Contents

Chapter 1

Introduction

This Is the Title of This Story

Which Is Also Found Several Times in the Story Itself.

David Moser's story in *Metamagical Themas* (Douglas Hofstadter)

1.1 Motivation and overview

The accident at the Three Mile Island nuclear plant near Harrisburg, Pennsylvania, that took place in March 1979 nearly caused a catastrophe. In the subsequent investigation into the accident, the operators (agents) were blamed for 'errors' that were actually sensible defeasible conjectures based on default rules provided by their training. They were acting to avoid the risks associated with a 'loss of coolant' system state, unaware that the system was already in that state. This example is illustrative of the more general scenario in which control room agents, whether human or non-human, need to reason and act in a changing environment with only limited information available to them. It highlights the importance of determining the current state of the environment and the importance of having an internal representation (called an epistemic state) which reflects the state of the environment.

The 'environments' of control room agents form what we shall call the class of diagrammable systems, which, broadly speaking, are systems that are directly representable (in a sense to be made precise) with finitely many components. The key features of diagrammable systems are that they are dynamic and discrete. Control room agents themselves may be summarised as first-order intentional systems having specific informational and motivational attitudes, about which more will be said later. Control room agents are further characterised by a set of principles which may be viewed as a specialisation of more general principles for rational (ideal) agents. A key specialisation is the principle of Qualitativeness, which, motivated from cognitive science, advocates a qualitative approach, as opposed to the quantitative approaches underlying probabilistic or fuzzy methods. The quantitative approaches represent a large and important body of work but they violate our insistence on qualitativeness and therefore are not included in our comparisons.

The information-theoretic approach that is developed to address the questions raised in this thesis is purely qualitative in nature. The central construct used in the information-theoretic approach is the concept of a templated ordering, or t-ordering for short. T-orderings have their origins in the dichotomous 'partitions' associated with semantic information theory (Carnap and Bar-Hillel, 1952) and in the strict modular partial orders associated with the approach of Lehmann and Magidor (1992) to nonmonotonic logic. Unlike the weaker semantic representations of information afforded by these approaches, t-orderings provide a unified semantic framework for the combination of both definite information (or knowledge) and indefinite information (or belief). T-orderings are closely related to the ordinal conditional functions of Spohn (1988) but, unlike ordinal conditional functions which rely on the arithmetic of ordinals, t-orderings are purely qualitative.

Finding appropriate representations of the epistemic states of agents is an important issue in the area of knowledge representation and reasoning. Representing the epistemic state of a control room agent by a (regular) t-ordering signals a departure from the many approaches in which only the beliefs of an agent are represented. T-orderings allow for the representation of both knowledge and belief, as advocated by the proposed principle of Duality. The distinction between knowledge and belief is not based on the traditional Platonic view of 'true justified belief' but instead related to a psychological notion of 'entrenchment of belief' in which knowledge arises as a special case when some threshold of entrenchment is surpassed. The triangular connection that is shown to exist between the information-theoretic semantics for epistemic logic of Labuschagne and Ferguson (2002), epistemic entrenchment (Gärdenfors and Makinson, 1988), and t-orderings, provides an information-theoretic justification for representing the epistemic state of a control room agent by a (regular) t-ordering.

The area of belief change, which has been an active research area for at least two decades, has as its focus the issue of how an agent changes its epistemic state upon receiving new information. The classical approaches of Alchourrón, Gärdenfors, and Makinson (1985) and Katsuno and Mendelzon (1992) continue to serve as the comparative benchmarks for belief revision and belief update respectively. The problem of iterated revision has received a lot of attention in recent years (in stark contrast to the problem of iterated update which has largely been ignored) and the approach of Darwiche and Pearl (1997) has proved to be very influential, despite continued criticism. In the context of control room agents, the notions of templated revision and templated update represent an information-theoretic approach to iterated epistemic change in which an epistemic change operation may be followed by a different epistemic change operation on the basis of a clear selection between the epistemic change operations. A proposed epistemic change algorithm provides the mechanism for control room agents to select between templated revision and templated update. The selection stems from the (agent-oriented) distinction between knowledge and belief: broadly speaking, templated revision is selected when newly received information is consistent with the control room agent's knowledge (though not necessarily consistent with the agent's beliefs), templated update otherwise. Under the principle of Trustworthiness, information received is trustworthy, and inconsistency with the control room agent's knowledge therefore indicates that a change of state has occurred in the system, which may be radically different from the previous state because of the discreteness of diagrammable systems. In contrast to the standard approach to belief update in which the underlying system is (implicitly) assumed to be continuous, templated update makes no such assumption, leading to unexpected results. From the (agent-oriented) distinction between knowledge and belief an alternative but comparable distinction between revision and update emerges of revision as an epistemic change operation in which the agent's knowledge grows monotonically while its beliefs may change non-monotonically and of update as an epistemic change operation in which the agent's knowledge changes non-monotonically and consequently, its beliefs too.

In recent times, research in the area of belief change has tended to shift more towards belief merging (Delgrande et al., 2005) and the field of information fusion (Grégoire and Konieczny, 2006) where the focus has been on the combination (merging) of possibly conflicting information (epistemic states). The notion of templated merging is strongly influenced by the proposal of Meyer (2001a) for merging epistemic states and takes into account results from social choice theory (Arrow, 1963; Arrow, Sen, and Suzumura, 2002). Unlike other approaches that focus on the technical similarities between merging epistemic states and aggregating preferences in social choice theory, the notion of templated merging is guided by an evaluation of the applicability of the results from social choice theory for control room agents. Templated merging is an m-ary operation on regular t-orderings that combines (possibly conflicting) information from multiple sources. It is shown to be flexible enough to capture the existing families of merging operations for both knowledge base merging and epistemic state merging. However, some of these constructions do not satisfy the principle of Qualitativeness. The importance of the principle of Qualitativeness for control room agents prompted the development of a new family of purely qualitative merging operations.

Overall, the thesis is motivated by the hope that by formalising the epistemic functioning of control room agents one may provide a way to show whether or not agents act sensibly in terms of the available information. This might help to catch errors earlier and to prevent agents from being blamed for 'errors' when in fact they were reasoning sensibly given the available information.

1.2 A reader's guide

Chapter 1 gives an overview of this thesis. It also defines the class of finitely generated transparent propositional languages that serve as knowledge representation languages for control room agents.

In chapter 2, the notion of a diagrammable system is introduced in the form of an illustrative example before defining it formally and showing that finitely generated transparent propositional languages are suitable knowledge representation languages for diagrammable systems. The notion of an agent-oriented view is outlined by reviewing some of the literature on agent theory before characterising the class of control room agents by a set of principles which guide their epistemic functioning.

The focus of chapter 3 is on the semantic representation of information. The theory of semantic information (Carnap and Bar-Hillel, 1952) is formally described, highlighting the representation of definite information in the context of diagrammable systems. Similarly, the representation of indefinite information is highlighted in the formal description of the KLM approach to nonmonotonic logic (Kraus, Lehmann, and Magidor, 1990; Lehmann and Magidor, 1992). T-orderings are formally defined as instances of an abstract datatype and it is shown that every t-ordering can be transformed into a t-ordering in normal form that is order equivalent to the original t-ordering. Subclasses of t-orderings for the exclusive representation of definite and of indefinite information are defined too. The information represented by (different subclasses of) t-orderings is shown to be expressible (as different normal forms) in the knowledge representation language and, more importantly, it is shown that from every such syntactic expression, the original t-ordering may be recovered. The notion of (semantic) information content for t-orderings is defined and compared with similar notions from semantic information theory. A preferential model semantics based on t-orderings is defined and, using the concepts of plausibility and distrust of a sentence with respect to a t-ordering (Labuschagne et al., 2002), it is shown that the defeasible consequence relations induced by this semantics satisfy several restricted forms of transitivity. As a practical example to reasoning with t-orderings, the lottery paradox is considered.

Chapter 4 is concerned with the informational attitudes of knowledge and belief. The Kripke semantics of epistemic logic (Hintikka, 1962) is formally described before recalling the information-theoretic semantics of Labuschagne and Ferguson (2002) in the context of diagrammable systems. The information-theoretic semantics is a generalisation of the Kripke semantics in which the accessibility relation is replaced by an accessibility function that assigns a t-ordering to each state of the system. The information-theoretic semantics allows for a wide range of counterparts to the axiom schemas familiar from the Kripke semantics. Other approaches in the literature that allow knowledge and belief to be modelled in the same semantics, of which there are few, are reviewed and contrasted to the information-theoretic model. The original notion of epistemic entrenchment (Gärdenfors and Makinson,1988) is formally described and the link between epistemic entrenchment and the modal operators for belief under the information-theoretic semantics recalled. A plausibility ordering on sentences of the knowledge representation language is defined in terms of the notion of plausibility (of a sentence with respect to a t-ordering) and a direct link established between plausibility orderings and epistemic entrenchment orderings. These connections provide an information-theoretic justification for representing the epistemic state of a control room agent by a regular t-ordering plus an associated belief set and knowledge set. Finally, the representation of epistemic states by regular t-orderings is contrasted to the representation of epistemic states by ordinal conditional functions.

In chapter 5, the focus is on two kinds of epistemic change, namely, revision and update. The AGM approach to belief revision (Alchourrón, Gärdenfors, and Makinson, 1985) is described with the focus on different semantic constructions of belief revision operations and followed by a description of the KM approach (Katsuno and Mendelzon, 1992) to knowledge base (belief) update. Iterated revision, the problem of dealing with a succession of revisions to the epistemic state of an agent, is described by focussing on the approach of Darwiche and Pearl (1997) and highlighting some of the more recent trends in the literature. The notion of templated revision (as an information-theoretic approach to iterated revision) is defined by proposing a set of rationality postulates, providing a representation theorem, and constructing a concrete templated revision operation. The differences between templated revision and some of the more recent proposals around iterated revision in the literature are teased out. An epistemic change algorithm is provided to distinguish between templated revision and the notion of templated update, which is defined (as an information-theoretic approach to iterated update) by proposing a set of rationality postulates, providing a representation theorem, and constructing a concrete templated update operation. Other frameworks in the literature which allow an agent to perform iterated revision and iterated update in the same system, of which there are few, are reviewed and contrasted to the templated framework. Finally, some connections between templated revision and nonmonotonic logic are explored.

Chapter 6 is concerned with the merging of (possibly conflicting) information from multiple sources. Several proposals for knowledge base (belief) merging that have been presented in the literature, based on different intuitions, are reviewed with the focus on the logical properties of merging; followed by a comparison of these proposals in terms of three families of knowledge base merging operations. At the level of epistemic states, where fewer proposals have been presented in the literature, belief merging is formally described by focussing on the proposal of Meyer (2001a) and highlighting some of the others. Technically, the problem of merging epistemic states may be viewed as similar to the problem of aggregating preferences in social choice theory (Maynard-Zhang and Lehmann, 2003). With this in mind, the conditions imposed by Arrow (1951, 1963) on the process for aggregating individual preferences into a social preference (social welfare functions) are evaluated to determine their applicability for control room agents. The notion of templated merging (as the combination of possibly conflicting information represented by regular t-orderings) is defined by proposing a set of rationality postulates and by constructing several concrete instances of abstractly defined templated merging operations. Lastly, a proposal for content-based merging is provided.

The final chapter, chapter 7, summarises the contribution of the thesis and points to areas of further research.

1.3 Finitely generated transparent propositional languages

Control room agents are assumed to express their knowledge and beliefs in transparent propositional languages that are finitely generated. As the focus is on developing new semantic concepts, the languages themselves will be kept as simple as possible. The advantage of using a transparent (rather than an opaque) language is that the atomic sentences have structure provided by predicates and constants rather than being merely propositional symbols, which allows for finer-grained expression. A finitely generated transparent propositional language is composed of the following symbols: a finite set $Cons$ of constants, a finite set $Pred$ of predicate symbol-arity pairs, the Boolean connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$, and the punctuation symbols (and).

An *atom* of L is a string of the form $P(c_1, c_2, \ldots, c_k)$ where $(P, k) \in Pred$ and each $c_1, c_2, \ldots, c_k \in Cons$. The set of atoms of L is denoted by $Atom$. Every atom is a sentence of L. If α and β are sentences of L then so too are $\neg\alpha$, $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$, $(\alpha \rightarrow \beta)$, and $(\alpha \leftrightarrow \beta)$. L is the smallest set of sentences generated from the atoms using the connectives a finite number of times as described above.

A *literal* is a sentence of L that is either an atom or the negation of an atom. A set of literals in which every atom appears exactly once, either negated or unnegated, is a *diagram* of L[1]. Thus for the languages of interest, diagrams are finite sets. Suppose L is generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. A sentence σ of L is a *state description* iff $\sigma = \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge\ \alpha_n$ where for each α_j either $\alpha_j = \rho_j$ or $\alpha_j = \neg\rho_j$. A state description may be constructed by forming a conjunction using all the literals appearing in a diagram. A sentence β of L is in *state description normal form (SDNF)* iff $\beta = \sigma_1 \vee \sigma_2 \vee \ldots \vee\ \sigma_n$ where each σ_j is a distinct state description. As metavariables, lowercase Greek letters (with or without subscripts) will be used to denote sentences of L, while uppercase Greek letters will denote sets of sentences of L. Parentheses will be omitted when writing sentences of L provided no ambiguity arises.

Associated with each language L is the traditional truth-value semantics consisting of interpretations. An *interpretation* I consists of a non-empty set D (the domain of I)

[1] See Chang and Keisler (1990) for a more formal definition.

and a denotation function den that assigns to each constant of L an element of D and to each predicate symbol of L a set of n-tuples of D where n is the arity of the predicate symbol. A *term interpretation* is an interpretation $I = (D, den)$ where $D = Cons$ and where the denotation function den associates with every constant of L the constant itself. The set of all term interpretations is denoted by U_T.

Every interpretation $I = (D, den)$ determines a unique valuation. A valuation is a function $v : Atom \rightarrow \{1, 0\}$; thus a valuation assigns a truth value to every atom. The set of all valuations of L is denoted by V. The valuation determined by an interpretation $I = (D, den)$ is denoted by v_I and obtained as follows: for every atom $P(c_1, c_2, \ldots, c_k) \in Atom$, $v_I(P(c_1, c_2, \ldots, c_k)) = 1$ iff $\langle den(c_1), den(c_2), \ldots, den(c_k)\rangle \in den(P, k)$. There is a bijective correspondence between the set V of valuations and the set U_T of term interpretations. Suppose that the set of atoms is $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. It will often be convenient to abbreviate a valuation v by writing the sequence of values $v(\rho_1)v(\rho_2)\ldots v(\rho_n)$.

A *possible worlds interpretation* of L is a pair $M = \langle S, l\rangle$ such that

- S is a non-empty set (the elements of which are called states) and
- l is a function from S to U_T (referred to as the labelling function).

If $A \subseteq Atom$, then $l(s)$ and $l(s')$ are said to *agree on* A, denoted by $l(s) \approx_A l(s')$, iff $l(s)$ and $l(s')$ assign the same truth value to every atom of A. Given a possible worlds interpretation $M = \langle S, l\rangle$, a state $s \in S$ satisfies a sentence α of L in M, denoted by $M, s \Vdash \alpha$, iff one of the following is the case:

- $\alpha \in Atom$ and $v_I(\alpha) = 1$ where $I = l(s)$
- $\alpha = \neg\beta$ and $M, s \nVdash \beta$
- $\alpha = (\beta \wedge \gamma)$ and $M, s \Vdash \beta$ and $M, s \Vdash \gamma$

- $\alpha = (\beta \vee \gamma)$ and $M, s \Vdash \beta$ or $M, s \Vdash \gamma$ or both
- $\alpha = (\beta \rightarrow \gamma)$ and $M, s \nVdash \beta$ or $M, s \Vdash \gamma$
- $\beta = (\beta \leftrightarrow \gamma)$ and $M, s \Vdash \beta$ and $M, s \Vdash \gamma$, or $M, s \nVdash \beta$ and $M, s \nVdash \gamma$

In general, the states of a possible worlds interpretation $M = \langle S, l \rangle$ may be *intensional*, by which is meant that two different states s and s' may, under l, correspond to the same valuation. In cases where the labelling function l is injective, the states of M are *extensional* in the sense that the only thing that matters is what atoms hold at that state. An *extensional* interpretation is a possible worlds interpretation $M = \langle S, l \rangle$ such that l is injective. There is historical precedent for singling out extensional interpretations for special attention. According to Ben-Naim (2005), such 'preferential structures without copies' were first singled out for attention by Hanson (1969) in the context of deontic logics. Subsequently, and of great interest to us, Shoham (1987, 1988) used them to give a semantics for nonmonotonic logics. The extensional interpretation $M_0 = \langle S, l \rangle$ for $S = U_T$ and l the identity function, which corresponds to the traditional truth-value semantics of propositional logic, will be referred to as the *classical interpretation*.

If a sentence $\alpha \in L$ is satisfied at a state s in a possible worlds interpretation $M = \langle S, l \rangle$ of L, then s is called a (local) *model* of α and α is said to be (locally) *satisfiable*. Given M, the set of models of α is denoted by $Mod_M(\alpha)$. For a set of sentences $\Gamma \subseteq L$, $Mod_M(\Gamma) = \{s \in S \mid \forall \gamma \in \Gamma,\ s \in Mod_M(\gamma)\}$. The nonmodels of α form the set $S - Mod_M(\alpha)$, denoted by $NMod_M(\alpha)$. Any sentence α such that $Mod_M(\alpha) \neq \varnothing$ is *M-satisfiable* and if $Mod_M(\alpha) = S$, then α is *M-valid*. In particular, tautologies are M-valid for every M, where by a *tautology* is understood any sentence $\top$ that is M_0-valid where $M_0 = \langle S, l \rangle$ is the classical interpretation. Any sentence α such that $Mod_M(\alpha) = \varnothing$ is *M-unsatisfiable*. In particular, contradictions are M-unsatisfiable for every M, where by a *contradiction* is understood any sentence $\bot$ that is M_0-unsatisfiable.

Given a possible worlds interpretation $M = \langle S, l \rangle$, sentences α and β are *semantically equivalent* (under M), denoted by $\alpha \equiv_M \beta$, iff $Mod_M(\alpha) = Mod_M(\beta)$. By a *proposition* is understood an equivalence class of sentences relative to $\equiv_M$. For any sentence α, the equivalence class to which it belongs is $[\alpha] = \{\beta \in L \mid \alpha \equiv_M \beta\}$.

The entailment relation (or semantic consequence relation) induced by a possible worlds interpretation $M = \langle S, l \rangle$ is the relation $\models_M$ on L given by $\alpha \models_M \beta$ iff $Mod_M(\alpha) \subseteq Mod_M(\beta)$. If $\alpha \models_M \beta$ then α is said to *entail* β (under M) and β is said to be a *semantic consequence* of α (under M). The entailment relation $\models_M$ is associated with a (unary) semantic consequence operation Cn_M on L which is defined in terms of $\models_M$ as $Cn_M(\alpha) = \{\beta \in L \mid \alpha \models_M \beta\}$. The entailment relation $\models_M$ can also be defined in terms of Cn_M by taking $\alpha \models_M \beta$ iff $\beta \in Cn_M(\alpha)$. For a set of sentences Γ, $Cn_M(\Gamma) = \{\beta \in L \mid Mod_M(\Gamma) \subseteq Mod_M(\beta)\}$. If a set of sentences Γ is M-unsatisfiable, then $Cn_M(\Gamma) = L$. A sentence α is *M-inconsistent* with a set of sentences Γ iff $\Gamma \cup \{\alpha\}$ is M-unsatisfiable.

Every semantic consequence operation Cn_M induced by a possible worlds interpretation $M = \langle S, l \rangle$ satisfies, for every sentence $\alpha, \beta \in L$, the properties of *reflexivity* (i.e. $\alpha \in Cn_M(\alpha)$), *idempotence* (i.e. $Cn_M(Cn_M(\alpha)) \subseteq Cn_M(\alpha)$), and *monotonicity* (i.e. if $\alpha \models_M \beta$ then $Cn_M(\beta) \subseteq Cn_M(\alpha)$). *Compactness* (i.e. $\alpha \in Cn_M(\Gamma)$ iff $\alpha \in Cn_M(\Gamma')$ for some finite subset $\Gamma' \subseteq \Gamma$) follows from the finiteness of L. Every semantic consequence operation Cn_M satisfies the *deduction theorem* (i.e. $\beta \in Cn_M(\Gamma \cup \{\alpha\})$ iff $(\alpha \to \beta) \in Cn_M(\Gamma)$). Lastly, every semantic consequence operation Cn_M is *supraclassical* in the sense that $Cn_{M_0}(\alpha) \subseteq Cn_M(\alpha)$ where $M_0 = \langle S, l \rangle$ is the classical interpretation.

A *theory* is a set of sentences Γ closed under entailment (under M), i.e. $\Gamma = Cn_M(\Gamma)$. For every $X \subseteq S$, the theory determined by X, denoted by $Th_M(X)$, is the set $Th_M(X) = \{\alpha \in L \mid X \subseteq Mod_M(\alpha)\}$. Suppose that $M = \langle S, l \rangle$ is a possible worlds interpretation of L and that $X \subseteq S$. If there exists a set of sentences Γ of L such that $Mod_M(\Gamma) = X$, then Γ is called an *axiomatisation* of X. In an extensional interpretation of a finitely generated language, the theory determined by X is an axiomatisation of X since $X = Mod_M(Th_M(X))$. If there exists a sentence α of L such that $Mod_M(\alpha) = X$, then α is called a *finite axiomatisation* of X.

In contexts where the classical interpretation $M_0 = \langle S, l\rangle$ is used, the semantic notions defined above will usually be given without any form of subscripting.

Proposition 1.1 *Suppose L is generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. Let $M = \langle S, l\rangle$ be an extensional interpretation of L. Then for every state $s \in S$, the set $\{s\}$ has a finite axiomatisation in the form of a unique state description.*

Proof. Pick any state $s \in S$. Let $I = l(s)$. A state description $\sigma = \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n$ is constructed by taking $\alpha_i = \rho_i$ if $v_I(\rho_i) = 1$ and $\alpha_i = \neg\rho_i$ if $v_I(\rho_i) = 0$ for each $i \leq n$. Clearly, $\{s\} = Mod_M(\sigma)$. Conversely, suppose $s' \in Mod_M(\sigma)$. Let $J = l(s')$. So for each $i \leq n$, s' satisfies α_i, i.e. if $\alpha_i = \rho_i$ then $v_J(\rho_i) = 1$ and if $\alpha_i = \neg\rho_i$ then $v_J(\rho_i) = 0$. But then $v_J = v_I$ and so $l(s') = l(s)$. Since l is injective it follows that $s' = s$. Hence $Mod_M(\sigma) \subseteq \{s\}$. But then state description σ is a finite axiomatisation of $\{s\}$. Suppose there is another state description $\sigma' = \delta_1 \wedge \delta_2 \wedge \ldots \wedge \delta_n$ such that $Mod_M(\sigma') = \{s\}$. So for each $i \leq n$, if $v_I(\rho_i) = 1$ then $\delta_i = \rho_i$ and if $v_I(\rho_i) = 0$ then $\delta_i = \neg\rho_i$. But then $\sigma' = \sigma$. Thus $\{s\}$ has a finite axiomatisation in the form of a unique state description. ■

Note that the converse of proposition 1.1 does not hold, i.e. it is not the case that for every state description σ there exists a unique state $s \in S$ such that $Mod_M(\sigma) = \{s\}$. The following example illustrates.

