]\phi$ is true at a state w, this will denote that ϕ is in the theory associated with w.

Expansions. To define the operator $[+\phi]\psi$ (ψ is in every result of expanding the current theory with ϕ) observe that expanding with ϕ involves a move along the information order to a state where ϕ is in theory, and moreover, this move should be a minimal one:

$$[+\phi]\psi \quad := \quad \neg\mathsf{dom}\Big(\mu\text{-}\mathsf{exp}([\trianglelefteq]\phi); (\neg[\trianglelefteq]\psi)?\Big)$$

In other words: it should not possible to first extend the current theory with ϕ (i.e., make a minimal move to a state w where $[\trianglelefteq]\phi$ is true), and then find that ψ is not in the result (i.e., find that $[\trianglelefteq]\psi$ is false at w).

Contractions. Our definition of a modal contraction operation mirrors the one of the modal expansion operation: $[-\phi]\psi$ is true if ψ is in every result of contracting with ϕ.

$$[-\phi]\psi \quad := \quad \neg\mathsf{dom}\Big(\mu\text{-}\mathsf{con}([\trianglelefteq]\phi); (\neg[\trianglelefteq]\psi)?\Big)$$

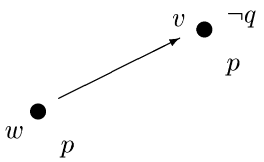
Given this modal contraction operation, we now try and translate the AGM postulates into DML. We can then see whether or not the (translated) AGM postulates come out valid, and if they don't, what kind of additional constraints we have to impose on our models.

By way of example we will translate postulates $(K-2)$–$(K-6)$ into DML. As with Fuhrmann's logic, the idea is that set theoretic inclusion translates into an implication, and that statements of the form $K \dagger \phi$ ($\dagger \in \{+,-\}$) translate into $[\dagger\phi]$.

$(K-2)$ $[-\phi]\psi \to [\trianglelefteq]\psi$.
$(K-3)$ $\neg[\trianglelefteq]\phi \wedge [\trianglelefteq]\psi \to [-\phi]\psi$.
$(K-4)$ $[-\phi]\phi \to [-\psi]\phi$.
$(K-5)$ $[\trianglelefteq]\psi \to [-\phi][+\phi]\psi$.
$(K-6)$ if $\vdash \phi \leftrightarrow \psi$ then $\vdash [-\phi]\chi \leftrightarrow [-\psi]\chi$.

As in the case of Fuhrmann's logic, the fourth postulate resists a direct translation into DML.

We will now look at the translation of the second postulate, and determine whether or not it is valid on all information models. Let w be a state that refutes an instance of the translation of $(K-2)$. That is, assume $w \models [-p]q, \neg[\trianglelefteq]q$. Then there exists a $\trianglelefteq$-successor v of w with $v \not\models q$. Now, to guarantee that $w \models [-p]q$, we need to ensure that every minimal way of giving up $[\trianglelefteq]p$ leads to a $[\trianglelefteq]q$-state. But, if there is no way whatsoever of giving up p, then this requirement is vacuously true — so let us require that $v, w \models p$. Putting things together, the following model refutes the translation of $(K-2)$.



What can we do about the failure of $(K-2)$? First of all, we can observe (again) that, according to AGM, contractions with non-theorems are always possible and defined; this is part of the functional reading of the operation $-$ in the AGM theory. Next, we can try to impose a similar constraint on our information models: in every state, and for every non-theorem ϕ it should be possible to move back along the information order to a state where ϕ is no longer part of the theory, that is, to a state where $[\trianglelefteq]\phi$ is false. Here is a somewhat unorthodox derivation rule to that effect:

$$\text{if} \vdash \neg\langle\trianglerighteq\rangle\neg[\trianglelefteq]\phi \text{ then} \vdash \phi. \qquad (26)$$

That is: if it is a theorem that $[\trianglelefteq]\phi$ is true down every information order, then ϕ itself must be a theorem.

We leave it to the reader to determine whether the translations of postulates $(K-3)$–$(K-6)$ are valid on all information models.

Revisions. To arrive at a modal counterpart of revisions, we will use the Levi Identity according to which $K * \phi = (K - \neg\phi) + \phi$. This leads to the following definition:

$$[*\phi]\psi \quad := \quad [-\neg\phi][+\phi]\psi.$$

We leave it to the reader to translate the AGM postulates for revision into DML, and to determine the validity of the translations.

Notes

Parts of this section are based on [2] and [30]; see also [31]. The material on Fuhrmann's logic of theory change is taken from [7].

9 Further Reading

To conclude these notes we give a few pointers to material related to the themes discussed here.

The descriptive approach to theory change put forward in Section 8 is closely related to descriptive approaches to other dynamic phenomena in logic, language and information, such as natural language semantics and artificial intelligence; Jaspars and Krahmer [17] provide a systematic approach. See Van Benthem [3] for further themes along these lines.

As pointed out in Section 5, the extension of the basic AGM paradigm to the setting of multiple agents is an active area of research. In addition to the references listed in Section 5 we should mention Jaspars [16], where multi-agent systems for (partial) descriptions of information change are studied, and Van Linder et al. [23], in which the authors combine the descriptive approach of Section 8 with ideas from agent-oriented programming.

The link between reasoning about changing information and actual implementations of database update operations is an underdeveloped area that deserves further exploration; see Winslett [34] for relevant references. We are confident that the flexibility of the descriptive approach will be beneficial here.