Example 1.1 *Suppose L is generated by $Atom = \{P(a), P'(a)\}$ and $M = \langle S, l\rangle$ is an extensional interpretation such that $S = \{s_1, s_2, s_3\}$ and $l = \{(s_1, I), (s_2, J), (s_3, K)\}$ with $v_I = 11$, $v_J = 10$, and $v_K = 01$ (where the atoms are considered in the given order, so that 10 corresponds to the valuation rendering $P(a)$ true but $P'(a)$ false). Then state* *descriptions* $P(a) \wedge P'(a)$*,* $P(a) \wedge \neg P'(a)$*, and* $\neg P(a) \wedge P'(a)$ *correspond to states* s_1*,* s_2*, and* s_3 *respectively. However, state description* $\neg P(a) \wedge \neg P'(a)$ *has no corresponding state, i.e. there is no state* $s_i \in S$ *such that* $Mod_M(\neg P(a) \wedge \neg P'(a)) = \{s_i\}$*.*

Definition 1.1 *A sentence* $\beta = \sigma_1 \vee \sigma_2 \vee \ldots \vee \sigma_n$ *of* L *in SDNF is* ***extensional*** *iff for every state description* σ_j *in* β*, there exists a corresponding state* $s_j \in S$ *such that* $Mod_M(\sigma_j) = \{s_j\}$*.*

Proposition 1.2 *Suppose* L *is generated by* $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$*. Let* $M = \langle S, l \rangle$ *be an extensional interpretation of* L*. Then every* $X \subseteq S$ *has a finite axiomatisation, and in fact, if* $X \neq \varnothing$ *then* X *has a finite axiomatisation in the form of a sentence* β *in extensional SDNF.*

Proof. If $X = \varnothing$ then $\perp$ is a finite axiomatisation of X. Suppose $X = \{s_1, s_2, \ldots, s_k\}$. By proposition 1.1, for every $s_j \in X$, the set $\{s_j\}$ has a finite axiomatisation in the form of a state description. For each $s_j \in X$, let σ_j be the state description such that $Mod_M(\sigma_j) = \{s_j\}$. Take $\beta = \sigma_1 \vee \sigma_2 \vee \ldots \vee \sigma_k$. So β is in extensional SDNF. Every state in X satisfies one of the disjuncts and thus satisfies β. So $X \subseteq Mod_M(\beta)$. Conversely, suppose $s' \in Mod_M(\beta)$. So s' must satisfy one of the disjuncts, say σ_i. Thus, by proposition 1.1, $\{s'\} = Mod_M(\sigma_i)$. But then $s' \in X$ and so $Mod_M(\beta) \subseteq X$. Thus the sentence β in extensional SDNF is a finite axiomatisation of X. ■

Propositional languages comprising infinitely many atoms have subsets of states that have no axiomatisation (Brink and Heidema, 1989). It is possible to establish this by a general cardinality argument. A more direct proof, using a transparent propositional language generated by an infinite set of atoms, is given below.

Proposition 1.3 *Suppose* L *is generated by* $Atom = \{\rho_1, \rho_2, \ldots\}$*. Let* $M_0 = \langle S, l \rangle$ *be the classical interpretation. Then there exists a set* $X \subseteq S$ *such that* X *has no axiomatisation.*

Proof. Pick any state $s \in S$ and let $I = l(s)$. Let $X = S - \{s\}$. Suppose there is some sentence α of L such that $Mod(\alpha) = X$. So $s \notin Mod(\alpha)$. Let $\{\rho_1, \rho_2, \ldots, \rho_k\}$ be such that every atom occurring in α is among $\rho_1, \rho_2, \ldots, \rho_k$. Let $s' \in S$ be the state such that $v_J(\rho_i) = v_I(\rho_i)$ for all $i \leq k$ and $v_J(\rho_i) \neq v_I(\rho_i)$ for all $i > k$ where $J = l(s')$. Then $s' \in Mod(\alpha)$ iff $s \in Mod(\alpha)$. But $s \notin Mod(\alpha)$ and so it must be the case that $s' \notin Mod(\alpha)$. But this contradicts the assumption that every state in $X = S - \{s\}$ is a model of α. Thus there is no sentence such that $Mod(\alpha) = X$. But then there can be no set Γ of sentences such that $Mod(\Gamma) = X$. So X has no axiomatisation. ■

Chapter 2

Systems and agents

Science is a system of statements based on direct experience,
and controlled by experimental verification.
The Unity of Science (Rudolf Carnap, *translation by M. Black*)

2.1 An illustrative example

As an illustrative example, the Three Mile Island (TMI) nuclear plant near Harrisburg, Pennsylvania in the USA and the sequence of events that took place during the accident of March 28, 1979 are considered. In the following brief account of the sequence of events that took place during the TMI accident, it is illustrated how the system changed state dynamically and discretely from 'normal' to 'turbine trip', to 'reactor scram', to 'LOCA', to 'core melting', to 'cooling restored', to 'hydrogen explosion'. Also indicated is how the operators (agents) reasoned from their own internal conceptual model of the system based on incomplete information about the state of the system. For a more detailed account, see Perrow (1984).

The accident in the TMI Unit 2 reactor started in the secondary cooling system when moisture entered the instrument air system of the plant causing two feedwater pumps to stop and the circulation of the secondary cooling system to halt, so that heat was no longer removed from the primary coolant. Automatically, at precisely 4:00 AM, the turbine tripped and the emergency feedwater pumps came on. However, two emergency feedwater block valves (EF-V12A and EF-V12B) that were supposed to be open had been left in a closed position after maintenance some days before, which meant that there was still no heat being removed from the core by the primary coolant, causing a rise in core temperature and pressure. The operators were unaware of any problems with emergency feedwater and, never expecting valves EF-V12A and EF-V12B to be closed, because they were always open during operation, did not immediately check the indicators on the TMI control panel which showed that the valves were closed. Besides, one of the indicators was obscured by a repair tag attached to a nearby switch.

Since no heat was being removed from the core, the reactor scrammed and graphite control rods dropped into the core to stop the chain reaction, but the residual decay heat of the reactor continued to build up temperature and pressure. Within seconds of the scram, the pilot-operated relief valve (PORV) automatically opened to relieve the pressure in the core. It failed to reseat, allowing the coolant in the core, which was under high pressure, to escape from the pressuriser through the stuck valve down into a drain tank. This was thirteen seconds into the accident. The operators believed that the PORV had reseated because the indicator on the control panel showed no warning to the contrary. It turned out that the indicator merely reflected that the PORV had received the signal to reseat.

With the loss of coolant, the pressure in the reactor went down and the temperature up, a dangerous situation that prompted the reactor coolant pumps to start up automatically followed shortly by the high-pressure injection (HPI) device which forces water into the core at a rapid rate. This was two minutes into the accident. On the control panel, one dial subsequently showed that the pressure in the reactor was still falling while another dial showed that the pressure in the pressuriser was rising. Normally, these dials move together and their readings were seen as contradictory. Believing the reactor dial would have meant that the core was being uncovered, something unheard of, while believing the pressuriser dial would have meant that the pressuriser was being flooded, called 'going solid'. Still unaware of any problem with emergency feedwater and having been diligently trained to avoid going solid, the operators decided to throttle back the HPI because if too much water is forced into the core it may cause the pressuriser to flood risking a loss of coolant accident, or LOCA. They were unaware that they were already in a LOCA. Shortly after throttling back the HPI, the reactor coolant pumps started to cavitate and were shut down by the operators thereby effectively ending forced cooling of the core.

With the arrival of a new shift at approximately 6:20 AM, the stuck PORV was discovered and operators closed a blocked valve to shut off the flow to the PORV. At that stage, incredible damage had been done and substantial parts of the core were uncovered and melting. In the control room, alarms were sounding, indicator lights were on or blinking, and radiation monitoring alarms were coming on. So many alarms were registered that the computer fell far behind schedule printing them; in fact, it took several hours before the message that something might be wrong with the PORV was finally printed. Fortunately, with the flow to the PORV shut off and valves EF-V12A and EF-V12B that had been blocking the emergency feedwater opened some time earlier, the operators were able over the next couple of hours to restore forced cooling of the reactor core.

Thirty-three hours into the accident, a muffled bang was heard in the control room registering a pressure spike in the containment building. The bang was caused by the explosion of a hydrogen bubble, which formed after the zirconium lining of the fuel rods became too hot and reacted with the surrounding water to free hydrogen by binding with oxygen. The operators could not explain the spike and wrote it off as an instrument malfunction of some sort. At the time, nuclear physicists disagreed whether hydrogen could be formed in this way. Prior to the accident, a paper mentioning the danger had been published in the Bulletin of the Atomic Scientist (Gulbransen, 1975) but the nuclear physicist who served as adviser to the company running the reactor at TMI had written a rebuttal (Palladino, 1976) denying these claims.

In the subsequent investigation into the accident, the operators were blamed for 'errors' that were actually sensible defeasible conjectures based on default rules provided by their training, as pointed out by Perrow (1984). For example, the Kemeny Commission indicated in their report (Kemeny, 1979) that the operators should have known that the PORV was open and the manufacturers of the reactor, Babcock and Wilson, agreed; 'this was the sole cause of the accident'. One of the indicators that might have alerted operators to the fact that the PORV was open was the drain tank temperature indicator (Perrow, 1984):

> "Another indicator showed the temperature of the drain tank; with hundreds of gallons of hot coolant spewing out and going to the drain tank, that temperature reading should be way up. It was indeed up. But they had been having trouble with a leaky PORV for some weeks, meaning that there was always some coolant going through it, so it was usual for it to be higher than normal. It did shoot up at one point, they noted, but that was shortly after the PORV opened, and when it didn't come down fast that was comprehensible, because the pipe heats up and stays hot. 'That hot?' a commissioner interrogating an operator asked, in effect. The operator replied, in effect, 'Yes, if it were a LOCA I would expect it to be much higher.' It was not the LOCA they were trained for on the simulators that are used for training ses-

sions, since it had some coolant coming in through an emergency system, and some coming in through HPI, which was only throttled back, not stopped. Their training never imagined a multiple accident with a stuck PORV, and blocked valves."

2.2 Diagrammable systems

Consider the collection of all raptors in Southern Africa. This collection of birds is unstable in the sense that birds come and go over a relatively short period of time due to the natural occurrence of birth and death and due to the seasonal migration of species such as the Black Kite and the Booted Eagle. The collection is anonymous in the sense that it is impossible to name or ring each individual bird in the collection. As a system, this collection of birds is undiagrammable. One consequence of this is that if one were to attempt to describe a state of the system, it would be difficult to decide where to draw the line, since so much could arguably influence the state: weather, the use of poisons, automobiles, the growth of the human population.

Consider, in contrast, a museum collection of raptors in Southern Africa, say, for example, the collection in the Albany Museum in Grahamstown, South Africa. This collection of birds is stable over a relatively long period of time and it is possible to distinguish and label each individual bird in the collection. As a result it is easier to demarcate relevance: the fact that specimen #113 is a Lammergeier (or Bearded Vulture) shot on the farm Moddervlei on June 6, 1946 is relevant and contributes to determining the state of the collection; facts such as that farmers have recently begun using a new pesticide, or even facts closer to home such as that the Albany museum was founded on September 11, 1855 are irrelevant as far as the state of the collection of raptors is concerned. The museum collection of raptors is an example of a diagrammable system, and the states of such a system are determined by the properties of and relationships between components of the system, rather than by external factors.

Definition 2.1 *A system is **directly representable** if*

- *each component is sufficiently stable and distinctive to be representable by a constant symbol of the knowledge representation language, and*
- *a state of the system is sufficiently determined by the properties of and relationships between components to be representable by the atoms of the knowledge representation language that hold in that state.*

Definition 2.2 *A **diagrammable system** is a directly representable system in which there are finitely many components.*

A nuclear power plant is a diagrammable system in the sense of having a stable configuration of finitely many components that are typically identified and named in a collection of complex drawings and diagrams. Diagrammable systems may be partially observable or fully observable, dynamic or static, discrete or continuous, nondeterministic or deterministic, sequential or episodic, multi-agent or single-agent, and cooperative or competitive[1]. Following Russell and Norvig (2003) a system is said to be

- *partially observable* if the agent's sensory apparatus provides incomplete information about the state of the system, and *fully observable* if it provides complete information about the state of the system;
- *dynamic* relative to the agent if the system can change state independently of the agent's actions, and *static* otherwise;

[1] In the first edition (Russel and Norvig, 1995), the terms 'inaccessible' and 'accessible' are used instead of 'partially observable' and 'fully observable', and 'nonepisodic' instead of 'sequential'.

- *discrete* if a change of state can result in a state of the system that differs radically from its predecessor, and *continuous* if changes in state are gradual;
- *nondeterministic*[2] to the agent if the next state of the system is not completely determined by the current state and the action of the agent, and *deterministic* otherwise;
- *sequential* if the agent's current decision and action could affect all future decisions, and *episodic* if it affects only the current episode of perception, decision, and action;
- *multi-agent* if more than one agent can interact with the system, and *single-agent* otherwise; and
- *cooperative* if the nature of the system demands cooperation between agents, and *competitive* if it encourages competition between agents.

Our focus will be on the class of diagrammable systems in which every system is partially observable, dynamic, discrete, sequential, multi-agent, and cooperative. A nuclear power plant is a positive example of such a diagrammable system.

In a diagrammable system, states of the system can be mapped injectively to valuations of the knowledge representation language. Hence the importance, for our purposes, of extensional interpretations. However, the mapping may not be direct and obvious. Consider a device containing n balls, with each ball painted one of k colours, which operates by changing the colouring of the balls in a random fashion. Consider the scenario in which the balls are distinguishable, say by virtue of a number painted on them in an invariant colour different from the k colours. An intuitive conceptualisation would have each state of the system represented by an n-tuple whose coordinates may have any of the k colour values. This natural representation provides a total of k^n states. However,

[2] *The term 'stochastic' is preferred over 'nondeterministic' in Russel and Norvig (2003).*

an (obvious) knowledge representation language for this system should allow one to say things like 'Ball j has colour m' and so there must be nk atoms and thus 2^{nk} valuations. Not only are the number of states and the number of valuations different, but it may take a little thought to decide which valuation corresponds to a particular state. For example, a state $\langle$red, blue$\rangle$ corresponds to the valuation which makes the atom 'Ball 1 is red' true, all other atoms talking about ball 1 false, the atom 'Ball 2 is blue' true, and all other atoms involving ball 2 false.

Finitely generated transparent propositional languages are suitable knowledge representation languages for diagrammable systems. The stability and distinctiveness of components mean that the systems lend themselves to the definitions of the sets $Cons$ and $Pred$ and that term interpretations suffice. To make explicit the correspondence between states of the system and valuations of the knowledge representation language, possible worlds interpretations are used. And since a state of a diagrammable system is determined by the elementary facts about the components, the possible worlds interpretations may be restricted to extensional interpretations.

Proposition 2.1 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Every state of a diagrammable system is axiomatisable by a unique (finite) diagram.*

Proof. Pick any state $s \in S$. By proposition 1.1, the set $\{s\}$ is finitely axiomatisable by a unique state description, say $\beta = \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n$. Thus $\{s\} = Mod_M(\beta)$. A (finite) diagram is constructed by taking $\Gamma = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. But $Mod_M(\beta) = Mod_M(\Gamma)$ and so $\{s\}$ is finitely axiomatisable by a unique (finite) diagram. ■

While a diagram has been defined to be a certain kind of set of sentences, ordinary language uses 'diagram' for a kind of stylised picture, in some contexts called a blueprint. The picture and the sentential diagram are closely related. Whether sentential or pictorial, the purpose of a diagram is to display at least some of the relevant facts that determine states of the system. From the pictorial version one can extract a set of invariant facts or constraints, limiting the realisable states of the system.

By way of example, consider a full adder, that is, an adder circuit of a digital computer that can handle a carry signal as well as the binary elements that are to be added. One possible configuration of a full adder is depicted by the circuit diagram in figure 2-1.

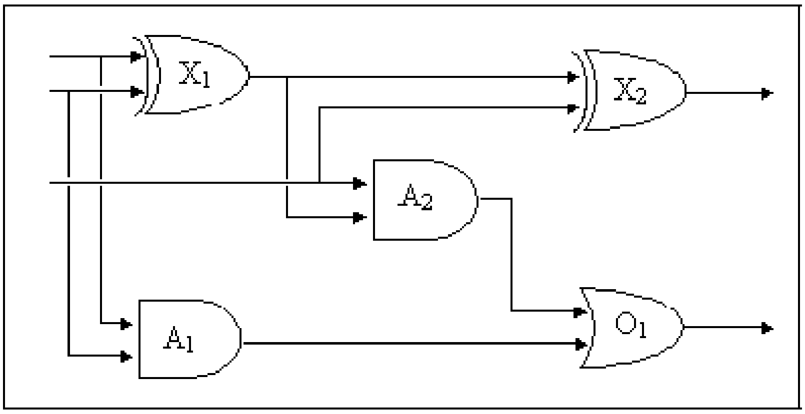

Figure 2-1: *Full adder*

The circuit diagram serves as the blueprint of the full adder: it identifies components X_1 and X_2 as XOR-gates, components A_1 and A_2 as AND-gates, and component O_1 as an OR-gate, and it shows how these components fit together, for example, it shows that the first input to component A_2 is the same as the second input to component X_2. From the circuit diagram it is clear that certain states, for example, states in which the output from component X_1 differs from either the first input to component X_2 or the second input to component A_2, are not realisable.

In focussing on a subset of invariant facts or constraints, a system diagram (or blueprint) typically provides an agent with information about those states of the system that are unrealisable.

2.3 Agents in context

An *agent* is an entity capable of interacting with its environment and changing it. Pollock (1995) memorably contrasts agents and rocks. Each is a stable entity in a dynamically changing environment. But whereas a rock achieves its stability passively, by virtue of being hard to destroy, an agent achieves its stability by interacting with its surroundings and changing these. An agent may be human or non-human, with non-human agents typically being hardware agents (robots) or software agents.

Software agents, as defined by Genesereth and Ketchpel (1994), are software components that communicate with one another by exchanging messages in a standardised agent communication language (ACL). The components of an ACL were originally defined as part of the DARPA Knowledge Sharing Effort (Neches et al., 1991) as consisting of a vocabulary, a Knowledge Interchange Format (KIF) language, and a Knowledge Query and Manipulation Language (KQML). KIF is based on first-order predicate logic and allows for the expression of simple information (Genesereth and Fikes, 1992) while KQML, which is based on speech acts, provides a linguistic layer for dialogue between sender and receiver (Finin et al., 1994, 1997). A critique of KQML may be found in Cohen and Levesque (1995, 1997). The effort of standardising agent communication languages has subsequently been taken over by FIPA, the Foundation of Intelligent Physical Agents, which is a standards body concerned with the interoperation of heterogeneous software agents. For a more recent collection of papers on software agents, see Bradshaw (1997) or Huhns and Singh (1998).

The concept of a software agent can be traced back to the Actor formalism of Hewitt (1977). Based on Hewitt's Actor model, Shoham (1990, 1993) introduced the computational framework of *agent-oriented programming* (*AOP*) as a specialisation of object-oriented programming (OOP). The basic units of computation in OOP are modules (objects) that communicate with one another by passing messages and that respond to incoming messages by using individual methods. In AOP, objects are specialised to agents that have mental states, which in turn consist of components such as beliefs, capabilities, choices, and commitments. (The tradition in AI of ascribing mental qualities to computer programs dates back to McCarthy (1979).) Agents communicate with each other by performing communicative acts such as informing, requesting, offering, accepting, and rejecting as opposed to passing unconstrained messages. These communicative acts are based on the theory of speech acts which, based on earlier work by Grice (1957), was introduced by Austin (1962) and further developed by Searle (1969, 1979). In AOP, agents no longer merely respond to incoming messages as in OOP but, with greater independence, are able to plan and act autonomously by using their mental components according to individual constraints. A more recent survey of agent-oriented programming may be found in Shoham (1997) or Thomas (1999).

A computer program endowed with mental attitudes may be seen as an *intentional system* in the sense of Dennett (1987). Dennett distinguishes between first-order and second-order intentional systems. A *first-order intentional system* has, for example, beliefs and desires but no beliefs and desires about beliefs and desires while a *second-order intentional system* has, in addition to, say, beliefs and desires, also beliefs and desires about beliefs and desires. Mental attitudes may be partitioned into four broad categories of which informational and motivational attitudes are, for the applications envisaged for agent-oriented programming, the most important, with social and emotional attitudes less so (Shoham and Cousins, 1994).

Informational attitudes relate to the information that an agent has about its environment and consists of knowledge and belief. The standard philosophical literature attaches more importance to knowledge than belief (Hintikka, 1962; Halpern, 1986, 1995) as is compatible with a 'true justified belief' account of knowledge (Schmitt, 1992) in which belief is merely the germ from which knowledge may spring. Until recently, very few formalisations in logic allowed the coexistence of knowledge and belief in the same language and semantics. A notable exception is the account given by Moses and Shoham (1993) in which belief is viewed as 'defeasible' knowledge.

Goals, desires, plans, intentions, and commitments are examples of *motivational attitudes* that relate to the way in which an agent makes decisions. In an influential paper, Cohen and Levesque (1990) formalised the interrelationship between beliefs, goals, and intentions by adopting Bratman's philosophical theory of intention (Bratman, 1987). Rao and Georgeff (1991) provided an alternative formalisation in which intention is defined, not implicitly as a commitment to goals, but explicitly as consistent with an agent's goals (or desires). This formalisation was further developed over a series of papers and culminated in their paper (Rao and Georgeff, 1998) on systems of belief-desire-intention (BDI) logics.

Social attitudes relate to the social and moral reasons (or obligations) for agents to behave in a certain way. Although the notions of obligation and permission have been formalised in philosophy in the context of deontic logic (von Wright, 1951, 1963; Åqvist, 1984, 2002; Makinson, 1999), the formalisation of social attitudes within the context of agent theory is still emerging (Castelfranchi, 1998; van der Torre, 2003).

The formalisation of the *emotional attitudes* of an agent, which relate to the influence of emotional analogs on the behaviour of the agent in achieving its goals, has received even less attention. Reilly and Bates (1992) were among the first to informally investigate the notion of emotional agents, largely based on the cognitive model of emotion of Ortony, Clore, and Collins (1988). The work of Picard (1997) on affective computing, influenced by research in neuroscience (Damasio, 1994; LeDoux, 1996) and cognitive science (Simon, 1967; Oatley and Johnson-Laird, 1987), has sparked renewed interest in the formalisation of emotional attitudes within the context of agent theory.

Most of the work in agent theory is concerned with giving a good account of the interrelationships between various mental attitudes and an early survey of different approaches may be found in Wooldridge and Jennings (1995). More recent investigations into the theory of rational agents may be found in Wooldridge and Rao (1999), Wooldridge (2000), and van der Hoek and Wooldridge (2003). It should be noted that a desire to render tractable the notion of accountability of agents has led to FIPA-sponsored attempts to devise useful models of agents that dispense with the epistemic components, namely, knowledge and belief. The motivation for doing so is set out in Singh (1998). Since our concern is with the epistemic functioning of agents, we remain true to the AOP tradition as conceived by Shoham.

2.4 Control room agents

Control room agents are agents whose environments are the class of diagrammable systems in which every system is partially observable, dynamic, discrete, sequential, multi-agent, and cooperative. They may be human agents or software agents. Whether human or non-human, the class of *control room agents* may be described as first-order intentional systems having both informational attitudes and motivational attitudes, with the focus of this thesis exclusively on the informational attitudes of knowledge and belief. To provide a broader framework for describing the informational attitudes of control room agents, we briefly consider one specific mental model, namely, the *Decision Ladder* as proposed by Rasmussen (1986) in the context of cognitive engineering. It is a mental model applicable for process control operators and therefore suitable for investigating control room agents.

The Decision Ladder focuses on the mental processing performed by control operators without making any assumptions as to the nature of the underlying information processes themselves. Eight information processes are identified. The first three processes relate to the sensing, observation, and identification of the current state of the system. The next two processes concern the interpretation and evaluation of goals to determine the desired state of the system. The last three processes bear on the selection, planning, and execution of tasks to effect the necessary change in the state of the system.

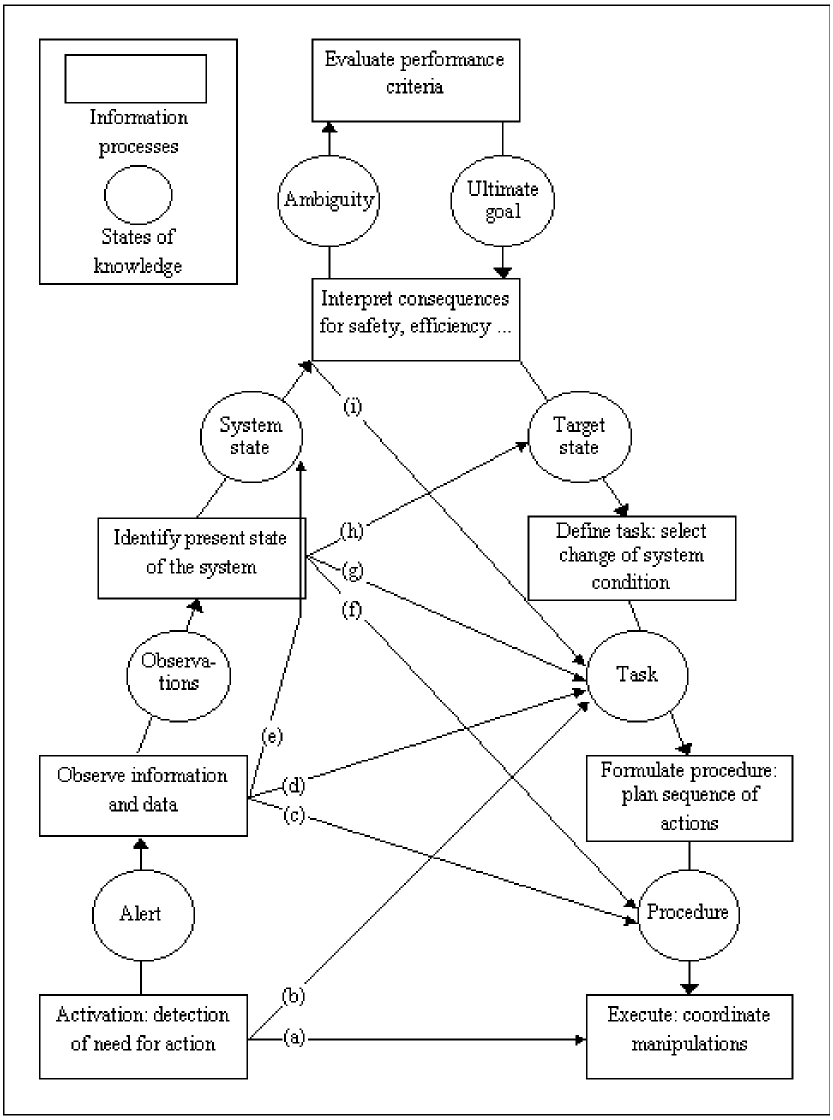

Figure 2-2: *Rasmussen's Decision Ladder*

The outcome of an information process is a ‘state of knowledge’, which may differ depending on the experience of the operator and the uncertainty of the situation, resulting in different sequences of mental processing. The unlabelled arcs in figure 2-2 show the sequence typically associated with decision making in unfamiliar situations. In more familiar situations, solutions from previous experiences are used resulting in various shortcut sequences as indicated by the labelled arcs.

Within the broader framework provided by the Decision Ladder, the focus on the informational attitudes of knowledge and belief may be viewed as concentrating on the ‘states of knowledge’ and on the information processes related to the identification of the current state of the system. In other words, the focus is on the representation of a control room agent’s ‘epistemic state’ and on the processes of changing the agent’s epistemic state based on new information.

Control room agents are further characterised by a set of principles which guide their epistemic functioning. This characterisation may be viewed as a specialisation of more general principles for rational (ideal) agents. A distinction is made between principles guiding the representation of a control room agent’s epistemic state, still to be defined, and principles guiding the manner in which a control room agent changes its epistemic state upon receiving new information.

As specific principles for guiding the representation of a control room agent’s epistemic state, the principle of Duality is proposed as a specialisation for control room agents while the general principle of Logical Closure is adopted.

- *Principle of Duality*

The principle of Duality says that the epistemic state of a control room agent should allow for the representation of both knowledge and belief and that knowledge should constitute more than mere tautologies. It is strongly motivated by the argument of Girle

(1998) that philosophico-logical systems (agents) should distinguish between knowledge and belief and, moreover, in a way that ensures it is not only logical and mathematical propositions (tautologies) that can be known.

- *Principle of Logical Closure*

The principle of Logical Closure says that if a sentence is expressing the current knowledge and beliefs of an agent in an epistemic state, then the logical consequences of the sentence should also be an expression of the current knowledge and beliefs of the agent in the epistemic state. Formulated as a rationality criterion (Gärdenfors, 1988), it is generally viewed as a minimal requirement on representations of epistemic states.

The general principle of Consistency is applicable to both epistemic state representations and changes, and is adopted for control room agents. The proposed principle of Qualitativeness is a key specialisation for control room agents and advocates a qualitative approach. It, too, is applicable to both epistemic state representations and changes.

- *Principle of Consistency*

The principle of Consistency says that the set of sentences expressing an agent's current knowledge and beliefs in an epistemic state should be consistent (where we may take 'consistent' to be a synonym for 'satisfiable'). This is the other rationality criterion that has been proposed as a minimal requirement on representations of epistemic states (Gärdenfors, 1988). In the context of epistemic change, it says that if an epistemic state is consistent before an epistemic change has taken place, then, provided the new information is satisfiable, the epistemic state after the change has taken place should be consistent too.

- *Principle of Qualitativeness*

The principle of Qualitativeness says that the representation of a control room agent's epistemic state should not depend on arithmetic and neither should any changes to its epistemic state depend thereon. It is motivated by results from cognitive science (Tversky and Kahneman, 1974, 1983; Kahneman, Slovic, and Tversky, 1982) where it has been shown that ordinary reasoning often violates the principles of probability theory and instead employs heuristics such as representativeness (stereotypes), availability of scenarios, and anchoring. Arguments against the use of numbers (Piattelli-Palmarini,1994; Gigerenzer, 2002) provide additional support for a qualitative model of thinking.

The fundamental problem underlying any quantitative model of thinking is 'Where do the numbers come from?' In the case of a large, complex, and very expensive system (e.g. a nuclear powerplant), it is simply not possible to experiment with various scenarios and set up the relative frequency tables that are required if a probability measure is to be based on anything more than intuition. (We do not include simulations under the term 'experiment', for simulations must generally assume the probabilities that the experiments would be intended to reveal.) In the absence of reliable statistics, we may as well accept that intuition is the best we can do, and proceed to model it. Given that intuition provides not numerically precise values such as 17.835% but softer indefinite indications such as the qualitative classification 'rather unlikely', it follows that a non-numerical model is sufficient.

To drive the point home, consider the hydrogen explosion that occurred during the Three Mile Island disaster. Potentially, such an explosion could breach containment and lead to a spread of contamination. What did the control room operators know about the likelihood of such explosions? At the time, nuclear physicists had formed no consensus as to whether the hydrogen formation could occur. There was complete ignorance of any probability for the event. Being ignorant of the probability is not the same as assuming the probability to be 50%, and the event therefore had no numerical probability that could justifiably be associated with it. Nevertheless, the operators were faced with data to which they needed to respond. Any analysis seeking to account for the fact that the operators wrote the explosion off "as an instrument malfunction of some sort" must take into account their (non-statistical) predisposition to regard such an explosion as impossible or very unlikely. This predisposition can be explained by recalling that it reflected the opinion of the scientist who served as official adviser to the company employing the operators.

Such informed opinions are a common source of predispositions forming the basis for action in many everyday situations. Generally, such opinions are in the nature of a categorisation, and the category contains members that are either more or less typical (Rosch et al., 1976). The principle of Qualitativeness may therefore be regarded as an undertaking to provide models of reasoning that make no assumptions beyond the assumption that control room agents are willing and capable of categorisation.

Adhering to the principles of Duality, Logical Closure, and Consistency imply that control room agents will be capable of distinguishing between knowledge and belief, capable of knowing all the consequences of their knowledge and believing all the consequences of their beliefs, and incapable of believing (or knowing) anything which is false. Adhering to the principle of Qualitativeness implies that the formation and change of control room agents' knowledge and beliefs will be modelled by non-arithmetic mechanisms rather than assuming the availability of numerical measures as in probability theory or fuzzy sets theory.

The manner in which a control room agent changes its epistemic state upon receiving new information will, likewise, be guided by a number of specific principles. The following general principles are adopted:

- *Principle of Minimal Change*

The principle of Minimal Change says that an epistemic change operation should change an epistemic state as little as necessary to accommodate the new information.

- *Principle of Informational Economy*

The principle of Informational Economy is related to the principle of Minimal Change. It says that an epistemic change operation should, in the process of ensuring that an epistemic state remains consistent, remove as little information as possible from the epistemic state.

- *Principle of Irrelevance of Syntax*

The principle of Irrelevance of Syntax says that an epistemic change operation should be independent of the syntactic form of an epistemic state and of the new information. The principle was first formulated by Dalal (1988).

- *Principle of Success*

The principle of Success says that an epistemic state should contain the new information after an epistemic change has taken place.

- *Principle of Categorical Matching*

The principle of Categorical Matching says that the representation of an epistemic state after an epistemic change has taken place should be of the same format as the representation of the epistemic state before the change. The principle is sometimes referred to as the principle of *Adequacy of Representation* (Dalal, 1988) but in the present formulation, it is generally attributed to Gärdenfors and Rott (1995). While almost too obvious, it is particularly relevant when iterations of epistemic change are at stake.

The principles of Consistency, Minimal Change, Irrelevance of Syntax, and Success are often adopted as 'basic' principles (Dubois and Prade, 1998). However, none of these principles constrain the new information in any way and should the new information be a contradiction, the agent would be obliged by the principle of Success to accept the contradiction, making its epistemic state inconsistent. While this may be appropriate in some domains, in the context of diagrammable systems and control room agents it is not.

The principle of Trustworthiness is proposed as a specialisation for control room agents to guide epistemic state changes and supports Friedman and Halpern's (1999) argument against accepting $\perp$ as new information.

- *Principle of Trustworthiness*

The principle of Trustworthiness says that the information received by a control room agent is trustworthy.

Under the principle of Trustworthiness a control room agent would neither observe $\perp$ nor receive the communication $\perp$ from another agent. But the principle makes an even stronger assumption. Firstly, it assumes that the control room agent would not hear from another agent that a component of the system is, say on, when in fact it is (believed by the other agent to be) off. Given the cooperative nature of diagrammable systems, this is realistic. Secondly, the principle makes the assumption that the control room agent would not observe that a component of the system is, say on, when in fact it is off. Whilst certainly possible for human senses (or artificial sensors) to fail at times, these are ordinarily reliable, and it does not seem unreasonable to adopt as a convenient simplification the assumption that a control room agent can trust its senses (or sensors).

Having established the notion of a control room agent, it will be convenient henceforth to refer to a control room agent simply as an agent.

Chapter 3

Semantic representations

'If there's no meaning in it,' said the King, 'that saves a world of trouble, you know, as we needn't try to find any. ...'

Alice's Evidence in *Alice's Adventures in Wonderland* (Lewis Carroll)

3.1 Semantic information theory

The basic intuition underlying semantic information theory is that the more information an agent has about a system the more states the agent is able to exclude, an intuition that dates back to Popper's idea that the more a statement forbids, the more it says about the world of experience (Popper, 1934 - see the opening sentence of Section 35). In this chapter we shall outline formalisations that rest on the intuition that a state may be ruled out of consideration either definitely or tentatively. The former case is that of definite information (or knowledge), while the latter brings us to indefinite information and the use of default rules to form (defeasible) beliefs. By focussing on semantic formalisations, we shall pull together apparently disparate strands from semantic information theory and nonmonotonic logic as the cornerstones for the unifying framework of templated orderings.

The *theory of semantic information* was formulated by Carnap and Bar-Hillel over a series of papers (Carnap and Bar-Hillel, 1952; Bar-Hillel and Carnap, 1953; and Bar-Hillel, 1955) in an attempt to address the semantic aspects of information, which were not taken into account in the mathematical theory of communication (Shannon, 1949). In its original form, the theory of semantic information applies to transparent propositional languages[1] generated by a finite number of atoms. In this convenient form, semantic information theory thus applies to diagrammable systems where the states of the system are of course determined by a finite number of elementary facts about the components that may or may not hold. In semantic information theory, it is assumed that states which are excluded, are ruled out definitely rather than tentatively. In this sense, semantic information theory deals with information which is definite.

Definite information is information that is 'hard' or 'certain', and from which conclusions may be drawn which are undisputable. It can be expressed by sentences of the knowledge representation language and is provided by *fixed information*, typically in the form of a system diagram (or blueprint), and *evidential information*, typically in the form of both observations and communications from trusted agents. Whereas fixed information is state-independent, evidential information is state-dependent. As an example of definite information, the following extract from the Three Mile Island account, provided earlier, is considered:

> "On the control panel, one dial subsequently showed that the pressure in the reactor was still falling while another dial showed that the pressure in the pressuriser was rising. Normally, these dials move together and their readings were seen as contradictory".

[1] The theory of semantic information also applies to predicate logic in an extended form developed by Hintikka (see, for example, Hintikka (1970)).

In this example, the definite information, in the form of an (indirect) observation, is firstly, that the pressure in the reactor was falling and secondly, that the pressure in the pressuriser was rising.

A semantic representation of definite information may be accomplished by a division of the finite set of states (each of which corresponds to a valuation of the language) into a set of included or positive states and a complementary set of excluded or negative states, providing a kind of dichotomous 'partition' on the set of states. (Visually, this semantic representation may be pictured by two boxes one of which is positioned on top of the other, the lower box containing the set of included states and the upper box the set of excluded states.) In keeping with the intuition behind semantic information theory, a lack of information is indicated by the exclusion of no state and too much (or contradictory) information by the exclusion of every state. An agent has complete information about the system when it is able to rule out all but one state of the system. Under the principle of Trustworthiness of information, the remaining non-excluded state would have to be the actual state of the system. (We are considering only the definite ruling out of states at this point.) However, for most of the time an agent has to deal with information that is incomplete.

Example 3.1 *Consider an agent observing a simple Light-Fan system where each component may be on or off. As a diagrammable system, the Light-Fan system may be represented by a transparent propositional language L generated by $Atom = \{P(a), P(b)\}$ where P is the predicate representing the property that a component is on and a and b are the constants representing the components the light and the fan respectively. The states of the system, represented by $S = \{11, 10, 01, 00\}$, correspond directly to the valuations of L, with 10 corresponding to the valuation making $P(a)$ true and $P(b)$ false, and so on. Suppose the agent can see only the light and observes that it is on. This example of definite information may be represented visually as shown in figure 3-1.*

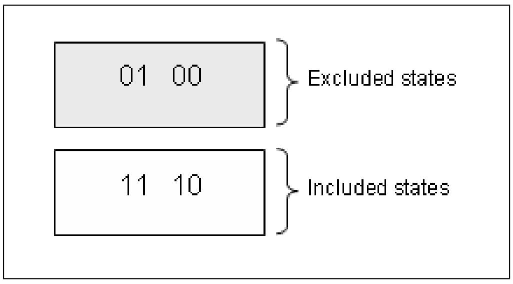

Figure 3-1: Definite information

A syntactic expression of definite information may be extracted from the semantic division in either of two ways, since specifying either the set of included states or the set of excluded states suffices to determine the remaining set. Inclusion of a state may be expressed by a state description since, by proposition 1.1, for every state $s \in S$, there exists a unique state description $\sigma \in L$ such that $Mod_M(\sigma) = \{s\}$ where $M = \langle S, l \rangle$ is an extensional interpretation of L. Exclusion of a state may be indicated by a corresponding content element: a sentence that is the negation of a state description and, by application of De Morgan's identities, a disjunction of atoms in which every atom appears exactly once, either negated or unnegated. It will subsequently be shown that for every state $s \in S$, there exists a unique content element $\varepsilon \in L$ such that $NMod_M(\varepsilon) = \{s\}$ where $M = \langle S, l \rangle$ is an extensional interpretation of L.

Definition 3.1 *Suppose L is a transparent propositional language generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. A sentence $\varepsilon \in L$ is a **content element** iff $\varepsilon = \alpha_1 \vee \alpha_2 \vee \ldots \vee \alpha_n$ where for each α_j either $\alpha_j = \rho_j$ or $\alpha_j = \neg\rho_j$. A sentence $\beta \in L$ is in **semantic content normal form (SCNF)** iff $\beta = \varepsilon_1 \wedge \varepsilon_2 \wedge \ldots \wedge \varepsilon_m$ where each ε_j is a distinct content element.*

Proposition 3.1 *Suppose L is a transparent propositional language generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. Let $M = \langle S, l \rangle$ be an extensional interpretation of L. Then for every* *state $s \in S$, there exists a unique content element $\varepsilon \in L$ such that $NMod_M(\varepsilon) = \{s\}$.*

Proof. Pick any state $s \in S$. Let $I = l(s)$. A content element $\varepsilon = \alpha_1 \vee \alpha_2 \vee \ldots \vee \alpha_n$ is constructed by taking $\alpha_i = \rho_i$ if $v_I(\rho_i) = 0$ and $\alpha_i = \neg\rho_i$ if $v_I(\rho_i) = 1$ for each $i \leq n$. Clearly, $\{s\} \subseteq NMod_M(\varepsilon)$. Conversely, suppose $s' \in NMod_M(\varepsilon)$. Let $J = l(s')$. So for each $i \leq n$, s' fails to satisfy α_i, i.e. if $\alpha_i = \rho_i$ then $v_J(\alpha_i) = 0$ and if $\alpha_i = \neg\rho_i$ then $v_J(\alpha_i) = 1$. But then $v_J = v_I$ and so $l(s') = l(s)$. Since l is injective, it follows that $s' = s$. Hence $NMod_M(\varepsilon) \subseteq \{s\}$. But then $NMod_M(\varepsilon) = \{s\}$. Suppose there is another content element $\varepsilon' = \delta_1 \vee \delta_2 \vee \ldots \vee \delta_n$ such that $NMod_M(\varepsilon') = \{s\}$. So if $v_I(\rho_i) = 0$ then $\delta_i = \rho_i$ and if $v_I(\rho_i) = 1$ then $\delta_i = \neg\rho_i$ for each $i \leq n$. But then $\varepsilon' = \varepsilon$. So for every state $s \in S$, there exists a unique content element $\varepsilon \in L$ such that $NMod_M(\varepsilon) = \{s\}$. ■

Note that the converse of proposition 3.1 does not hold, i.e. it is not necessarily the case that for every content element $\varepsilon \in L$ there exists a unique state $s \in S$ such that $NMod_M(\varepsilon) = \{s\}$. The following example illustrates.

Example 3.2 *Suppose L is generated by $Atom = \{P(a), P'(a)\}$ and $M = \langle S, l\rangle$ is an extensional interpretation such that $S = \{s_1, s_2, s_3\}$ and $l = \{(s_1, I), (s_2, J), (s_3, K)\}$ with $v_I = 11$, $v_J = 10$, and $v_K = 01$ (where the atoms are considered in the given order, so that 10 corresponds to the valuation rendering $P(a)$ true but $P'(a)$ false). Then content elements $\neg P(a) \vee \neg P'(a)$, $\neg P(a) \vee P'(a)$, and $P(a) \vee \neg P'(a)$ correspond to states s_1, s_2, and s_3 respectively. However, content element $P(a) \vee P'(a)$ has no corresponding state, i.e. there is no state $s_i \in S$ such that $NMod_M(P(a) \vee P'(a)) = \{s_i\}$.*

Definition 3.2 *A sentence $\beta = \varepsilon_1 \wedge \varepsilon_2 \wedge \ldots \wedge \varepsilon_m$ of L in SCNF is* ***extensional*** *iff for every content element ε_j in β, there exists a corresponding state $s_j \in S$ such that $NMod_M(\varepsilon_j) = \{s_j\}$.*

Proposition 3.2 *Suppose L is generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. Let $M = \langle S, l \rangle$ be an extensional interpretation of L. Then every $X \subset S$ has a finite axiomatisation in the form of a sentence β in extensional SCNF.*

Proof. Let $Y = S - X$. So $Y \neq \varnothing$. Suppose $Y = \{s_1, s_2, \ldots, s_k\}$. By proposition 3.1, for every $s_j \in Y$, there exists a unique content element ε such that $NMod_M(\varepsilon) = \{s_j\}$. For each $s_j \in Y$, let ε_j be the content element such that $NMod_M(\varepsilon_j) = \{s_j\}$. Take $\beta = \varepsilon_1 \wedge \varepsilon_2 \wedge \ldots \wedge \varepsilon_k$. So β is in extensional SCNF. But now every state in Y fails to satisfy one of the conjuncts and thus fails to satisfy β. So $Y \subseteq NMod_M(\beta)$. Conversely, suppose $s' \in NMod_M(\beta)$. So s' must fail to satisfy one of the conjuncts, say ε_i. Thus, by proposition 3.1, $\{s'\} = NMod_M(\varepsilon_i)$. But then $s' \in Y$ and so $NMod_M(\beta) \subseteq Y$. Thus $NMod_M(\beta) = Y$ from which it follows that $Mod_M(\beta) = X$. But then β, which is a sentence in extensional SCNF, is a finite axiomatisation of X. ■

For a sentence to be an accurate expression of the semantic representation of definite information it must axiomatise the set of included states. Let C be the set of included states and $\overline{C}$ the set of excluded states. By taking the state descriptions corresponding to C, a sentence in extensional SDNF may be formed which axiomatises C as shown in proposition 1.2 (provided C is non-empty). Similarly, by taking the content elements corresponding to $\overline{C}$, a sentence in extensional SCNF may be constructed which axiomatises C as illustrated in proposition 3.2 (provided $\overline{C}$ is non-empty). So for every semantic representation of definite information as a division of the finite set of states into a set of included states and a complementary set of excluded states there exists a syntactic expression in the form of a sentence in extensional SDNF or SCNF or both. Conversely, every sentence $\alpha \in L$ induces a semantic representation of definite information by dividing the finite set of states into a set of states included by α (the models of α) and a set of states excluded by α (the nonmodels of α).