References

1. C.E. Alchourrón, P. Gärdenfors, and D. Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic*, 50:510–530, 1985.
2. J. van Benthem. Modal Logic as a Theory of Information. Technical Report LP-91-10, ILLC, University of Amsterdam, 1991.
3. J. van Benthem. *Exploring Logical Dynamics*. Studies in Logic, Language and Information. CSLI Publications, Stanford, 1996.
4. C. Brink. Verisimilitude: Views and Reviews. *History and Philosophy of Logic*, 10:181–201, 1989.
5. C. Brink and J. Heidema. A Verisimilar Ordering of Theories Based in a Propositional Language. *The British Journal for the Philosophy of Science*, 38:533–549, 1987.
6. J. Doyle. A Truth Maintenance System. *Artificial Intelligence*, 12:231–272, 1979.

7. A. Fuhrmann. On the Modal Logic of Theory Change. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change*, pages 259–281. LNAI 465, Springer, 1991.
8. A. Fuhrmann. *An Essay on Contraction*. Habilitationsschrift, Dept. of Philosophy, University of Konstanz, 1994.
9. A. Fuhrmann and S.-O. Hansson. A Survey of Multiple Contractions. *Journal of Logic, Language and Information*, 3:39–76, 1994.
10. P. Gärdenfors. *Knowledge in Flux*. The MIT Press, 1988.
11. P. Gärdenfors and D. Makinson. Revisions of Knowledge Systems Using Epistemic Entrenchement. In M. Vardi, editor, *Proc. 2nd TARK*, pages 83–95. Morgan Kaufmann, Los Altos CA, 1988.
12. P. Gärdenfors and H. Rott. Belief Revision. In D. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in AI, Vol. 4*, pages 35–132, Oxford University Press, 1993.
13. S.-O. Hansson. New Operators for Theory Change. *Theoria*, 55:114–132, 1989.
14. S.-O. Hansson. Reversing the Levi Identity. *Journal of Philosophical Logic*, 22:637–669, 1993.
15. W. van der Hoek and M. de Rijke. Interleaved Contractions. Report RR305, Dept. of Computer Science, University of Warwick, 1996.
16. J. Jaspars. *Calculi for Constructive Communication*. PhD thesis, ILLC Dissertation Series 1994-4 and ITK Dissertation Series 1994-1, Amsterdam and Tilburg, 1994.
17. J. Jaspars and E. Krahmer. A Program of Modal Unification of Dynamic Theories. In P. Dekker and M. Stokhof, editors, *Proceedings 10th Amsterdam Colloquium*, pages 425–444, ILLC, University of Amsterdam, 1996.
18. H. Katsuno and A.O. Mendelzon. Knowledge Base Revision and Minimal Change. *Artificial Intelligence*, 52:263–294, 1991.
19. H. Katsuno and A.O. Mendelzon. On the Difference Between Updating a Knowledge Base and Revising It. In P. Gärdenfors, editor, *Belief Revision*, pages 183–203. Cambridge University Press, 1992.
20. N.E. Kfir-Dahav and M. Tennenholtz. Multiagent Belief Revision. In *Proceedings TARK 96*, 1996, pages 175–194.
21. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, 44:167–207, 1990.
22. D. Lehmann. Belief Revision Revised. In *Proceedings IJCAI 95*, 1995.
23. B. van Linder, W. van der Hoek, and J.-J. Meyer. Actions That Make You Change Your Mind. In A. Laux and H. Wansing, editors, *Knowledge and Belief in Philosophy and Artificial Intelligence*, pages 103–146, Akademie Verlag, Berlin, 1995.
24. S. Lindström and W. Rabinowicz. On Probabilistic Representation of Non-probabilistic Belief Revision. *Journal of Philosophical Logic*, 18:69–101, 1989.
25. D. Makinson. On the Status of the Postulate of Recovery in the Logic of Theory Change. *Journal of Philosophical Logic*, 16:282–394, 1987.
26. D. Makinson. Five Faces of Minimality. *Studia Logica*, 52:339–379, 1993.
27. D. Makinson and P. Gärdenfors. Relations between the Logic of Theory Change and Nonmonotonic Logic. In A. Fuhrmann and M. Morreau, editors, *The Logic of Theory Change*, pages 185–205. LNAI 465, Springer-Verlag, Berlin, 1991.
28. Z. Manna and A. Pnueli. *The Temporal Logic of Concurrent and Reactive Systems*. Springer-Verlag, 1992.
29. B. Nebel. Syntax-based Approaches to Belief Revision. In P. Gärdenfors, editor, *Belief Revision*, pages 52–88. Cambridge University Press, 1992.
30. M. de Rijke. Meeting Some Neighbours. In J. van Eijck and A. Visser, editors, *Logic and Information Flow*, pages 170–195, MIT Press, Cambridge (Mass.), 1994.
31. M. de Rijke. A System of Dynamic Modal Logic. *Journal of Philosophical Logic*, to appear.
32. M. Ryan and P.-Y. Schobbens. Belief Revision and Verisimilitude. *Notre Dame Journal of Formal Logic*, 36:15–29, 1995.
33. K. Segerberg. Belief Revision from the Point of View of Doxastic Logic. Technical Report Number 9, Department of Philosophy, Uppsala University, 1994.
34. M. Winslett. Epistemic Aspects of Database. In D.M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 4*, pages 133–240, Oxford University Press, 1995.

Acknowledgements: Earlier versions of this paper were used in various forms during courses given at the University of Amsterdam, the University of Cape Town, and the University of the Saarland, Saarbrücken. A big thanks to all the colleagues and students who commented on earlier versions.