Proposition 3.3 *Suppose L is a transparent propositional language generated by $Atom = \{\rho_1, \rho_2, \ldots, \rho_n\}$. Let $M_0 = \langle S, l \rangle$ be the classical interpretation. Then for every content element $\varepsilon \in L$, if $\varepsilon \models_M \alpha$, then either $\alpha \equiv_M \top$ or $\alpha \equiv_M \varepsilon$ (but not both).*

Proof. Pick any content element $\varepsilon = \alpha_1 \vee \alpha_2 \vee \ldots \vee \alpha_n \in L$ and suppose that $\varepsilon \models_M \alpha$. It must be shown that either $Mod(\alpha) = S$ or else $Mod(\alpha) = Mod(\varepsilon)$ but $Mod(\varepsilon) \neq S$. The proof relies on the result that under the classical interpretation, for every content element $\varepsilon \in L$ there exists a unique state $s \in S$ such that $NMod(\varepsilon) = \{s\}$. To see this, we construct a valuation v_I, where $I = l(s)$, by taking $v_I(\rho_i) = 0$ if $\alpha_i = \rho_i$ and $v_I(\rho_i) = 1$ if $\alpha_i = \neg\rho_i$ for each $i \leq n$. So for each $i \leq n$, s fails to satisfy α_i and thus s fails to satisfy ε, i.e. $s \in NMod(\varepsilon)$. Suppose there is another state $s' \in NMod(\varepsilon)$. Let $J = l(s')$. So for each $i \leq n$, s' fails to satisfy α_i, i.e. if $\alpha_i = \rho_i$ then $v_J(\alpha_i) = 0$ and if $\alpha_i = \neg\rho_i$ then $v_J(\alpha_i) = 1$. But then $v_J = v_I$ and so $l(s') = l(s)$. But under the classical interpretation, $S = U_T$ and l is the identity function. So $s' = s$ and thus $NMod(\varepsilon) = \{s\}$. But then $Mod(\varepsilon) = S - \{s\}$. Thus $Mod(\varepsilon) \neq S$. Since, by assumption, $Mod(\varepsilon) \subseteq Mod(\alpha)$, it then follows that either $Mod(\alpha) = S$ or $Mod(\alpha) = S - \{s\}$. ■

Under the traditional truth-value semantics of finitely generated propositional languages, a content element is the weakest sentence that conveys information or, alternatively, the smallest syntactic unit of information, as suggested by proposition 3.3. Using this property, Carnap and Bar-Hillel (syntactically) defined the *information carried by a sentence* to be the set of content elements entailed by the sentence. However, under the more general possible worlds semantics, a content element need not convey any information at all, as illustrated by example 3.2 (on page 41) where $NMod_M(P(a) \vee P'(a)) = \varnothing$. This is because the natural correspondence between excluded states and content elements that holds under the traditional truth-value semantics of finitely generated propositional languages, namely, that the set of content elements entailed by a sentence corresponds to the set of states excluded by the sentence, does not hold in general under the possible worlds semantics. Within the framework provided by the possible worlds semantics, the concept of *information content* is a semantic one rather than a syntactic one.

Definition 3.3 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and $\alpha \in L$. The information carried by α, or the* **content** *of α, is defined as $Cont_M(\alpha) = NMod_M(\alpha)$.*

Proposition 3.4 (Carnap and Bar-Hillel, 1952) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and $\alpha, \beta \in L$. Then the following holds:*

1. $\varnothing \subseteq Cont_M(\alpha) \subseteq S$
2. $Cont_M(\alpha) = \varnothing$ *iff* α *is* M*-valid*
3. $Cont_M(\alpha) = S$ *iff* α *is* M*-unsatisfiable*
4. $\alpha \models_M \beta$ *iff* $Cont_M(\alpha) \supseteq Cont_M(\beta)$
5. $\alpha \equiv_M \beta$ *iff* $Cont_M(\alpha) = Cont_M(\beta)$
6. $Cont_M(\alpha \wedge \beta) = Cont_M(\alpha) \cup Cont_M(\beta)$
7. $Cont_M(\alpha \vee \beta) = Cont_M(\alpha) \cap Cont_M(\beta)$
8. $Cont_M(\neg\alpha) = S - Cont_M(\alpha)$

The notion of information content is closely linked to the notion of entailment: whenever a sentence α entails a sentence β, α asserts all that is asserted by β (and possibly more), which means that the information carried by α includes the information carried by β. And since an M-valid sentence is entailed by every sentence and an M-unsatisfiable sentence entails every sentence, it is not unexpected that the former carries no information beyond that already provided by M while the latter carries all of the available information. On the other hand, sentences which are semantically equivalent entail each other and must therefore carry the same information.

In developing the theory of semantic information, Carnap and Bar-Hillel proceeded to develop numerical measures largely because they wanted to find a connection between the probabilistic foundation for inductive logic (Carnap, 1950) and the notions of semantic information. We mention the numerical aspects merely so that the overview of semantic information theory remains true to the spirit of the original work.

Carnap and Bar-Hillel (syntactically) defined the *amount of information carried by a sentence* using content elements, not directly, but indirectly via state descriptions[2]. Under the traditional truth-value semantics of finitely generated propositional languages, the amount of information carried by a sentence may be formulated using content elements by defining a measure for the smallest syntactic unit of information, i.e. the content element, and then taking the sum of the measures attributed to the elements of the information content of a sentence, i.e. the content elements entailed by the sentence, as the amount of information carried by the sentence. However, under the possible worlds semantics, the amount of information carried by a sentence will be formulated semantically by defining a measure for the smallest semantic unit of information, i.e. the state.

Definition 3.4 *The* ***content measure*** *of a state $s \in S$ is defined as $cont(s) = 1/card(S)$.*

Definition 3.5 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and $\alpha \in L$. The amount of information carried by α, or the* ***content measure*** *of α, is defined as $cont_M(\alpha) = \sum_{i=1}^{m}(cont(s_i) \mid s_i \in Cont_M(\alpha))$ where $m = card(Cont_M(\alpha))$.*

[2] In an alternative explicatum of content measure, Carnap and Bar-Hillel defined the content measure of a sentence α as $inf(\alpha) = Log(1/(1 - cont(\alpha)))$. This alternative explicatum was motivated directly by a desire to show that it was possible to use a numerical measure for semantic content that was closely analogous to the probabilistic formula for entropy used by Shannon in the mathematical theory of communication.

Proposition 3.5 (Carnap and Bar-Hillel, 1952) *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$ and $\alpha, \beta \in L$. Then the following holds:*

1. $0 \leq cont_M(\alpha) \leq 1$
2. $cont_M(\alpha) = 0$ *iff* α *is* M*-valid*
3. $cont_M(\alpha) = 1$ *iff* α *is* M*-unsatisfiable*
4. *if* $\alpha \models_M \beta$ *then* $cont_M(\alpha) \geq cont_M(\beta)$
5. *if* $\alpha \equiv_M \beta$ *then* $cont_M(\alpha) = cont_M(\beta)$
6. $cont_M(\alpha \wedge \beta) = cont_M(\alpha) + cont_M(\beta) - cont_M(\alpha \vee \beta)$
7. $cont_M(\alpha \vee \beta) = cont_M(\alpha) + cont_M(\beta) - cont_M(\alpha \wedge \beta)$
8. $cont_M(\neg\alpha) = 1 - cont_M(\alpha)$

The content measures of sentences range between 0 and 1 with 0 reserved for M-valid sentences (those having no content in the context of M) and 1 reserved for M-unsatisfiable sentences (those having contradictory content in the context of M). As expected, whenever the content of a sentence α includes the content of a sentence β, the content measure of α is never less that the content measure of β and whenever α and β have the same content, their content measures are the same.

Proposition 3.6 (Carnap and Bar-Hillel, 1952) *Suppose L is a transparent propositional language generated by n atoms. Let $M_0 = \langle S, l\rangle$ be the classical interpretation. Then the following holds:*

1. $cont(\alpha) = 1/2$ *for* α *a literal*
2. $cont(\alpha) = 1/2^n$ *for* α *a content element*

3. $cont(\alpha) = 1 - 1/2^n$ *for* α *a state description*

4. $cont(\alpha) = m/2^n$ *for* α *a sentence in SCNF with* m *conjuncts*

5. $cont(\alpha) = 1 - m/2^n$ *for* α *a sentence in SDNF with* m *disjuncts*

Under the more general possible worlds semantics, proposition 3.6 need not hold as illustrated by example 3.2 (on page 41) where $Cont_M(P(a) \vee P'(a)) = \varnothing$ for content element $P(a) \vee P'(a)$ so that $cont_M(P(a) \vee P'(a)) = 0$ thus violating properties 2 and 4. For state description $P(a) \wedge P'(a)$ where $Cont_M(P(a) \wedge P'(a)) = \{s_1, s_2, s_3\}$, properties 3 and 5 are violated since $cont_M(P(a) \wedge P'(a)) = 1$ (rather than 2/3). Lastly, property 1 is violated by literal $P(a)$, for example, since $Cont_M(P(a)) = \{s_3\}$ resulting in $cont_M(P(a)) = 1/3$ (and not 1/2).

In summary, semantic information theory is concerned with definite information, the semantic representation of which as a dichotomous 'partition' on the set of states is expressible as a sentence in either extensional SDNF or SCNF, and the syntactic representation of which as a sentence of the knowledge representation language induces a set of included states (the models of the sentence) and a complementary set of excluded states (the content of the sentence). The semantic representation of definite information provides a mechanism for the formation of semantic consequences. The formation may be viewed as a movement from a hypothesis to the set of states included by the hypothesis and from there to a semantic consequence that is true at each of the included states, or alternatively, as a movement from a hypothesis to the set of states excluded by the hypothesis and from there to a semantic consequence, the content of which is found among the set of excluded states.

3.2 KLM nonmonotonic reasoning

In defeasible reasoning, an agent may draw conclusions that are plausible but not indisputable. Traditional truth-value semantics of propositional logic and its associated semantic consequence relation do not permit such defeasible conjectures to be drawn. Nonmonotonic logics attempt to formalise defeasible reasoning. To define nonmonotonic consequence relations suitable for diagrammable systems, a form of the preferential model semantics of Kraus, Lehmann, and Magidor will be used. This influential approach, often referred to as the KLM approach, evolved, on the one hand, from earlier work by McCarthy on circumscription (McCarthy, 1980) and by Shoham on preferred models (Shoham, 1987, 1988), and, on the other hand, from the study of nonmonotonic consequence relations, pioneered by Gabbay (1985). It resulted in the development of preferential models to define preferential consequence relations (Kraus, Lehmann, and Magidor, 1990) and ranked models to define rational consequence relations (Lehmann and Magidor, 1992). In the context of diagrammable systems, nonmonotonic logic deals with information which is indefinite.

Indefinite information is information that is 'soft' or 'uncertain' and from which conclusions may be drawn that are plausible but not indisputable. Indefinite information is provided by *default rules*. Default rules arise naturally in the presence of incomplete information and are a common occurrence in daily life (Davis, 1990). In natural language, a default rule may be expressed in various ways, for example, by a sentence including the word 'normally', by a sentence including the word 'typically', by a sentence including the word 'probably', or by a sentence indicating a subjective preference. Unlike definite information, indefinite information can (in general) not be expressed by sentences of a formal knowledge representation language. For example, attempts to introduce generalised quantifiers like 'few', 'many', and 'most' have been shown to be subject to paradoxes. By way of illustration, suppose we introduce a quantifier Q that stands for 'normally' and tackle the semantics in the manner of Mostowski (1957). Then we might take $QxP(x)$ to be true in a term interpretation (D, den) if $P(d)$ were satisfied for more than half of the constants d in D. Now we can construct the paradox. Suppose P is a predicate symbol expressing the property of being a cat, and that P' is a predicate symbol expressing the property of having only three legs. The standard way to formalise the idea that 'Normally, cats have only three legs' would be $Qx(P(x) \rightarrow P'(x))$, in other words 'Normally, if x is a cat then x has only three legs'. Consider an interpretation in which there are, say, ten member of the domain D, and let den be such that four objects are four-legged cats and the remaining six objects are three-legged stools. Thus more than half of the objects satisfy the unquantified formula $P(x) \rightarrow P'(x)$, and so the original quantified formula is true in term interpretation (D, den). But this makes no sense, since D contains no 3-legged cats at all!

As an example of indefinite information, the extract from the Three Mile Island account, provided earlier, is considered again:

> "On the control panel, one dial subsequently showed that the pressure in the reactor was still falling while another dial showed that the pressure in the pressuriser was rising. Normally, these dials move together and their readings were seen as contradictory".

In this example, the indefinite information, in the form of a default rule, states that firstly, if the pressure in the reactor falls, then the pressure in the pressuriser typically falls as well (and vice versa), and secondly, that if the pressure in the reactor rises, then, normally, the pressure in the pressuriser rises as well (and vice versa). In a transparent propositional language, a default rule cannot be expressed by sentences of the language in any obvious way. In formalising the indefinite information available to an agent, it will thus be necessary to find an alternative representation of indefinite information; either a syntactic representation via inference rules, as in Reiter's default logic (Reiter, 1980), or, as we shall instead prefer, a semantic representation.

Whereas definite information is represented semantically by a division of the finite set of states into at most two complementary subsets, thus permitting subsets of states to be matched up with sentences via the notion of a model, indefinite information is given by a division of the finite set of states into more than two subsets. This division may very naturally be represented by a strict linearly ordered partition where the subset of states lowest down in the ordering is regarded as most preferred (i.e. most normal, likely, or typical) and the subsets of states higher up as progressively less preferred (i.e. less normal, likely, or typical). A *strict linear order* R on a set X is a binary relation on X that is irreflexive, antisymmetric, transitive, and connected, where by *connected* it is understood that for all elements $x, y \in X$, either $(x, y) \in R$ or $(y, x) \in R$ or $x = y$.

Visually, this semantic representation may be pictured as a stack of boxes one on top of another, with a bottom box that has nothing under it and a top box that has nothing above it, and the set of states contained in each box corresponding to an element of the partition and the position of each box in the picture corresponding to that of its associated element in the ordering.

Example 3.3 *Consider again the Light-Fan system introduced in example 3.1 (on page 39). Suppose the agent has through experience observed that the light is normally on, and if the light is on then normally the fan is on as well. This example of indefinite information may be represented visually as shown in figure 3-2.*

A strict linearly ordered partition of states may be produced from a total preorder on the set of states as shown by the following (more general) proposition. A *preorder* R on a set X is a reflexive and transitive relation on X and is *total* iff for all elements $x, y \in X$, either $(x, y) \in R$ or $(y, x) \in R$ or both.

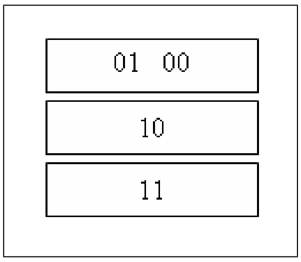

Figure 3-2: Indefinite information

Proposition 3.7 *Let X be a set and let R be a total preorder on X. Then R produces a strict linearly ordered partition of X.*

Proof. See proof in appendix A, section A.1. ■

For every preorder R on a set X, there is a corresponding strict order on X which is given by $R_S = R - \{(x, y) \mid (y, x) \in R\}$. An order relation is strict if it is irreflexive instead of reflexive. Preorders R and Q on X are order-equivalent iff $R_S = Q_S$. For a total preorder, the corresponding strict order is a (strict) modular partial order, as will be shown subsequently. A *partial order* R on a set X is a preorder that is antisymmetric.

Definition 3.6 (Lehmann and Magidor, 1992) *A partial order R on X is* ***modular*** *iff for every x, y, and $z \in X$ one of the following holds:*

- *if $(x, y) \notin R$, $(y, x) \notin R$, and $(z, x) \in R$ then $(z, y) \in R$*
- *if $(x, y) \in R$ then either $(z, y) \in R$ or $(x, z) \in R$*
- *there is a totally ordered set Ω and function $r : X \to \Omega$ such that $(x, y) \in R$ iff $r(x) < r(y)$*

Proposition 3.8 *Let X be a set. For every total preorder R on X there is a strict modular partial order Q on X such that R and Q are order-equivalent, and conversely,* *for every strict modular partial order Q on X there is a total preorder R on X such that Q and R are order-equivalent.*

Proof. Let R be a total preorder on X and let $Q = R_S$. But then Q is order-equivalent to R. By construction, Q is irreflexive on X and antisymmetric. To see that Q is transitive, suppose that $(x, y) \in Q$ and $(y, z) \in Q$. So $(x, y) \in R$ and $(y, z) \in R$ and (by transitivity of R) $(x, z) \in R$. But $(z, x) \notin R$ since otherwise $(y, x) \in R$ (by transitivity of R) and then $(x, y) \notin Q$. So $(x, z) \in Q$. Thus Q is a strict partial order on X. To see that Q is modular suppose that $(x, y) \notin Q$, $(y, x) \notin Q$, and $(z, x) \in Q$. Since R is total it must be the case that either $(x, y) \in R$ or $(y, x) \in R$. Since neither belongs to Q, it follows that both $(x, y) \in R$ and $(y, x) \in R$. But since $(z, x) \in Q$ it must hold that $(z, x) \in R$ and $(x, z) \notin R$. So $(z, y) \in R$ by transitivity of R but $(y, z) \notin R$. Thus $(z, y) \in Q$, i.e. Q is modular.

Conversely, let Q be a strict modular partial order on X and let $R = Q \cup \{(x, y) \mid (x, y) \notin Q \text{ and } (y, x) \notin Q\}$. Clearly $R_S = Q$ and hence R order-equivalent to Q. For every $x \in X, (x, x) \notin Q$. But then $(x, x) \in R$ for every $x \in X$, i.e. R is reflexive. Transitivity of R follows directly from transitivity of Q. Thus R is a preorder. To see that R is total, pick any $x, y \in X$. If $(x, y) \in Q$ then $(x, y) \in R$. Similarly, if $(y, x) \in Q$, then $(y, x) \in R$. If $(x, y) \notin Q$ and $(y, x) \notin Q$ then, by construction, $(x, y) \in R$ and $(y, x) \in R$. Thus R is total. ■

Proposition 3.8 suggests that a strict linearly ordered partition of states may also be produced from a strict modular partial order on the set of states. Furthermore, since a *total preorder* and its corresponding *strict modular partial order* are order-equivalent, one would expect them to produce the same linearly ordered partition of states.

Proposition 3.9 *Let X be a set and let Q be a strict modular partial order on X. Then Q produces a strict linearly ordered partition of X.*

Proof. See proof in appendix A, section A.1. ■

Proposition 3.10 *Let X be a set, R a total preorder on X, and Q the corresponding strict modular partial order on X. Then R and Q produce the same strict linearly ordered partition of X.*

Proof. See proof in appendix A, section A.1. ■

Semantically, indefinite information may be represented by either a total preorder or a strict modular partial order since the effect of both is to produce a strict linearly ordered partition of states as shown by proposition 3.10. For illustrative purposes, an example is provided below of a total preorder, its corresponding strict modular partial order, and the strict linearly ordered partition that they produce.

Example 3.4 *Suppose $S = \{11, 10, 01, 00\}$ is a set of states . The relation $R = \{(11, 11),$ $(10, 10), (01, 01), (00, 00), (11, 10), (10, 00), (11, 00), (11, 01), (01, 00), (10, 01), (01, 10)\}$ is a total preorder on S, and the relation $R_S = \{(11, 10), (10, 00), (11, 00), (11, 01), (01, 00)\}$ the corresponding strict modular partial order on S. Then $P = \langle\{11\}, \{10, 01\}, \{00\}\rangle$ is the strict linearly ordered partition produced by R (and also by R_S). See the illustration in figure 3-3.*

The KLM approach was formulated in terms of strict modular partial orders. Nonetheless, the visual representation of an ordered partition by boxes positioned on top of one another highlights the difference between indefinite information and definite information in a way not as apparent with either preorders or partial orders.

Visually, a dichotomous ordered partition on the set of states depicted by a box of least normal states positioned on top of a box of most normal states has the same appearance as the dichotomy associated with definite information and depicted by a box of excluded states positioned on top of a box of included states, but has a different

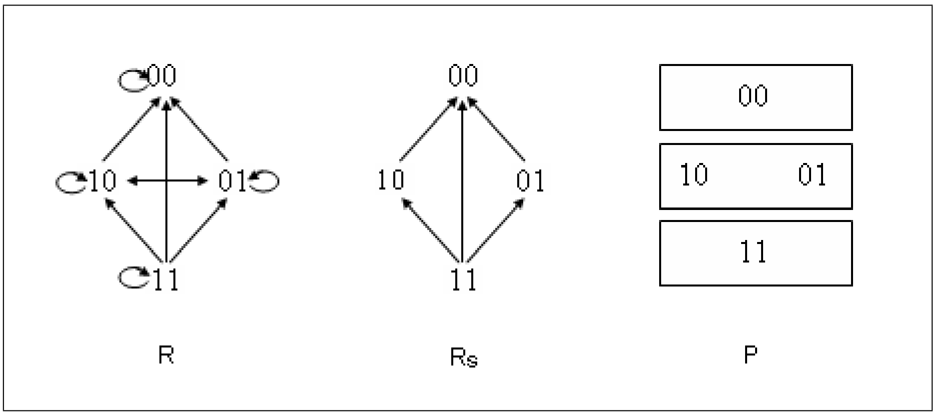

Figure 3-3: Orderings

meaning. In the case of indefinite information and default rules, we may view the less normal states as being tentatively, rather than definitely, excluded. Should we wish to treat both indefinite information and definite information, a visualisation will have to be developed that makes this distinction clear.

Strict modular partial orders (or total preorders) play an important role in the preferential model semantics of Kraus, Lehmann, and Magidor. Such a relation provides an ordering on the set of states from which the notion of a minimal element in a subset of states, and hence, the notion of a model of a sentence that is most preferred with respect to the ordering, may be defined.

Definition 3.7 *Let R be a total preorder on a set X.*

- *For every $Y \subseteq X$, an element $y \in Y$ is* ***minimal*** *in Y with respect to R iff there is no $x \in Y$ such that $(x, y) \in R$, unless $(y, x) \in R$. The set of minimal elements in Y with respect to R is denoted by $Min_R(Y)$.*

- *For every $Y \subseteq X$, the preorder R is Y-**smooth** iff for every element $y \in Y$ there* *is some $x \in Y$ such that $(x, y) \in R$ and x is minimal in Y*[3]*.*

In the context of diagrammable systems, the definition of a ranked model (or interpretation) may be simplified and adapted in the obvious way to a (finitely generated) propositional language that is transparent rather than opaque.

Definition 3.8 *A* ***ranked interpretation*** *of L is a triple $P = \langle S, R, l\rangle$ such that*

- *S is a non-empty set of states,*
- *$l : S \to U_T$ is a labelling function, and*
- *R is a total preorder on S*[4]*.*

The notion of satisfaction and of a model are defined in exactly the same way for ranked interpretations as for possible worlds interpretations. The notion of a minimal model is, however, only applicable to ranked interpretations. Given a ranked interpretation $P = \langle S, R, l\rangle$, a state $s \in S$ is called a minimal model of a sentence α iff $s \in Min_R(Mod_P(\alpha))$. Given P, the set of minimal models of α is denoted by $Min_R(\alpha)$. The essential difference between ranked interpretations and possible worlds interpretations is that, through the use of minimal models, every ranked interpretation P induces a defeasible consequence relation on L. A defeasible consequence relation is a consequence relation $\mid\!\sim$ on L that cannot be counted on to always satisfy monotonicity. *Monotonicity* in the case of a consequence relation, as opposed to a consequence operation, translates into saying that if $\alpha \mid\!\sim \beta$ then $\alpha \wedge \gamma \mid\!\sim \beta$ for all $\alpha, \beta, \gamma \in L$.

Definition 3.9 *Let $P = \langle S, R, l\rangle$ be a ranked interpretation of L. The* ***defeasible*** *consequence relation induced by P is the relation $\mid\!\sim_P$ on L given by $\alpha \mid\!\sim_P \beta$ iff $Min_R(\alpha) \subseteq$*

[3] *The smoothness condition is essentially the same as the limit assumption in Lewis (1973).*

[4] *Lehmann and Magidor also require R to be X-smooth for every sentence $\alpha \in L$ where $X = Mod_M(\alpha)$ and $M = \langle S, l\rangle$. However, this condition holds true since L is finite.*

*$Mod_P(\beta)$. If $\alpha \mid\!\sim_P \beta$ then α is said to **defeasibly entail** β (under P) and β is said to be a **defeasible consequence** of α (under P).*

The defeasible consequence relation $\mid\!\sim_P$ induced by a ranked interpretation $P = \langle S, R, l\rangle$ does not necessarily satisfy Monotonicity, as expected. However, in the limit, for the ranked interpretation $P = \langle S, R, l\rangle$ with $R = S \times S$, the defeasible consequence relation $\mid\!\sim_P$ induced by P is exactly the semantic consequence relation $\models_P$ and thus monotonic. The more general relationship between semantic consequence, semantic equivalence, and defeasible consequence is provided by the following proposition, the proof of which is routine. Note that *Supraclassicality* is defined in the version appropriate for consequence relations rather than for consequence operations.

Proposition 3.11 *Let $P = \langle S, R, l\rangle$ be a ranked interpretation of L and let $\mid\!\sim_P$ be the defeasible consequence relation on L induced by P. Then $\mid\!\sim_P$ satisfies the following properties:*

- *if $\alpha \models_P \beta$ then $\alpha \mid\!\sim_P \beta$ (Supraclassicality)*
- *if $\alpha \equiv_P \alpha'$, $\beta \equiv_P \beta^{'}$, and $\alpha \mid\!\sim_P \beta$ then $\alpha' \mid\!\sim_P \beta^{'}$ (Well-behavedness)*

The KLM approach focusses on the properties of defeasible consequence relations, based on the general assumption that all reasonable logical systems draw only sensible conclusions. Of particular interest are the family of preferential consequence relations and the family of rational consequence relations.

Definition 3.10 *A **preferential** consequence relation is a binary relation $\mid\!\sim$ on L that satisfies the following properties for all α, β, and $\gamma \in L$:*

- *for every $\alpha \in L$, $\alpha \mid\!\sim \alpha$ (Reflexivity)*

- *if* $\alpha \mid\!\sim \beta$ *and* $\alpha \equiv \gamma$ *then* $\gamma \mid\!\sim \beta$ *(Left Logical Equivalence)*
- *if* $\alpha \mid\!\sim \beta$ *and* $\beta \models \gamma$ *then* $\alpha \mid\!\sim \gamma$ *(Right Weakening)*
- *if* $\alpha \mid\!\sim \beta$ *and* $\alpha \mid\!\sim \gamma$ *then* $\alpha \mid\!\sim \beta \wedge \gamma$ *(And)*
- *if* $\alpha \mid\!\sim \gamma$ *and* $\beta \mid\!\sim \gamma$ *then* $\alpha \vee \beta \mid\!\sim \gamma$ *(Or)*
- *if* $\alpha \mid\!\sim \beta$ *and* $\alpha \mid\!\sim \gamma$ *then* $\alpha \wedge \gamma \mid\!\sim \beta$ *(Cautious Monotonicity)*

Reflexivity (in the case of a consequence relation) states that a hypothesis is a defeasible consequence of itself. *Left Logical Equivalence* stipulates that semantically equivalent hypotheses have the same defeasible consequences while *Right Weakening* stipulates that the defeasible consequences of a hypothesis include logically weaker sentences. The conjunction of two defeasible consequences of the same hypothesis is itself a defeasible consequence (by *And*) while a (common) defeasible consequence of two different hypotheses is also a defeasible consequence of the disjunction of the hypotheses (by *Or*). *Cautious Monotonicity* states that a hypothesis may be expanded without invalidating previous defeasible consequences by adding to it one of its defeasible consequences.

Definition 3.11 *A* ***rational*** *consequence relation is a preferential consequence relation* $\mid\!\sim$ *on* L *that satisfies the following property for all* α, β, *and* $\gamma \in L$*:*

- *if* $\alpha \mid\!\sim \beta$ *and* $\alpha \not\mid\!\sim \neg\gamma$ *then* $\alpha \wedge \gamma \mid\!\sim \beta$ *(Rational Monotonicity)*

Rational Monotonicity is stronger than *Cautious Monotonicity* in the sense that a hypothesis may be expanded without invalidating previous defeasible consequences by adding any sentence to the hypothesis, provided the negation of the added sentence is not already a defeasible consequence of the hypothesis. Rational consequence relations may be characterised in terms of ranked interpretations as shown by Lehmann and Magidor (1992) in the following theorem.

Theorem 3.1 (Lehmann and Magidor, 1992) *If $P = \langle S, R, l \rangle$ is a ranked interpretation of L, then the defeasible consequence relation $\mid\!\sim_P$ on L induced by P is rational. Conversely, if a binary relation $\mid\!\sim$ on L is a rational consequence relation, then it can be defined by some ranked interpretation of L.*

As mentioned earlier, the rational consequence relation induced by a ranked interpretation differs from the semantic consequence relation induced by a possible worlds interpretation primarily in not always satisfying monotonicity. However, a number of other important properties which are satisfied by semantic consequence relations, are not (in general) satisfied by rational consequence relations.

Proposition 3.12 *Let $P = \langle S, R, l \rangle$ be a ranked interpretation of L and let $\mid\!\sim_P$ be the rational consequence relation on L induced by P. Then $\mid\!\sim_P$ does not always satisfy the following properties (which are satisfied by $\models_P$):*

- *if $\alpha \models_P \beta$ and $\beta \models_P \alpha$ then $\alpha \equiv_P \beta$ (Antisymmetry)*
- *if $\alpha \models_P \beta$ and $\beta \models_P \gamma$ then $\alpha \models_P \gamma$ (Transitivity)*
- *if $\alpha \models_P \beta$ then $\neg\beta \models_P \neg\alpha$ (Contraposition)*

In summary, nonmonotonic reasoning is concerned with indefinite information, the semantic representation of which by some form of ordering is utilised in providing a mechanism for the formation of defeasible conjectures. The mechanism is essentially a movement from a hypothesis to those models that are minimal with respect to the ordering representing the indefinite information and from there to a defeasible consequence that is true at each of the minimal models. In cases where every model is minimal, the mechanism yields classic semantic consequences.

3.3 Templated orderings

Templated orderings have their origins in the strict modular partial orderings associated with nonmonotonic logic. Recall from the previous section that the effect of both a strict modular partial order and a total preorder is to produce a strict linearly ordered partition on a set of states. The visual representation of such an ordered partition by boxes stacked on top of one another, one box for each element of the partition, provides a useful analogue for introducing the concept of a templated ordering. Loosely speaking, a templated ordering may be seen as the distribution of a set of states over an arrangement of boxes stacked on top of one another with one box specially designated as the top box. This fixed construct of stacked boxes is what is referred to as the *template*. A template is determined by the size of the set of states in the sense that it contains, apart from the top box, a box for each element of the largest possible partition on the set of states. Thus if there are n states, the template consists of $n+1$ boxes. The top box in a template fulfils the same role as the top box encountered in semantic information theory: it is reserved for those states that are definitely excluded.

In the context of diagrammable systems, where the states of the system correspond to a subset of the valuations of some finitely generated propositional language, a template is of course influenced by the language itself. In the case where there is a bijective correspondence between states and valuations, a propositional language generated by n atoms will result in a template consisting of exactly $2^n + 1$ boxes, since the language has 2^n valuations.

A templated ordering may be regarded as an instance of an *abstract data type*, or *ADT* for short. An ADT is a collection of objects together with a collection of operations that may be performed on the objects and a collection of laws describing the behaviour of the objects. (See, for example, Aho and Ullman (1992). Alternatively, for a more formal treatment, see Guttag (1986) or Loeckx, Ehrich, and Wolf (1996).) In a templated ordering ADT, the objects are the templated orderings themselves, i.e. the distributions of a set of states across, or over, a template.

Definition 3.12 *Let S be a finite non-empty set of states with $card(S) = n$. The* ***template*** *for S is the set $B = \{0, 1, \ldots, n\}$ (of box-labels) together with the usual linear order $\leq$ on B. A* ***templated ordering*** *or,* ***t-ordering*** *for short, is a function $t : S \to B$. The class of all t-orderings having the same set S of states and template B for S is denoted by T_S.*

Choosing the set $\{0, 1, \ldots, n\}$ as the template B for S is a matter of convenience. The important characteristic of a template is that it permits the finest possible partition of a finite set of states and that a linear order is defined on the template. The basic idea underlying templates is qualitative in nature although, admittedly, choosing the set $\{0, 1, \ldots, n\}$ opens up the possibility of using t-orderings in a more quantitative manner. This temptation will be resisted.

Definition 3.13 *The collection of basic operations for a class T_S of t-orderings is defined as follows:*

- *get* $: T_S \times B \times B \to \wp S$ *by* $get(t, i, j) = \{s \in S \mid i \leq t(s) \leq j \text{ and } i \leq j\}$
- *top* $: T_S \to \wp S$ *by* $top(t) = \{s \in S \mid t(s) = n\}$
- *bottom* $: T_S \to \wp S$ *by* $bottom(t) = \{s \in S \mid t(s) = 0\}$
- *push* $: T_S \times S \times B \to T_S$ *by*
$$push(t, s, j) = \begin{cases} t \oplus \{(s, j)\} & \text{if } j > t(s), \\ t & \text{otherwise} \end{cases}$$
- *pull* $: T_S \times S \times B \to T_S$ *by*
$$pull(t, s, j) = \begin{cases} t \oplus \{(s, j)\} & \text{if } j < t(s), \\ t & \text{otherwise.} \end{cases}$$

where the overriding operation $\oplus$ on functions is given by $f \oplus g = \{(x, y) \mid$ *if* x *is in the domain of* g *then* $y = g(x)$*, else* $y = f(x)\}$

Operation '*get*' retrieves the sets of states between (and at) two specified box-labels while operation '*top*' simply retrieves the set of states at the top box and operation '*bottom*' the set of states at the bottom box. Operation '*push*' modifies the position of a state by pushing it up to a specified box-label while operation '*pull*' modifies the position of a state by pulling it down to a specified box-label.

Definition 3.14 *Specialisations of operation* $get : T_S \times B \times B \rightarrow \wp S$ *for a class* T_S *of t-orderings are defined as follows:*

- $get_{\rightarrow} : T_S \times B \rightarrow \wp S$ *by* $get_{\rightarrow}(t, j) = get(t, j, j)$
- $get_{\uparrow} : T_S \times B \rightarrow \wp S$ *by* $get_{\uparrow}(t, j) = get(t, 0, j)$
- $get_{\downarrow} : T_S \times B \rightarrow \wp S$ *by* $get_{\downarrow}(t, j) = get(t, j, n)$

Operation '$get_{\rightarrow}$' retrieves the set of states at a specified box-label. Operation '$get_{\uparrow}$' retrieves the sets of states starting from the bottom box up to (and including) a specified box-label while operation '$get_{\downarrow}$' retrieves the sets of states starting from the top box down to (and including) a specified box-label.

Definition 3.15 *Two auxiliary operations for a class* T_S *of t-orderings will prove useful and are defined as follows:*

- $first_{\uparrow} : T_S \rightarrow B$ *by* $first_{\uparrow}(t) = j$ *where* $get_{\uparrow}(t, j-1) = \varnothing$ *and* $get_{\rightarrow}(t, j) \neq \varnothing$
- $first_{\downarrow} : T_S \rightarrow B$ *by* $first_{\downarrow}(t) = j$ *where* $get_{\downarrow}(t, j+1) = \varnothing$ *and* $get_{\rightarrow}(t, j) \neq \varnothing$

Operation '$first_\uparrow$' retrieves the first box-label that is occupied by a set of states, starting from the bottom box going up while operation '$first_\downarrow$' retrieves the first box-label that is occupied by a set of states, starting from the top box going down.

Proposition 3.13 *Let $t \in T_S$, $i, j \in B$, and $s \in S$. Then the following constraints hold:*

1. $bottom(t) \subseteq get_\uparrow(t, j)$ *and* $top(t) \subseteq get_\downarrow(t, j)$
2. $get_\rightarrow(t, i) \subseteq get_\uparrow(t, j)$ *if* $i \leq j$ *and* $get_\rightarrow(t, i) \subseteq get_\downarrow(t, j)$ *if* $i \geq j$
3. $get_\uparrow(t, j) \cup get_\downarrow(t, j) = S$ *and* $get_\uparrow(t, j) \cap get_\downarrow(t, j) = get_\rightarrow(t, j)$
4. $pull(push(t, s, j), s, t(s)) = t$
5. $push(pull(t, s, j), s, t(s)) = t$

Proof. See proof in appendix A, section A.2. ■

Definition 3.16 *The **templated ordering ADT** is the class which is the union of all the classes T_S of t-orderings.*

Definition 3.17 *A t-ordering $t \in T_S$ is **strongly contradictory** iff $get_\uparrow(t, n-1) = \varnothing$ and **weakly contradictory** iff $get_\rightarrow(t, 0) = \varnothing$ but $get_\uparrow(t, n-1) \neq \varnothing$. A t-ordering $t \in T_S$ is **tautological** iff $get_\rightarrow(t, 0) = S$.*

Every t-ordering induces a strict modular partial order or, alternatively, a total preorder on the set of states.

Proposition 3.14 *Let $t \in T_S$. Then t induces a total preorder and a strict modular partial order on S.*

Proof. Let R_t be the relation on S induced by t-ordering $t \in T_S$ such that $(s, s') \in R_t$ iff $t(s) \leq t(s')$. Clearly, the relation R_t is reflexive, transitive, and total. Alternatively, let Q_t be the relation on S induced by t-ordering $t \in T_S$ such that $(s, s') \in Q_t$ iff $t(s) < t(s')$. It is easy to see that Q_t is irreflexive, antisymmetric, and transitive. To see that Q_t is modular note that B is a linearly ordered set and t a function from S to B such that $(s, s') \in Q_t$ iff $t(s) < t(s')$. ■

The total preorder induced by a t-ordering and the strict modular partial order induced by the same t-ordering are order-equivalent since the corresponding strict order of the total preorder is precisely the strict modular partial order. If a t-ordering is strongly contradictory then the total preorder induced by the t-ordering is the relation $R = S \times S$ while the strict modular partial order induced by the same t-ordering is the relation $Q = \varnothing$. Note that the same relations are induced by a t-ordering that is tautological (and by every t-ordering where $t(s) = i$ for every $s \in S$). Thus a t-ordering is more than a total preorder or strict modular partial order.

Definition 3.18 *Let $t, t' \in T_S$. Then t and t' are **order-equivalent** iff the total preorders induced by t and t' are order-equivalent.*

Each equivalence class of order-equivalent t-orderings has a member which is minimal in the sense that every state, except for the definitely excluded states in the top box, is as low as possible.

Definition 3.19 *A t-ordering $t \in T_S$ is in **normal form** iff for each (box-label) $i \in B$, if $i < n$ and $get_{\rightarrow}(t, i) = \varnothing$ then $get(t, i, n-1) = \varnothing$. The subclass of all t-orderings $t \in T_S$ in normal form is denoted by T_N.*

Included in the subclass of t-orderings in normal form are t-orderings which are tautological and t-orderings which are strongly contradictory, but t-orderings which are weakly contradictory are excluded. The subclass of t-orderings in normal form effectively defines a purely qualitative subclass of t-orderings because the notion of relative distance, which can be expressed by unrestricted t-orderings, is removed. T-orderings in normal form are an important subclass of t-orderings and, as will be shown, every t-ordering can be transformed into a t-ordering in normal form that is order equivalent to the original t-ordering. Such a transformation, which is called normalisation, is reliant on the fact that template B is a well-ordered set. An ordering is a *well-ordering* on a set X if every nonempty subset of X has a least element under the ordering. If X is finite and R is a linear order on X, then R is automatically a well-ordering on X. The *initial segment* of an element $x \in X$ is given by $seg(x) = \{y \in X \mid (y, x) \in R\}$. In a well-ordered set X, the elements in $seg(x)$ are followed by element x in the ordering in the sense that x is the least member of the complement of $seg(x)$. Given a well-ordered set X, x is referred to as the *successor* of $seg(x)$.

Definition 3.20 *Let A be a non-empty proper subset of template B. The normalise function for A is the function $g : A \to B$ given by*

- $g(x) = \begin{cases} card(seg(x)) & \text{if } x \neq n \\ n & \text{otherwise} \end{cases}$
  where $seg(x)$ is the initial segment of x.

Definition 3.21 *The* ***normalisation*** *of t-ordering $t \in T_S$ is the t-ordering of T_S given by the composition $g \circ t$ where g is the normalise function for $ran(t)$.*

To prove that the normalisation of a t-ordering produces a t-ordering in normal form that is order-equivalent to the original, several lemmas are subsequently introduced.

Lemma 3.1 *Let A be a non-empty proper subset of template B and let g be the normalise function for A. Then $g(x) < n$ for every $x \in A$ such that $x \neq n$.*

Proof. See proof in appendix A, section A.2. ■

Lemma 3.2 *Let A be a non-empty proper subset of template B and let g be the normalise function for A. Then g is injective.*

Proof. See proof in appendix A, section A.2. ■

Lemma 3.3 *Let A be a non-empty proper subset of template B and let g be the normalise function for A. If $j < n$ and $j \in ran(g)$, then $k \in ran(g)$ for every $0 \leq k < j$.*

Proof. See proof in appendix A, section A.2. ■

Proposition 3.15 *Let $t \in T_S$. Then $g \circ t$ is in normal form where g is the normalise function for $ran(t)$.*

Proof. Assume there is some $i \in B$ such that $i < n$ and $get_{\rightarrow}(g \circ t, i) = \varnothing$. It must be shown that $get(g \circ t, i, n-1) = \varnothing$. Suppose it is not the case. So there must be at least one $s \in S$ such that $(g \circ t)(s) = g(t(s)) = j$ for $i < j \leq n-1$. So $j < n$ and $j \in ran(g)$. But then $k \in ran(g)$ for every $0 \leq k < j$ by lemma 3.3. Since $i < j$ it follows that $i \in ran(g)$. But then there must be some $s' \in dom(t)$ such that $g(t(s')) = i$. But then $s' \in get_{\rightarrow}(g \circ t, i)$, i.e. $get_{\rightarrow}(g \circ t, i) \neq \varnothing$. Contradiction. So $get(g \circ t, i, n-1) = \varnothing$. Thus $g \circ t$ is in normal form. ■

Proposition 3.16 *Let $t \in T_S$. Then $g \circ t$ is order-equivalent to t where g is the normalise function for $ran(t)$.*

Proof. Let Q_t be the strict modular partial order induced by t-ordering t and let $Q_{g \circ t}$ be the strict modular partial order induced by t-ordering $g \circ t$. Pick any element $(s_i, s_j) \in Q_t$. So $t(s_i) < t(s_j)$. Suppose $t(s_j) \neq n$. Then $g(t(s_i)) = card(seg(t(s_i)))$ and $g(t(s_j)) = card(seg(t(s_j)))$. But $t(s_i) \in seg(t(s_j))$ and so $g(t(s_i)) < g(t(s_j))$. Thus

$(s_i, s_j) \in Q_{g\circ t}$. Suppose $t(s_j) = n$. Then $g(t(s_i)) = card(seg(t(s_i)))$ and $g(t(s_j)) = n$. But $t(s_i) \neq n$ and so (by lemma 3.1) $g(t(s_i)) < n$. Hence, $g(t(s_i)) < g(t(s_j))$ and so $(s_i, s_j) \in Q_{g\circ t}$. Since (s_i, s_j) was chosen arbitrarily, $Q_t \subseteq Q_{g\circ t}$.
Conversely, suppose $(s_i, s_j) \in Q_{g\circ t}$. So $g(t(s_i)) < g(t(s_j))$. Suppose $t(s_j) \neq n$. Then $card(seg(t(s_i))) < card(seg(t(s_j)))$, i.e. $seg(t(s_i)) \subset seg(t(s_j))$. But $ran(t)$ is well-ordered and so $t(s_i) = succ(seg(t(s_i)))$ and $t(s_j) = succ(seg(t(s_j)))$. But then $t(s_i) < t(s_j)$. Thus $(s_i, s_j) \in Q_t$. Suppose $t(s_j) = n$. Then $t(s_i) < n$ otherwise $g(t(s_i)) = g(t(s_j))$. So $t(s_i) < t(s_j)$ and thus $(s_i, s_j) \in Q_t$. Since (s_i, s_j) was chosen arbitrarily, $Q_{g\circ t} \subseteq Q_t$. So $Q_t = Q_{g\circ t}$, i.e. t-orderings t and $g \circ t$ are order-equivalent. ■

Proposition 3.16 shows that for every t-ordering in a class of t-orderings having the same set of states and template, there exists a t-ordering in normal form that is order-equivalent to the original. T-orderings in normal form provide the basis for representing definite and indefinite information semantically, and for expressing these semantic representations syntactically.

Definite t-orderings provide a means for representing definite information in an easily visualisable way. Recall from semantic information theory that definite information was represented semantically by a division of the finite set of states into a set of included states and a complementary set of excluded states. In a definite t-ordering, the set of included states is distributed below the top box (over the bottom box) and the complementary set of excluded states over the top box.

Example 3.5 *Consider the Light-Fan system of example 3.1 (on page 39). Using a definite t-ordering, the observation that the light is on may be represented semantically as shown in figure 3-4.*

Definition 3.22 *A t-ordering $t \in T_S$ is* ***definite*** *iff t is in normal form and $top(t) \cup bottom(t) = S$. The subclass of all definite t-orderings $t \in T_S$ is denoted by T_D.*

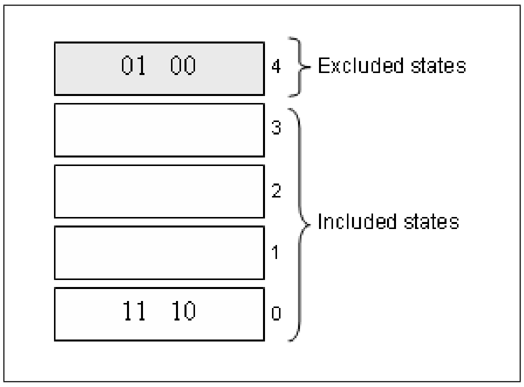

Figure 3-4: Definite t-ordering

When a definite t-ordering is tautological, it corresponds to the case in semantic information theory where the set of excluded states is empty, thus representing the instance of no information. In contrast, when a definite t-ordering is strongly contradictory, every state belongs to the set of excluded states thus representing the instance of too much (or contradictory) information.

Indefinite t-orderings provide a means for representing indefinite information in a visualisable way that is compatible with our visualisation of definite information. Recall from nonmonotonic logic that a default rule was represented semantically by a division of the finite set of states into a strict linearly ordered partition where the subset of states lowest down in the ordering was considered most typical (or normal) and the subsets of states higher up as progressively less typical (or normal). In an indefinite t-ordering, the subset of states considered most typical are distributed over the bottom box, the subset of states considered less typical over the next box, and so on, until the subset of states considered least typical has been distributed over a box. The set of *maximally preferred states* is taken to be the subset of most typical states, i.e. the set of states distributed over the bottom box. Note that the top box will remain empty. This is consistent with the idea that indefinite information warrants only the tentative ruling out of states, not definite exclusion.

Example 3.6 *Consider the Light-Fan system of example 3.1 (on page 39). Using an indefinite t-ordering, the default rule 'Normally the light is on, and if the light is on, then normally the fan is on as well' may be represented semantically as shown in figure 3-5.*

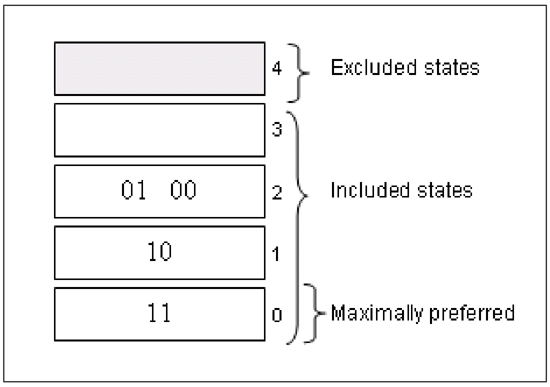

Figure 3-5: Indefinite t-ordering

Definition 3.23 *A t-ordering $t \in T_S$ is* ***indefinite*** *iff t is in normal form and $top(t) = \varnothing$ and $bottom(t) \neq S$. The subclass of all indefinite t-orderings $t \in T_S$ is denoted by T_I.*

In defining an indefinite t-ordering as a t-ordering in normal form, indefinite information has been restricted to the kind of default rules encountered in nonmonotonic reasoning approaches such as the KLM approach. In section 3.7, an example of a default rule that cannot be represented by a t-ordering in normal form will be encountered, and it will be shown that such a representation leads to a kind of 'weak' paradox.

Definite and indefinite t-orderings are two specialisations of t-orderings in normal form in which either definite or indefinite information is represented, but not both. In contrast, regular t-orderings, which is another specialisation of t-orderings in normal form, provide a means for representing the combination of (non-contradictory) definite information and indefinite information.

Example 3.7 *Consider the Light-Fan system of example 3.1 (on page 39). Suppose the agent has available definite information in the form of an observation that the light is on (with ignorance about the fan) and indefinite information in the form of a default rule that says 'Normally the light is on, and if the light is on, then normally the fan is on as well'. Using a regular t-ordering, the combination of definite and indefinite information may be represented semantically as shown in figure 3-6.*

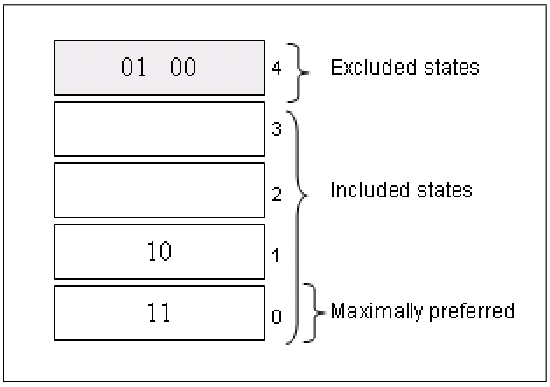

Figure 3-6: Regular t-ordering

Definition 3.24 *A t-ordering $t \in T_S$ is* ***regular*** *iff t is in normal form and t is not strongly contradictory. A regular t-ordering is said to be* ***pure*** *iff $t \notin T_D$ and $t \notin T_I$. The subclass of all regular t-orderings $t \in T_S$ is denoted by T_E.*

The subclass of regular t-orderings differs from the subclass of t-orderings in normal form only in the exclusion of t-orderings that are strongly contradictory. An important property of regular t-orderings is that if two regular t-orderings are order-equivalent, then they are identical. The same does not hold true (in general) for t-orderings in normal form. To see why, consider t-orderings $t = \{(11, 0), (10, 0), (01, 0), (00, 0)\}$ and t' $= \{(11, 4), (10, 4), (01, 4), (00, 4)\}$. The total preorders induced by both t-orderings are the relation $R = S \times S$. So we have a case in which two t-orderings in normal form are order-equivalent but, clearly, not identical.

Proposition 3.17 *Let $t, t' \in T_E$. If t and t' are order-equivalent then $t = t'$.*

Proof. Suppose that t and t' are order-equivalent. So $t(s) \leq t(s')$ iff $t'(s) \leq t'(s')$ for every $s, s' \in S$. It must be shown that $t(s) = t'(s)$ for every $s \in S$, i.e. it must be shown that $get_{\rightarrow}(t, i) = get_{\rightarrow}(t', i)$ for every $i \in B$. The proof is by induction.

Suppose $i = 0$. Choose any $s \in get_{\rightarrow}(t, 0)$. So $t(s) = 0$. But then $t(s) \leq t(s')$ for every $s' \in S$. But t and t' are order-equivalent and thus $t'(s) \leq t'(s')$ for every $s' \in S$. Since t' is regular, it follows that $t'(s) = 0$, i.e. $s \in get_{\rightarrow}(t', 0)$. But s was chosen arbitrarily and thus $get_{\rightarrow}(t, 0) \subseteq get_{\rightarrow}(t', 0)$. Similarly for the converse. So $get_{\rightarrow}(t, i) = get_{\rightarrow}(t', i)$ for $i = 0$.

Suppose that $get_{\rightarrow}(t, j) = get_{\rightarrow}(t', j)$ for some $0 < j < n$ (induction hypothesis).

It must be shown that $get_{\rightarrow}(t, j+1) = get_{\rightarrow}(t', j+1)$. Choose any $s \in get_{\rightarrow}(t, j+1)$. So $t(s) = j + 1$. Suppose that $t'(s) < j + 1$, say $t'(s) = i$ where $i \leq j$. So $s \in get_{\rightarrow}(t', i)$. Then, by the induction hypothesis, $s \in get_{\rightarrow}(t, i)$, i.e. $t(s) = i$. Contradiction. Suppose that $t'(s) > j + 1$, say $t'(s) = j + 1 + i$ where $i > 0$. Since t' is regular, it follows that $get_{\rightarrow}(t', j + 1) \neq \varnothing$. So there must be some $s' \in get_{\rightarrow}(t', j + 1)$ such that $t'(s') < t'(s)$. But t and t' are order-equivalent and thus $t(s') < t(s)$. Since $t(s) = j+1$, it follows that $t(s') < j + 1$. Suppose $t(s') = j$, i.e. $s' \in get_{\rightarrow}(t, j)$. Then, by the induction hypothesis, $s' \in get_{\rightarrow}(t', j)$, i.e. $t'(s) = j$. Contradiction. So $t'(s) \not> j + 1$. But then $t'(s) = j + 1$. So $s \in get_{\rightarrow}(t', j + 1)$. Since s was chosen arbitrarily it follows that $get_{\rightarrow}(t, j + 1) \subseteq get_{\rightarrow}(t', j + 1)$. Similarly for the converse. So $get_{\rightarrow}(t, j + 1) = get_{\rightarrow}(t', j + 1)$. ■

The visualisations that were developed for the semantic representations of information assisted us in pulling together apparently disparate strands from semantic information theory and (the KLM approach to) nonmonotonic logic. Unlike the semantic representations of information afforded by these approaches, t-orderings provide a unified semantic framework for the combination of both definite information and indefinite information. Henceforth, we shall refrain from speaking in terms of the visualisation of a t-ordering and shall simply speak of the levels (boxes) of a t-ordering and of the level (box-label) of a state in a t-ordering.

3.4 Syntactic expressions

In semantic information theory, every semantic representation of definite information is expressible as a sentence in either extensional SDNF or extensional SCNF, and from every such syntactic representation in the knowledge representation language, a semantic representation may be recovered. In the KLM approach to nonmonotonic reasoning, the semantic representation of indefinite information is (in general) not expressible as sentences of the knowledge representation language. Developing normal forms to syntactically express the representation of (semantic) information by t-orderings and recovering t-orderings from such normal forms, is the focus of this section.

The first step in developing normal forms is to decide what set of states the normal form should axiomatise. We could choose to axiomatise absolutely any subset of states, but four possibilities stand out as less arbitrary reference points than the rest: the set of all states, the empty set of states, the set of maximally preferred states at the bottom level, and the set of all included states. While it is tempting to axiomatise the set of all states by ensuring that the normal form is a tautology relative to the set S of states, thereby perturbing the landscape of tautologies in a potentially interesting way, we have chosen instead an alternative more compatible with an agent-oriented approach. The agent's beliefs, determined by the states at the bottom level as will be shown in the next chapter, encompass all the elements that have a bearing on decisions and actions, from the most daring defeasible beliefs to the most conservative and entrenched items of knowledge. Therefore we choose to axiomatise the states at the bottom level, in effect casting the agent's beliefs into a normal form.

Definition 3.25 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. A t-ordering $t \in T_S$ is **expressible** as a sentence $\beta \in L$ iff β is a finite axiomatisation of $bottom(t)$, the set of maximally preferred states. A t-ordering $t \in T_S$ is **recoverable** from a syntactic expression $\beta \in L$ iff t can be reconstructed from β in the manner illustrated below. A t-ordering $t \in T_S$ that is recovered from a syntactic expression $\beta \in L$ of t is denoted by t_β.*

The definite information from the observation that the light is on, as represented by the definite t-ordering $t = \{(11, 0), (10, 0), (01, 4), (00, 4)\}$ in figure 3-4 (on page 67), is expressible as the sentences

$\beta = (P(a) \wedge P(b)) \vee (P(a) \wedge \neg P(b))$ in extensional SDNF, and

$\beta = (P(a) \vee \neg P(b)) \wedge (P(a) \vee P(b))$ in extensional SCNF.

This much is evident. The trick is to cope with states that are tentatively excluded. We shall first formally establish that definite t-orderings are easily expressible and recoverable, and then move on to the more complex question of t-orderings representing indefinite information (i.e. default rules) and t-orderings representing the combination of definite and indefinite information.

Proposition 3.18 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then every definite t-ordering is expressible as a sentence in extensional SDNF or extensional SCNF (or both).*

Proof. Pick any definite t-ordering $t \in T_D$. It must be shown that there exists a sentence in extensional SDNF or extensional SCNF that is a finite axiomatisation of

$bottom(t)$, the set of maximally preferred states. If t is not strongly contradictory then, by proposition 1.2, $bottom(t)$ has an axiomatisation in the form of a sentence β in extensional SDNF. If t is not tautological then, by proposition 3.2, $bottom(t)$ has an axiomatisation in the form of a sentence β in extensional SCNF. Since t was chosen arbitrarily, it follows that every definite t-ordering is expressible as a sentence $\beta \in L$ in extensional SDNF or extensional SCNF (or both). ■

Proposition 3.19 *Let* $M = \langle S, l\rangle$ *be an extensional interpretation of* L *with* $card(S) = n$ *and let* $t \in T_D$ *be a definite t-ordering. Then* t *is recoverable from every syntactic expression of* t *in extensional SDNF or extensional SCNF.*

Proof. It must be shown that from every syntactic expression of t in extensional SDNF or extensional SCNF, a definite t-ordering t' can be constructed such that $t' = t$. Pick any finite axiomatisation $\beta \in L$ of t in extensional SDNF or extensional SCNF. Thus, by proposition 3.18, $bottom(t) = Mod_M(\beta)$ (and $top(t) = NMod_M(\beta)$). Let t'' be a t-ordering that is strongly contradictory. For every $s \in Mod_M(\beta)$ let $t' = pull(t'', s, 0)$. So $bottom(t') = Mod_M(\beta)$ and $top(t') = NMod_M(\beta)$. But then $t' = t$. ■

In general, since every sentence is semantically equivalent to a sentence in extensional SDNF or extensional SCNF (or both), it follows that under an extensional interpretation of L every definite t-ordering t is expressible as a sentence α and recoverable from α, and conversely, that from every sentence $\alpha \in L$ a definite t-ordering t may be recovered that is expressible as a sentence β in extensional SDNF or extensional SCNF, where β is semantically equivalent to α. Note that under an intensional interpretation, this need not be true, i.e. it is not necessarily the case that every definite t-ordering t is expressible as a sentence α. The following example illustrates.

Example 3.8 *Suppose* L *is generated by* $Atom = \{P(a), P'(a)\}$ *and* $M = \langle S, l\rangle$ *is an intensional interpretation such that* $S = \{s_1, s_2, s_3\}$ *and* $l = \{(s_1, I), (s_2, J), (s_3, J)\}$

with $v_I = 11$ *and* $v_J = 01$ *(where the atoms are considered in the given order, so that* 01 *corresponds to the valuation rendering* $P(a)$ *false but* $P'(a)$ *true). Let* $t = \{(s_1, 0), (s_2, 0), (s_3, 3)\}$ *be a definite t-ordering. Then there exists no sentence* α *such that* $Mod_M(\alpha) = bottom(t)$ *where* $bottom(t) = \{s_1, s_2\}$*. This is because if* $s_2 \in Mod_M(\alpha)$ *then it must be the case that* $s_3 \in Mod_M(\alpha)$ *since* $l(s_2) = l(s_3)$ *and thus* $Mod_M(\alpha) \neq \{s_1, s_2\}$*.*

Moving on to the more complex question of t-orderings representing indefinite information, we turn our attention to indefinite t-orderings and a new normal form, called *state description cumulative form*, or *SDCF* for short.

Definition 3.26 *Let* $M = \langle S, l\rangle$ *be an extensional interpretation of* L *with* $card(S) = n$*. A sentence* $\gamma \in L$ *is in* ***state description cumulative form (SDCF)*** *iff* $\gamma = \bigwedge_{i=0}^{n} \beta_i$ *where*

- *each* β_i *is a sentence in extensional SDNF,*
- *for each state description* σ_j *occurring in* β_i*,* σ_j *occurs in every* β_k *where* $k > i$*, and*
- β_n *contains all the state descriptions corresponding to* S*.*

From an indefinite t-ordering t, the β_i components of a sentence in SDCF are constructed so that each β_i axiomatises, starting from level 0, the cumulative subsets of states $get_\uparrow(t, i)$. Consider the default rule 'Normally the light is on, and if the light is on, then normally the fan is on as well', as represented by the indefinite t-ordering $t = \{(11, 0), (10, 1), (01, 2), (00, 2)\}$ in figure 3-5 (on page 68). Then

- $\beta_0 = (P(a) \wedge P(b))$ with $Mod(\beta_0) = \{11\}$,
- $\beta_1 = \beta_0 \vee (P(a) \wedge \neg P(b))$ with $Mod(\beta_1) = \{11, 10\}$,

- $\beta_2 = \beta_1 \vee (\neg P(a) \wedge P(b)) \vee (\neg P(a) \wedge \neg P(b))$ with $Mod(\beta_2) = S$,
- $\beta_3 = \beta_2$ with $Mod(\beta_3) = S$, and
- $\beta_4 = \beta_3$ with $Mod(\beta_4) = S$.

The fact that there are no change after β_2 shows that above the third level, the remaining levels are empty. Taking the conjunction of the β_i components yields the sentence $\gamma = \beta_0 \wedge \beta_1 \wedge \beta_2 \wedge \beta_3 \wedge \beta_4$ in SDCF. It is not hard to see that $Mod(\gamma) = Mod(\beta_0)$.

Lemma 3.4 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = \bigwedge_{i=0}^{n} \beta_i$ of L in SDCF, $Mod_M(\gamma) = Mod_M(\beta_0)$.*

Proof. See proof in appendix A, section A.3. ■

As a dual to state description cumulative form, which is based on extensional SDNF, we introduce *semantic content cumulative form*, or *SCCF* for short, which is a normal form based on extensional SCNF.

Definition 3.27 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. A sentence $\gamma \in L$ is in* ***semantic content cumulative form (SCCF)*** *iff $\gamma = \bigwedge_{i=0}^{m} \beta_i$ for $m < n$ where*

- *each β_i is a sentence in extensional SCNF, and*
- *for each content element ε_j occurring in β_i, ε_j occurs in every β_k where $k < i$.*

In the case of SCCF, the β_i components of a sentence in SCCF are semantically equivalent to the negation of state descriptions and may therefore be read as 'recording' which states are excluded from each cumulative subset $get_{\uparrow}(t, i)$. Consider the same indefinite t-ordering $t = \{(11, 0), (10, 1), (01, 2), (00, 2)\}$ from figure 3-5 (on page 68). Then

- $\beta_0 = \beta_1 \wedge (\neg P(a) \vee P(b))$ with $NMod(\beta_0) = \{01, 00, 10\}$ and
- $\beta_1 = (P(a) \vee \neg P(b)) \wedge (P(a) \vee P(b))$ with $NMod(\beta_1) = \{01, 00\}$.

The fact that there are only two β_i components indicates that by the third level, there are no more states to be excluded, i.e. $get_{\uparrow}(t, 2) = S$. Note that each β_i still axiomatises the cumulative subsets of states $get_{\uparrow}(t, i)$ since $Mod(\beta_0) = \{11\}$ and $Mod(\beta_1) = \{11, 10\}$. Taking the conjunction of the β_i components yields the sentence $\gamma = \beta_0 \wedge \beta_1$ in SCCF. It is not hard to see that $Mod(\gamma) = Mod(\beta_0)$.

Lemma 3.5 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = \bigwedge_{i=0}^{m} \beta_i$ of L in SCCF, $Mod_M(\gamma) = Mod_M(\beta_0)$.*

Proof. See proof in appendix A, section A.3. ■

Lemma 3.5 is key in showing that every indefinite t-ordering is expressible as a sentence in SCCF and, similarly, lemma 3.4 in showing the expressibility of every indefinite t-ordering as a sentence in SDCF.

Proposition 3.20 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Then every indefinite t-ordering is expressible as a sentence in SDCF and in SCCF.*

Proof. Pick any indefinite t-ordering $t \in T_I$. It must be shown that there exists a sentence in SDCF and in SCCF that is a finite axiomatisation of $bottom(t)$, the set of maximally preferred states.

Case 1 (SDCF): Let $X_i = get_{\uparrow}(t, i)$ for $i = 0, 1, \ldots, n$. Since t is indefinite, it follows that each X_i is non-empty. Thus, by proposition 1.2, there is a finite axiomatisation in the form of a sentence β_i in extensional SDNF for each X_i, i.e. $Mod_M(\beta_i) = X_i$. Let $\gamma = \bigwedge_{i=0}^{n} \beta_i$. But then γ is a sentence of L in SDCF. (To see that, pick any state description σ_j occurring in β_i. So there must be some $s \in X_i$ satisfying σ_j, say s_j. But, by construction, $X_i \subseteq X_k$ for every $k > i$. So $s_j \in X_k$ for every $k > i$. But then σ_j occurs in every β_k where $k > i$.) By lemma 3.4, $Mod_M(\gamma) = Mod_M(\beta_0)$. But $Mod_M(\beta_0) = X_0$ and $X_0 = get_{\uparrow}(t, 0) = bottom(t)$. So γ is a finite axiomatisation of $bottom(t)$.

Case 2 (SCCF): By definition, $bottom(t) \neq S$. Let $m = first_{\downarrow}(t) - 1$ and let $X_i = get_{\uparrow}(t, i)$ for $i = 0, 1, \ldots, m$. So each $X_i \subset S$. Thus, by proposition 3.2, there is a finite axiomatisation in the form of a sentence β_i in extensional SCNF for each X_i such that $Mod_M(\beta_i) = X_i$. Let $\gamma = \bigwedge_{i=0}^{m} \beta_i$. But then γ is a sentence of L in SCCF. (To see that, pick any content element ε_j occurring in β_i. So there must be some $s \in S - X_i$ that fails to satisfy ε_j, say s_j. But, by construction, $X_i \subseteq X_k$ for every $k > i$, i.e. $S - X_i \subseteq S - X_k$ for every $k < i$. So $s_j \in X_k$ for every $k < i$. But then ε_j occurs in every β_k where $k < i$.) By lemma 3.5, $Mod_M(\gamma) = Mod_M(\beta_0)$. But $Mod_M(\beta_0) = X_0$ and $X_0 = get_{\uparrow}(t, 0) = bottom(t)$. So γ is a finite axiomatisation of $bottom(t)$. ∎

Although an indefinite t-ordering is expressible by every sentence that axiomatises the set of maximally preferred states (and not just by a sentence in SDCF or SCCF), it is not the case that the t-ordering can be recovered from every such sentence. However, an indefinite t-ordering $t \in T_I$ is recoverable from every syntactic expression of t in SDCF or SCCF as shown by the following proposition.

Proposition 3.21 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and let $t \in T_I$ be an indefinite t-ordering. Then t is recoverable from every syntactic expression of t in SDCF or SCCF.*

Proof. It must be shown that from every syntactic expression of t in SDCF or SCCF, an indefinite t-ordering t' can be constructed such that $t' = t$.

Case 1 (SDCF): Pick any syntactic expression $\gamma \in L$ of t in SDCF. By definition of SDCF, $\gamma = \bigwedge_{i=0}^{n} \beta_i$ where each β_i is a sentence in extensional SDNF. Let $X_i = Mod_M(\beta_i)$. Since γ is in SDCF, it follows that for each state description σ_j occurring in β_i, σ_j

occurs in every β_k where $k > i$. Thus $X_i \subseteq X_k$ for every $k > i$. Let $Y_0 = X_0$ and $Y_i = X_i - X_{i-1}$ for $i = 1, \ldots, n$. Let t' be a t-ordering that is tautological. For each Y_i, let $t' = push(t', s_j, i)$ for every $s_j \in Y_i$. It is claimed that $t' = t$. To see this, pick any $s \in S$. By construction of γ, each β_i axiomatises $get_{\uparrow}(t, i)$. So $X_i = get_{\uparrow}(t, i)$. Suppose $t'(s) = 0$. So $s \in Y_0 = X_0$. But then $s \in get_{\uparrow}(t, 0)$, i.e. $t(s) = 0$. Suppose $t'(s) = i$ for $i > 0$. So $s \in Y_i = X_i - X_{i-1}$. But then $s \in get_{\uparrow}(t, i)$ and $s \notin get_{\uparrow}(t, i-1)$. So $s \in get_{\longrightarrow}(t, i)$, i.e. $t(s) = i$. Hence $t' = t$.

Case 2 (SCCF): Pick any syntactic expression $\gamma \in L$ of t in SCCF. By definition of SCCF, $\gamma = \bigwedge_{i=0}^{m} \beta_i$ where each β_i is a sentence in extensional SCNF. Let $X_i = Mod_M(\beta_i)$. Since γ is in SCCF, it follows that for each content element ε_j occurring in β_i, ε_j occurs in every β_k where $k < i$, whence it follows that $S - X_i \subseteq S - X_k$ for every $k < i$, i.e. $X_k \subseteq X_i$ for every $k < i$. Let $Y_0 = X_0$ and $Y_i = X_i - X_{i-1}$ for $i = 1, 2, \ldots, m$ and let $Y_{m+1} = S - X_m$. Let t' be a t-ordering that is tautological. For each Y_i, let $t' = push(t', s_j, i)$ for every $s_j \in Y_i$. It is claimed that $t' = t$. To see this, pick any $s \in S$. By construction of γ, each β_i axiomatises $get_{\uparrow}(t, i)$. So $X_i = get_{\uparrow}(t, i)$. Suppose $t'(s) = 0$. So $s \in Y_0 = X_0$. But then $s \in get_{\uparrow}(t, 0)$, i.e. $t(s) = 0$. Suppose $t'(s) = i$ for $0 < i \leq m$. So $s \in Y_i = X_i - X_{i-1}$. But then $s \in get_{\uparrow}(t, i)$ and $s \notin get_{\uparrow}(t, i-1)$. So $s \in get_{\longrightarrow}(t, i)$, i.e. $t(s) = i$. Suppose $t'(s) = m + 1$. So $s \in S - X_m$. By construction of γ, $m = first_{\downarrow}(t) - 1$. But then $s \in get_{\longrightarrow}(t, m+1)$ otherwise $s \in get_{\longrightarrow}(t, j)$ for some $j \leq m$ in which case $s \in X_i$ for every $j < i \leq m$ so that $t'(s) \neq m + 1$ or else $s \in get_{\longrightarrow}(t, j)$ for some $j > m + 1$ in which case $first_{\downarrow}(t) \neq m + 1$. Both cases lead to contradictions. Thus $t(s) = m + 1$. Hence $t' = t$.

■

Having established that indefinite t-orderings are expressible and recoverable, we turn our attention to regular t-orderings representing the combination of definite and indefinite information and in particular, pure regular t-orderings. We shall again introduce new normal forms, the first of which is called *state description templated form*, or *SDTF* for short.

Definition 3.28 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. A sentence γ of L is in* ***state description templated form (SDTF)*** *iff* $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ *where*

- *each β_i is a sentence in extensional SDNF,*
- *for each state description σ_j occurring in β_i, σ_j occurs in every β_k where $k > i$, and*
- *β_n contains all the state descriptions corresponding to S.*

The β_i components of a sentence in SDTF are constructed in exactly the same manner as the β_i components of a sentence in SDCF. Consider the pure regular t-ordering $t = \{(11, 0), (10, 1), (01, 4), (00, 4)\}$ in figure 3-6 (on page 69). Then

- $\beta_0 = (P(a) \wedge P(b))$ with $Mod(\beta_0) = \{11\}$,
- $\beta_1 = \beta_0 \vee (P(a) \wedge \neg P(b))$ with $Mod(\beta_1) = \{11, 10\}$,
- $\beta_2 = \beta_1$ with $Mod(\beta_2) = \{11, 10\}$,
- $\beta_3 = \beta_2$ with $Mod(\beta_3) = \{11, 10\}$, and
- $\beta_4 = \beta_3 \vee (\neg P(a) \wedge P(b)) \vee (\neg P(a) \wedge \neg P(b))$ with $Mod(\beta_4) = S$,

Taking the conjunction of all the β_i components yields the sentence $\gamma' = \beta_0 \wedge \beta_1 \wedge \beta_2 \wedge \beta_3 \wedge \beta_4$, which is in SDCF and thus axiomatises the set of maximally preferred states, i.e. $Mod(\gamma') = \{11\}$. Taking the disjunction of components β_0 to β_3 yields the sentence $\gamma'' = \beta_0 \vee \beta_1 \vee \beta_2 \vee \beta_3$. It is not hard to see that $Mod(\gamma'') = \{11, 10\}$. The sentence

$\gamma = \gamma'' \wedge \gamma'$ is in SDTF and since $Mod(\gamma) = \{11\}$, it follows that γ axiomatises the set of maximally preferred states. Note that the disjunctive part γ'' of the sentence γ in SDTF determines the set of excluded states (since it axiomatises the set of included states) and thus records the definite information represented by t.

Lemma 3.6 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ of L in SDTF, $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$.*

Proof. See proof in appendix A, section A.3. ■

The second normal form is called *semantic content templated form*, or *SCTF* for short. In contrast to SDTF, which is based on extensional SDNF, SCTF is based on extensional SCNF.

Definition 3.29 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. A sentence γ of L is in* ***semantic content templated form (SCTF)*** *iff $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ where*

- *each β_i is a sentence in extensional SCNF, and*
- *for each content element ε_j occurring in β_i, ε_j occurs in every β_k where $k < i$.*

As was the case with SDTF, the β_i components of a sentence in SCTF are constructed in exactly the same manner as the β_i components of a sentence in SCCF. Consider again the pure regular t-ordering $t = \{(11, 0), (10, 1), (01, 4), (00, 4)\}$ in figure 3-6 (on page 69). Then

- $\beta_0 = \beta_1 \wedge (\neg P(a) \vee P(b))$ with $NMod(\beta_0) = \{01, 00, 10\}$,
- $\beta_1 = \beta_2$ with $NMod(\beta_1) = \{00, 10\}$,

- $\beta_2 = \beta_3$ with $NMod(\beta_2) = \{00, 10\}$, and
- $\beta_3 = (P(a) \vee \neg P(b)) \wedge (P(a) \vee P(b))$ with $NMod(\beta_1) = \{01, 00\}$.

The fact that there are four β_i components indicates that by the fifth level, which is the top level, there are no more states to be excluded, i.e. $get_{\uparrow}(t, 4) = S$. Taking the conjunction of components β_0 to β_3 yields the sentence $\gamma' = \beta_0 \wedge \beta_1 \wedge \beta_2 \wedge \beta_3$, which is in SCCF (with $m = n - 1$) and axiomatises the set of maximally preferred states, i.e. $Mod(\gamma') = \{11\}$. Taking the disjunction of components β_0 to β_3 yields the sentence $\gamma'' = \beta_0 \vee \beta_1 \vee \beta_2 \vee \beta_3$. It is not hard to see that $Mod(\gamma'') = \{11, 10\}$. The sentence $\gamma = \gamma'' \wedge \gamma'$ is in SCTF and since $Mod(\gamma) = \{11\}$, it follows that γ axiomatises the set of maximally preferred states. As was the case with SDTF, the disjunctive part γ'' of the sentence γ in SCTF determines the set of excluded states (since it axiomatises the set of included states) and thus records the definite information represented by t.

Lemma 3.7 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ of L in SCTF, $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$.*

Proof. See proof in appendix A, section A.3. ■

Proposition 3.22 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then every pure regular t-ordering is expressible as a sentence in SDTF and in SCTF.*

Proof. Pick any pure regular t-ordering $t \in T_E$. It must be shown that there exists a sentence in SDTF and in SCTF that is a finite axiomatisation of $bottom(t)$, the set of maximally preferred states.

Case 1 (SDTF): Let $X_i = get_{\uparrow}(t, i)$ for $i = 0, 1, \ldots, n$. Since t is pure regular, it follows that each X_i is non-empty. Thus, by proposition 1.2, there is a finite axiomatisation in the form of a sentence β_i in extensional SDNF for each X_i, i.e. $Mod_M(\beta_i) = X_i$.

Let $\gamma' = \bigwedge_{i=0}^{n} \beta_i$. But then γ' is a sentence of L in SDCF (as shown in proposition 3.20, case 1). But then, by proposition 3.20, γ' is a finite axiomatisation of $bottom(t)$, i.e. $Mod_M(\gamma') = bottom(t)$. Let $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$. By lemma 3.6, $Mod_M(\gamma'') = Mod_M(\beta_{n-1})$, i.e. $Mod_M(\gamma'') = get_{\uparrow}(t, n-1)$. So $Mod_M(\gamma'') \cap Mod_M(\gamma') = bottom(t)$. Let $\gamma = \gamma'' \wedge \gamma'$. But then γ is in SDTF and $Mod_M(\gamma) = bottom(t)$. So γ is a finite axiomatisation of $bottom(t)$.

Case 2 (SCTF): Since t is pure regular, it follows that t is not tautological and $first_{\downarrow}(t') = n$. Let $X_i = get_{\uparrow}(t, i)$ for $i = 0, 1, \ldots, n-1$. So each $X_i \subset S$. Thus, by proposition 3.2, there is a finite axiomatisation in the form of a sentence β_i in extensional SCNF for each X_i such that $Mod_M(\beta_i) = X_i$. Let $\gamma' = \bigwedge_{i=0}^{n-1} \beta_i$. But then γ' is a sentence of L in SCCF (as shown in proposition 3.20, case 2) with $m = first_{\downarrow}(t) - 1 = n-1$. But then, by proposition 3.20, γ' is a finite axiomatisation of $bottom(t)$, i.e. $Mod_M(\gamma') = bottom(t)$. Let $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$. By lemma 3.7, $Mod_M(\gamma'') = Mod_M(\beta_{n-1})$, i.e. $Mod_M(\gamma'') = get_{\uparrow}(t, n-1)$. So $Mod_M(\gamma'') \cap Mod_M(\gamma') = bottom(t)$. Let $\gamma = \gamma'' \wedge \gamma'$. But then γ is in SCTF and $Mod_M(\gamma) = bottom(t)$. So γ is a finite axiomatisation of $bottom(t)$. ■

Although a pure regular t-ordering is expressible by every sentence that axiomatises the set of maximally preferred states (and not just by a sentence in SDTF or SCTF), it is not the case that the t-ordering can be recovered from every such sentence. However, a pure regular t-ordering $t \in T_E$ is recoverable from every syntactic expression of t in SDTF or SCTF as shown by the following proposition.

Proposition 3.23 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and let $t \in T_E$ be a pure regular t-ordering. Then t is recoverable from every syntactic expression of t in SDTF or SCTF.*

Proof. It must be shown that from every syntactic expression of t in SDTF or SCTF, a t-ordering t' can be constructed such that $t' = t$.

Case 1 (SDTF): Pick any syntactic expression $\gamma \in L$ of t in SDTF. By definition $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ where each β_i is a sentence in extensional SDNF. Let $X_i = Mod_M(\beta_i)$. Now let $\gamma = \gamma'' \wedge \gamma'$ where $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$ and $\gamma' = \bigwedge_{i=0}^{n} \beta_i$. Since γ' is in SDCF, it follows that for each state description σ_j occurring in β_i, σ_j occurs in every β_k where $k > i$. Thus $X_i \subseteq X_k$ for every $k > i$. Let $Y_0 = X_0$ and $Y_i = X_i - X_{i-1}$ for $i = 1, 2, \ldots, n$. Let t' be a t-ordering that is tautological. For each Y_i, let $t' = push(t', s_j, i)$ for every $s_j \in Y_i$. It is claimed that $t' = t$. To see this, pick any $s \in S$. By construction of γ', each β_i axiomatises $get_\uparrow(t, i)$. So $X_i = get_\uparrow(t, i)$. Suppose $t'(s) = 0$. So $s \in Y_0 = X_0$. But then $s \in get_\uparrow(t, 0)$, i.e. $t(s) = 0$. Suppose $t'(s) = i$ for $i > 0$. So $s \in Y_i = X_i - X_{i-1}$. But then $s \in get_\uparrow(t, i)$ and $s \notin get_\uparrow(t, i-1)$. So $s \in get_\longrightarrow(t, i)$, i.e. $t(s) = i$. Hence $t' = t$.

Case 2 (SCTF): Pick any syntactic expression $\gamma \in L$ of t in SCTF. By definition $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ where each β_i is a sentence in extensional SCNF. Let $X_i = Mod_M(\beta_i)$. Now let $\gamma = \gamma'' \wedge \gamma'$ where $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$ and $\gamma' = \bigwedge_{i=0}^{n-1} \beta_i$. Since γ' is in SCCF with $m = first_\downarrow(t) - 1 = n - 1$, it follows that for each content element ε_j occurring in β_i, ε_j occurs in every β_k where $k < i$, whence it follows that $S - X_i \subseteq S - X_k$ for every $k < i$, i.e. that $X_k \subseteq X_i$ for every $k < i$. Let $Y_0 = X_0$ and $Y_i = X_i - X_{i-1}$ for $i = 1, 2, \ldots, n-1$ and let $Y_n = S - X_{n-1}$. Let t' be a t-ordering that is tautological. For each Y_i, let $t' = push(t', s_j, i)$ for every $s_j \in Y_i$. It is claimed that $t' = t$. To see this, pick any $s \in S$. By construction of γ', each β_i axiomatises $get_\uparrow(t, i)$. So $X_i = get_\uparrow(t, i)$. Suppose $t'(s) = 0$. So $s \in Y_0 = X_0$. But then $s \in get_\uparrow(t, 0)$, i.e. $t(s) = 0$. Suppose $t'(s) = i$ for $0 < i \leq n - 1$. So $s \in Y_i = X_i - X_{i-1}$. But then $s \in get_\uparrow(t, i)$ and $s \notin get_\uparrow(t, i-1)$. So $s \in get_\longrightarrow(t, i)$, i.e. $t(s) = i$. Suppose $t'(s) = n$. So $s \in S - X_{n-1}$. But then $s \in get_\longrightarrow(t, n)$ otherwise $s \in get_\longrightarrow(t, j)$ for some $j \leq n - 1$ in which case $s \in X_i$ for every $j < i \leq n - 1$ so that $t'(s) \neq n$, leading to a contradiction. Thus $t(s) = n$. Hence $t' = t$.

■

The normal forms that were developed to syntactically express the semantic representation of information by t-orderings and from which t-orderings are recoverable, illustrate that the semantic representations of information afforded by semantic information theory and (the KLM approach to) nonmonotonic logic are weaker than that of t-orderings. Definite t-orderings are expressible and recoverable through extensional SDNF (extensional SCNF) in a way similar to the dichotomous 'partitions' associated with semantic information theory. Indefinite t-orderings, which are expressible and recoverable through SDCF (SCCF), represent the strict modular partial orders associated with (the KLM approach to) nonmonotonic logic. In addition, pure regular t-orderings, which are expressible through SDTF (SCTF), have no counterpart as a semantic representation of information in either of these approaches. In the next section we return to one of the cornerstones of t-orderings, namely, semantic information theory.

3.5 Semantic information content

The concept of the information carried by a sentence (as the set of states excluded by the sentence) and the concept of the amount of information carried by a sentence (as a numerical measure of its information content) were introduced in section 3.1. In this section, the notion of (semantic) information content for t-orderings will be defined. For completeness, numerical measures for the notion of information content will be defined too, but, because our approach is qualitative in nature, will not be developed any further. The information content of a t-ordering may be formulated in terms of the exclusion of states, either definitely or tentatively.

Definition 3.30 *Let $t \in T_S$. The definite information represented by t, or the* ***definite content*** *of t, is defined as $Cont_D(t) = top(t)$.*

Definition 3.31 *Let $t \in T_S$. The amount of definite information represented by t, or the*

definite measure *of* t, *is defined as* $cont_D(t) = \sum_{i=1}^{m}(cont(s_i) \mid s_i \in Cont_D(t))$ *where* $m = card(Cont_D(t))$.

Proposition 3.24 *Let* $t \in T_S$. *Then the following holds for* $0 \leq i < n$:

1. $\varnothing \subseteq Cont_D(t) \subseteq S$
2. *if* t *is indefinite, then* $Cont_D(t) = \varnothing$
3. $Cont_D(t) = S$ *iff* t *is strongly contradictory*

Proof. The results follow directly from the definitions. ■

Having defined the notion of definite content, we are now in a position to compare the notion of the information carried by the syntactic expression of a t-ordering with the notion of the definite content of the t-ordering recovered from such a sentence. As mentioned in section 3.4, under an extensional interpretation of L every definite t-ordering is expressible as and recoverable from every sentence $\alpha \in L$ and not merely sentences in extensional SDNF or extensional SCNF.

Proposition 3.25 *Let* $M = \langle S, l\rangle$ *be an extensional interpretation of* L *with* $card(S) = n$. *Let* $t_\alpha \in T_D$ *be the definite t-ordering recovered from a sentence* $\alpha \in L$. *Then* $Cont_M(\alpha) = Cont_D(t_\alpha)$.

Proof. Pick any $s \in Cont_M(\alpha)$. By definition of $Cont_M$, $s \in NMod_M(\alpha)$. But α is a syntactic expression of t_α and thus $Mod_M(\alpha) = bottom(t_\alpha)$. Since t_α is definite it follows that $s \in top(t_\alpha)$. Since s was chosen arbitrarily it follows that $Cont_M(\alpha) \subseteq Cont_D(t_\alpha)$. Conversely, pick any $s \in Cont_D(t_\alpha)$. By definition of $Cont_D$, $s \in top(t_\alpha)$. But α is a syntactic expression of t_α and so $Mod_M(\alpha) = bottom(t_\alpha)$. Since t_α is definite it follows that $s \in NMod_M(\alpha)$, i.e. $s \in Cont_M(\alpha)$. Since s was chosen arbitrarily it follows that $Cont_D(t_\alpha) \subseteq Cont_M(\alpha)$. Hence $Cont_M(\alpha) = Cont_D(t_\alpha)$. ■

When restricting ourselves to definite t-orderings, the basic results from semantic information theory apply as illustrated by the following proposition.

Proposition 3.26 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\alpha, t_\beta \in T_D$ be the definite t-orderings recovered from $\alpha, \beta \in L$ respectively. Then $\alpha \models_M \beta$ iff $Cont_D(t_\alpha) \supseteq Cont_D(t_\beta)$.*

Proof. Suppose $\alpha \models_M \beta$. So $Mod_M(\alpha) \subseteq Mod_M(\beta)$. But then $NMod_M(\alpha) \supseteq NMod_M(\beta)$, and, by definition of $Cont_M$, $Cont_M(\alpha) \supseteq Cont_M(\beta)$. But t_α and t_β are the definite t-orderings recovered from α and β respectively and, thus, by proposition 3.25, it follows that $Cont_D(t_\alpha) \supseteq Cont_D(t_\beta)$. Conversely, suppose $Cont_D(t_\alpha) \supseteq Cont_D(t_\beta)$. But t_α and t_β are the definite t-orderings recovered from α and β respectively and, thus, by proposition 3.25, it follows that $Cont_M(\alpha) \supseteq Cont_M(\beta)$. So, by definition of $Cont_M$, $NMod_M(\alpha) \supseteq NMod_M(\beta)$. But then $Mod_M(\alpha) \subseteq Mod_M(\beta)$, i.e. $\alpha \models_M \beta$. ■

The result shows that the information carried by a sentence coincides with the definite content of the definite t-ordering recovered from that sentence. The other results from semantic information theory (proposition 3.4) similarly apply under the restriction to definite t-orderings.

We now turn our attention to pure regular t-orderings and formally show that the disjunctive part of a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ in SDTF determines the set of excluded states and thus records the definite information represented by t_γ. It is done by showing that the information carried by the sentence $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$ coincides with the definite content of the pure regular t-ordering recovered from γ. Similarly, it will be shown that the disjunctive part of a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ in SCTF records the definite information represented by t_γ.

Proposition 3.27 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_E$ be the pure regular t-ordering recovered from a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge$*

$(\bigwedge_{i=0}^{n} \beta_i)$ in SDTF. Then $Cont_M(\bigvee_{i=0}^{n-1} \beta_i) = Cont_D(t_\gamma)$.

Proof. Let $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$. By definition of SDTF, each β_i is a sentence in extensional SDNF and by construction of γ, each β_i axiomatises $get_\uparrow(t_\gamma, i)$. So $Mod_M(\beta_{n-1}) = get_\uparrow(t_\gamma, n-1)$. But $Mod_M(\gamma'') = Mod_M(\beta_{n-1})$ by lemma 3.6 and thus $Mod_M(\gamma'') = get_\uparrow(t_\gamma, n-1)$. Pick any $s \in Cont_M(\gamma'')$. By definition $s \in NMod_M(\gamma'')$, i.e. $s \in S - Mod_M(\gamma'')$ and thus $s \in S - get_\uparrow(t_\gamma, n-1)$. So $s \in top(t_\gamma)$. But then $s \in Cont_D(t_\gamma)$. Since s was chosen arbitrarily it follows that $Cont_M(\gamma'') \subseteq Cont_D(t_\gamma)$. Conversely, pick any $s \in Cont_D(t_\gamma)$. By definition $s \in top(t_\gamma)$, i.e. $s \notin get_\uparrow(t_\gamma, n-1)$. So $s \notin Mod_M(\gamma'')$ and thus $s \in NMod_M(\gamma'')$. But then $s \in Cont_M(\gamma'')$. Since s was chosen arbitrarily it follows that $Cont_D(t_\gamma) \subseteq Cont_M(\gamma'')$. But then $Cont_M(\bigvee_{i=0}^{n-1} \beta_i) = Cont_D(t_\gamma)$. ■

Proposition 3.28 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_E$ be the pure regular t-ordering recovered from a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ in SCTF. Then $Cont_M(\bigvee_{i=0}^{n-1} \beta_i) = Cont_D(t_\gamma)$.*

Proof. Let $\gamma'' = \bigvee_{i=0}^{n-1} \beta_i$. By definition of SCTF, each β_i is a sentence in extensional SCNF and by construction of γ, each β_i axiomatises $get_\uparrow(t_\gamma, i)$. So $Mod_M(\beta_{n-1}) = get_\uparrow(t_\gamma, n-1)$. But $Mod_M(\gamma'') = Mod_M(\beta_{n-1})$ by lemma 3.7 and thus $Mod_M(\gamma'') = get_\uparrow(t_\gamma, n-1)$. Pick any $s \in Cont_M(\gamma'')$. By definition $s \in NMod_M(\gamma'')$, i.e. $s \in S - Mod_M(\gamma'')$ and thus $s \in S - get_\uparrow(t_\gamma, n-1)$. So $s \in top(t_\gamma)$. But then $s \in Cont_D(t_\gamma)$. Since s was chosen arbitrarily it follows that $Cont_M(\gamma'') \subseteq Cont_D(t_\gamma)$. Conversely, pick any $s \in Cont_D(t_\gamma)$. By definition $s \in top(t_\gamma)$, i.e. $s \notin get_\uparrow(t_\gamma, n-1)$. So $s \notin Mod_M(\gamma'')$ and thus $s \in NMod_M(\gamma'')$. But then $s \in Cont_M(\gamma'')$. Since s was chosen arbitrarily it follows that $Cont_D(t_\gamma) \subseteq Cont_M(\gamma'')$. But then $Cont_M(\bigvee_{i=0}^{n-1} \beta_i) = Cont_D(t_\gamma)$. ■

As is to be expected, the information carried by a sentence γ in SDCF or SCCF does not coincide with the definite content of the indefinite t-ordering recovered from γ.

Proposition 3.29 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_I$ be the indefinite t-ordering recovered from a sentence γ in SDCF or SCCF. Then $Cont_M(\gamma) \neq Cont_D(t_\gamma)$.*

Proof. Since t_γ is indefinite it follows that $Cont_D(t_\gamma) = \varnothing$ and $bottom(t_\gamma) \neq S$. By proposition 3.20, $Mod_M(\gamma) = bottom(t_\gamma)$. Since $bottom(t_\gamma) \neq S$ it follows that $Mod_M(\gamma) \neq S$, i.e. $NMod_M(\gamma) \neq \varnothing$, i.e. $Cont_M(\gamma) \neq \varnothing$. But then $Cont_M(\gamma) \neq Cont_D(t_\gamma)$. ■

This is because a sentence γ in SDCF or SCCF determines the maximal set of tentatively excluded states (since it axiomatises the set of maximally preferred states) and thus records the indefinite information represented by t_γ. The indefinite information represented by a t-ordering is captured by the notion of *indefinite content*. We allow for varying degrees of tentative exclusion, the usefulness of which will be illustrated in section 4.3 where the information-theoretic semantics for epistemic logic (Labuschagne and Ferguson, 2002) is applied. The indefinite content of a t-ordering t to degree 0 will often be referred to simply as the indefinite content of t.

Definition 3.32 *Let $t \in T_S$. The indefinite information represented by t, or the **indefinite content** of t, to degree i, is defined as $Cont_i(t) = get(t, i+1, n)$ for $0 \leq i < n$.*

Definition 3.33 *Let $t \in T_S$. The amount of indefinite information represented by t, or the **indefinite measure** of t, to degree i, is defined as $cont_i(t) = \sum_{j=1}^{m}(cont(s_j) \mid s_j \in Cont_i(t))$ where $m = card(Cont_i(t))$.*

Proposition 3.30 *Let $t \in T_S$. Then the following holds for $0 \leq i < n$:*

1. $\varnothing \subseteq Cont_i(t) \subseteq S$
2. $Cont_0(t) = \varnothing$ *iff t is tautological*

3. *if t is weakly contradictory, then $Cont_0(t) = S$*

4. *if t is definite, then $Cont_0(t) = Cont_D(t)$*

5. $Cont_i(t) \subseteq Cont_{i-1}(t)$

Proof. The results follow directly from the definitions. ■

Having defined the notion of indefinite content, it can now be formally shown that a sentence γ in SDCF or SCCF determines the maximal set of tentatively excluded states (since it axiomatises the set of maximally preferred states) and thus records the indefinite information represented by t_γ. It is done by showing that the information carried by γ coincides with the indefinite content of the indefinite t-ordering recovered from γ. In addition, it will be shown that the conjunctive parts of a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ in SDTF and a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ in SCTF record the indefinite information represented by the pure regular t-ordering recovered from γ.

Proposition 3.31 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_I$ be the indefinite t-ordering recovered from a sentence γ in SDCF or SCCF. Then $Cont_M(\gamma) = Cont_0(t_\gamma)$.*

Proof. A sentence γ in SDCF or SCCF axiomatises $bottom(t_\gamma)$, by proposition 3.20. So $Mod_M(\gamma) = bottom(t_\gamma)$. But then $NMod_M(\gamma) = S - bottom(t_\gamma) = get(t_\gamma, 1, n)$. So $Cont_M(\gamma) = get(t_\gamma, 1, n)$. But $Cont_0(t_\gamma) = get(t_\gamma, 1, n)$ and thus $Cont_M(\gamma) = Cont_0(t_\gamma)$. ■

Proposition 3.32 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_E$ be the pure regular t-ordering recovered from a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ in SDTF. Then $Cont_M(\bigwedge_{i=0}^{n} \beta_i) = Cont_0(t_\gamma)$.*

Proof. Let $\gamma' = \bigwedge_{i=0}^{n} \beta_i$. By definition of SDTF, each β_i is a sentence in extensional SDNF and by construction of γ, each β_i axiomatises $get_{\uparrow}(t_\gamma, i)$. Since γ' is a sentence in SDCF, it follows by lemma 3.4 that $Mod_M(\gamma') = Mod_M(\beta_0)$. But $Mod_M(\beta_0) = get_{\uparrow}(t_\gamma, 0)$ and thus $Mod_M(\gamma') = bottom(t_\gamma)$. But then $NMod_M(\gamma') = S-bottom(t_\gamma) = get(t_\gamma, 1, n)$. So $Cont_M(\gamma') = get(t_\gamma, 1, n)$. But $Cont_0(t_\gamma) = get(t_\gamma, 1, n)$ and thus $Cont_M(\gamma') = Cont_0(t_\gamma)$. Hence $Cont_M(\bigwedge_{i=0}^{n} \beta_i) = Cont_0(t_\gamma)$. ■

Proposition 3.33 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_\gamma \in T_E$ be the pure regular t-ordering recovered from a sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ in SCTF. Then $Cont_M(\bigwedge_{i=0}^{n-1} \beta_i) = Cont_0(t_\gamma)$.*

Proof. Let $\gamma' = \bigwedge_{i=0}^{n-1} \beta_i$. By definition of SCTF, each β_i is a sentence in extensional SCNF and by construction of γ, each β_i axiomatises $get_{\uparrow}(t_\gamma, i)$. Since γ' is a sentence in SCCF with $m = first_{\downarrow}(t) - 1 = n - 1$, it follows by lemma 3.5 that $Mod_M(\gamma') = Mod_M(\beta_0)$. But $Mod_M(\beta_0) = get_{\uparrow}(t_\gamma, 0)$ and thus $Mod_M(\gamma') = bottom(t_\gamma)$. But then $NMod_M(\gamma') = S-bottom(t_\gamma) = get(t_\gamma, 1, n)$. So $Cont_M(\gamma') = get(t_\gamma, 1, n)$. But $Cont_0(t_\gamma) = get(t_\gamma, 1, n)$ and thus $Cont_M(\gamma') = Cont_0(t_\gamma)$. Hence $Cont_M(\bigwedge_{i=0}^{n-1} \beta_i) = Cont_0(t_\gamma)$. ■

The notion of indefinite content has no counterpart in semantic information theory, one of the cornerstones of t-orderings. On the other hand, the notion of definite content is closely connected with semantic information theory in the sense that the information carried by a sentence coincides with the definite content of the definite t-ordering recovered from that sentence. In the next section we return to the other cornerstone of t-orderings, namely, nonmonotonic logic.

3.6 Reasoning with t-orderings

As mentioned earlier, nonmonotonic reasoning is concerned with indefinite information, the semantic representation of which by some form of ordering is utilised in providing a mechanism for the formation of defeasible conjectures. In the KLM approach to nonmonotonic reasoning, the semantic representation of indefinite information is by means of a total preorder or its corresponding strict modular partial order. Both orderings produce a strict linearly ordered partition on the set of states. T-orderings may be seen as a distribution of the elements of such a partition of states over a template that maintains the strict linear order on the partition of states (but keeping the top-most position for states that are definitely excluded). The mechanism involved in nonmonotonic reasoning is essentially a movement from a hypothesis to those models that are minimal with respect to the ordering representing the indefinite information and from there to a defeasible consequence that is true at each of the minimal models.

Definition 3.34 *Let $t \in T_S$. For every $Y \subseteq S$, an element $s \in Y$ is **minimal** in Y with respect to t iff there is no $s' \in Y$ such that $t(s') \leq t(s)$, unless $t(s) \leq t(s')$. The set of minimal elements in Y with respect to t is denoted by $Min_t(Y)$.*

Definition 3.35 *A **templated interpretation** of L is a triple $T = \langle S, t, l \rangle$ such that*

- *S is a finite non-empty set of states with $card(S) = n$,*
- *$l : S \to U_T$ is a labelling function, and*
- *$t \in T_S$ is a t-ordering.*

As before, the notions of satisfaction and of a model are defined in exactly the same way for templated interpretations as for possible worlds interpretations. The notion of a minimal model is defined as for ranked interpretations, but using t-orderings instead of strict modular partial orders (or total preorders). Thus, given a templated interpretation $T = \langle S, t, l \rangle$, a state $s \in S$ is called a *minimal model* of a sentence α iff $s \in Min_t(Mod_T(\alpha))$. Given T, the set of minimal models of α is denoted by $Min_t(\alpha)$ and the defeasible consequence relation induced by T on L defined in exactly the same way as for ranked interpretations, i.e. $\alpha \mid\!\sim_T \beta$ iff $Min_t(\alpha) \subseteq Mod_T(\beta)$. When the t-ordering t is definite, the defeasible consequence relation $\mid\!\sim_T$ induced by a templated interpretation $T = \langle S, t, l \rangle$ is exactly the semantic consequence relation $\models_T$. More importantly, defeasible consequence relations induced by templated interpretations are rational.

Proposition 3.34 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $P = \langle S, R, l \rangle$ be the corresponding (finite) ranked interpretation of L where R is the total preorder induced by t. Then $\alpha \mid\!\sim_T \beta$ iff $\alpha \mid\!\sim_P \beta$ for every $\alpha, \beta \in L$.*

Proof. It must be shown that $Min_t(\alpha) = Min_R(\alpha)$. Choose any $s \in Min_t(\alpha)$, i.e. $s \in Min_t(Mod_T(\alpha))$. So $s \in Mod_T(\alpha)$ and there is no $s' \in Mod_T(\alpha)$ such that $t(s') \leq t(s)$, unless $t(s) \leq t(s')$. But then, by proposition 3.14, there can be no $s' \in Mod_T(\alpha)$ such that $(s', s) \in R$, unless $(s, s') \in R$. Since $Mod_T(\alpha) = Mod_P(\alpha)$, it follows that $s \in Min_R(Mod_P(\alpha))$, i.e. $s \in Min_R(\alpha)$. So $Min_t(\alpha) \subseteq Min_R(\alpha)$. Similarly, it can be shown that $Min_R(\alpha) \subseteq Min_t(\alpha)$. Hence $Min_t(\alpha) = Min_R(\alpha)$. ■

Proposition 3.35 *If $T = \langle S, t, l \rangle$ is a templated interpretation of L, then the defeasible consequence relation $\mid\!\sim_T$ on L induced by T is rational. Conversely, if a binary relation $\mid\!\sim$ on L is a rational consequence relation, then it can be defined by some templated interpretation of L.*

Proof. Suppose that $T = \langle S, t, l \rangle$ is a templated interpretation of L. Then, by proposition 3.34, a corresponding (finite) ranked interpretation $P = \langle S, R, l \rangle$ of L may be defined such that $\mid\!\sim_T = \mid\!\sim_P$. But $\mid\!\sim_P$ is rational by theorem 3.1 and hence, it follows that $\mid\!\sim_T$ is rational.

Conversely, suppose that $\mid\!\sim$ is a rational consequence relation on L. Then, by theorem 3.1, it can be defined by some ranked interpretation $P = \langle S, R, l\rangle$. Since L is finitely generated, $P = \langle S, R, l\rangle$ can be defined such that S is finite. A templated interpretation $T = \langle S, t, l\rangle$ can now be defined from $P = \langle S, R, l\rangle$ by taking $t(s) \leq t(s')$ iff $(s, s') \in R$. ■

Given a t-ordering and a sentence, the concepts of *plausibility* and *distrust* (of the sentence with respect to the t-ordering) can be defined. These concepts were originally defined within the context of epistemic doxastic logic as the *entrenchment* of a sentence and the *disbelief* in a sentence (Labuschagne et al., 2002). With the plausibility of a sentence with respect to a t-ordering is meant the degree of confidence with which the agent whose epistemic state is given by the t-ordering may regard the sentence to be true and adopt it as a basis for further action. A high degree of plausibility suggests that the sentence may be adopted with a high degree of confidence that it is true in the current state of the system. In essence, if the plausibility of a sentence with respect to a t-ordering has the highest possible level, the agent would consider the sentence to be irrefutable with respect to the t-ordering. A T-valid sentence is irrefutable with respect to every t-ordering.

The distrust in a sentence with respect to a t-ordering refers to the degree to which the agent may lack confidence that the sentence is satisfiable given the t-ordering. The higher the degree of distrust, the lower the degree of confidence in the satisfiability of a sentence, with the highest degree of distrust indicating a judgment of unsatisfiability. A T-unsatisfiable sentence is (as would be expected) unsatisfiable with respect to every t-ordering. A formalisation of these concepts is subsequently presented.

Definition 3.36 *Let $T = \langle S, t, l\rangle$ be a templated interpretation of L and let $\alpha \in L$. Then* *the **plausibility** of α with respect to t, is the function $pl : T_S \times L \to \{-1, 0, \ldots, n\}$ given by*

- $pl(t,\alpha) = \begin{cases} \text{the greatest } j \text{ such that } get_{\uparrow}(t,j) \subseteq Mod_T(\alpha) \\ -1 \text{ if there is no such } j \end{cases}$

The plausibility of a sentence with respect to a t-ordering is defined as the greatest level in the t-ordering such that the sentence is satisfied at every state (in the templated interpretation) lying at or below that level. From the definition it is clear that every T-valid sentence has a plausibility of n with respect to every t-ordering, since $Mod_T(\alpha) = S$ for every T-valid sentence α. A sentence that cannot be adopted with any degree of confidence has a plausibility of -1.

Definition 3.37 *Let $T = \langle S, t, l\rangle$ be a templated interpretation of L and let $\alpha \in L$. Then the **distrust** in α with respect to t, is the function $dt : T_S \times L \to \{0, \ldots, n\}$ given by*

- $dt(t,\alpha) = \begin{cases} \text{the least } j \text{ such that } get_{\uparrow}(t,j) \cap Mod_T(\alpha) \neq \varnothing \\ n \text{ if there is no such } j \end{cases}$

The distrust in a sentence with respect to a t-ordering is defined as the least level in the t-ordering such that the sentence is satisfied at some state (in the templated interpretation) lying at or below that level. From the definition, it is clear that the distrust in every T-unsatisfiable sentence is n with respect to every t-ordering, since $Mod_T(\alpha) = \varnothing$ for every T-unsatisfiable sentence α.

Proposition 3.36 *Let $T = \langle S, t, l\rangle$ be a templated interpretation of L and let $\alpha, \beta \in L$. Then the following hold:*

1. *If t is tautological, then $pl(t,\alpha) = -1$ for every $\alpha \in L$ that is not T-valid and $dt(t,\alpha) = 0$ for every $\alpha \in L$ that is not T-unsatisfiable*

2. *If t is strongly contradictory, then $pl(t, \alpha) = n - 1$ for every $\alpha \in L$ that is not T-valid and $dt(t, \alpha) = n$ for every $\alpha \in L$*

3. *If t is weakly contradictory, then $pl(t, \alpha) > -1$ for every $\alpha \in L$*

4. *If t is in normal form and $pl(t, \alpha) > -1$, then $dt(t, \alpha) = 0$*

5. *If $dt(t, \alpha) < dt(t, \beta)$ then $pl(t, \alpha) \geq pl(t, \beta)$*

Proof. 1. Let t be tautological. So $get_{\rightarrow}(t, 0) = S$. Suppose α is not T-valid, i.e. $Mod_T(\alpha) \neq S$. Then there can be no $j < n$ such that $get_{\uparrow}(t, j) \subseteq Mod_T(\alpha)$ and thus $pl(t, \alpha) = -1$. Suppose α is not T-unsatisfiable, i.e. $Mod_T(\alpha) \neq \varnothing$. Then, since $get_{\rightarrow}(t, 0) = S$ it follows that $get_{\uparrow}(t, 0) \cap Mod_T(\alpha) \neq \varnothing$, i.e. $dt(t, \alpha) = 0$.

2. Let t be strongly contradictory. So $get_{\uparrow}(t, n-1) = \varnothing$. But then $get_{\uparrow}(t, n-1) \subseteq Mod_T(\alpha)$ for every $\alpha \in L$ and there can be no greater $j < n$ such that $get_{\uparrow}(t, j) \subseteq Mod_T(\alpha)$ unless α is T-valid. Hence $pl(t, \alpha) = n - 1$ for every $\alpha \in L$ that is not T-valid. Since $get_{\uparrow}(t, n-1) = \varnothing$ it follows that $get_{\uparrow}(t, n) = S$ and thus, for every $\alpha \in L$ that is not T-unsatisfiable, the least j such that $get_{\uparrow}(t, j) \cap Mod_T(\alpha) \neq \varnothing$ is n. If α is unsatisfiable, then $dt(t, \alpha) = n$ by definition. Hence $dt(t, \alpha) = n$ for every $\alpha \in L$.

3. Let t be weakly contradictory. So $get_{\rightarrow}(t, 0) = \varnothing$. But then $get_{\rightarrow}(t, 0) \subseteq Mod_T(\alpha)$ for every $\alpha \in L$ and hence $pl(t, \alpha) > -1$.

4. Let t be in normal form. Suppose $pl(t, \alpha) > -1$. But then $get_{\uparrow}(t, 0) \neq \varnothing$ and so there must be some $j > 0$ such that $get_{\uparrow}(t, j) \subseteq Mod_T(\alpha)$. So $get_{\uparrow}(t, 0) \cap Mod_T(\alpha) \neq \varnothing$ and hence $dt(t, \alpha) = 0$.

5. Let $dt(t, \alpha) < dt(t, \beta)$. If $dt(t, \beta) = n$ then there is no j such that $get_{\uparrow}(t, j) \cap Mod_T(\beta) \neq \varnothing$. But then there can be no j such that $get_{\uparrow}(t, j) \subseteq Mod_T(\beta)$, i.e. $pl(t, \beta) = -1$. So $pl(t, \alpha) \geq pl(t, \beta)$. Let $dt(t, \alpha) = i$ and $dt(t, \beta) = j$. So i is the least level such that $get_{\uparrow}(t, i) \cap Mod_T(\alpha) \neq \varnothing$ and j is the least level such that $get_{\uparrow}(t, j) \cap Mod_T(\beta) \neq \varnothing$.

Since $i < j$ it follows that $get_{\uparrow}(t, i) \cap Mod_T(\beta) = \varnothing$. But then $get_{\uparrow}(t, i) \not\subseteq Mod_T(\beta)$, i.e. $pl(t, \beta) = -1$. So $pl(t, \alpha) \geq pl(t, \beta)$. ■

When restricted to t-orderings in normal form, a sentence which is plausible with respect to a t-ordering, i.e. a sentence that the agent feels able to adopt with some degree of confidence as a basis for action, has the least possible distrust with respect to the t-ordering, as shown by case (4) of the previous proposition. However, for sentences that are not plausible, the level of distrust may vary with respect to the t-ordering.

Example 3.9 *Consider a transparent propositional language L generated by $Atom = \{P(a), P(b)\}$ and where the states of the system, represented by $S = \{11, 10, 01, 00\}$, correspond directly to the valuations of L, with 10 corresponding to the valuation making $P(a)$ true and $P(b)$ false, and so on. Let $T = \langle S, t, l \rangle$ be a templated interpretation of L where $t = \{(11, 1), (10, 0), (01, 2), (00, 4)\}$ is a t-ordering in normal form and visually depicted in figure 3-7.*

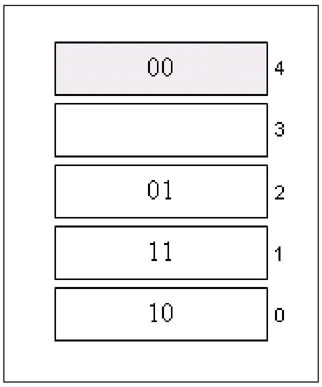

Figure 3-7: T-ordering t

The following sentences are not plausible with respect to t:

Plausibility	***Sentences***
-1	$\neg P(a) \vee P(b)$, $\neg P(a)$, $P(a) \leftrightarrow P(b)$, $P(b)$, $\neg P(a) \wedge \neg P(b)$, $\neg P(a) \wedge P(b)$, $P(a) \wedge P(b)$, $P(a) \wedge \neg P(a)$

However, the distrust in these sentences differs with respect to t:

Distrust	***Sentences***
1	$\neg P(a) \vee P(b)$, $P(a) \leftrightarrow P(b)$, $P(b)$, $P(a) \wedge P(b)$
2	$\neg P(a)$, $\neg P(a) \vee P(b)$
4	$\neg P(a) \wedge \neg P(b)$, $P(a) \wedge \neg P(a)$

The notion of the distrust in a sentence with respect to a t-ordering is used to define a restricted form of transitivity, called *Rational Transitivity* (Labuschagne et al., 2002), which is subsequently shown to be satisfied by every defeasible consequence relation induced by a templated interpretation.

Lemma 3.8 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha \in L$. Then $Min_t(\alpha) = \{s \in Mod_T(\alpha) \mid t(s) = dt(t, \alpha)\}$.*

Proof. See proof in appendix A, section A.4. ■

Lemma 3.9 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha \in L$ and $s \in S$. If $s \in Mod_T(\alpha \wedge \beta)$ and $s \in Min_t(\alpha)$, then $s \in Min_t(\alpha \wedge \beta)$.*

Proof. See proof in appendix A, section A.4. ■

Proposition 3.37 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha, \beta, \gamma \in L$. Then $\mid\!\sim_T$ satisfies the following property:*

- *if $\alpha \mid\!\sim_T \beta$ and $\beta \mid\!\sim_T \gamma$, and $dt(t, \alpha) \not> dt(t, \beta)$, then $\alpha \mid\!\sim_T \gamma$ (Rational Transitivity)*

Proof. Let $\alpha \mid\!\sim_T \beta$ and $\beta \mid\!\sim_T \gamma$. So $Min_t(\alpha) \subseteq Mod_T(\beta)$ and $Min_t(\beta) \subseteq Mod_T(\gamma)$. Let $dt(t,\alpha) \not> dt(t,\beta)$. Choose any $s \in Min_t(\alpha)$. By lemma 3.8, $s \in Mod_T(\alpha)$ and $t(s) = dt(t,\alpha)$. But $dt(t,\alpha) \not> dt(t,\beta)$. So either $t(s) < dt(t,\beta)$ or $t(s) = dt(t,\beta)$. But $s \in Mod_T(\beta)$ and thus $t(s) \not< dt(t,\beta)$. So $t(s) = dt(t,\beta)$. But then, by lemma 3.8, $s \in Min_t(\beta)$. But $Min_t(\beta) \subseteq Mod_T(\gamma)$ and thus $s \in Mod_T(\gamma)$. Since s was chosen arbitrarily it follows that $Min_t(\alpha) \subseteq Mod_T(\gamma)$. But then $\alpha \mid\!\sim_T \gamma$. ■

Other restricted forms of transitivity were studied by Freund, Lehmann, and Morris (1991). They showed that for preferential consequence relations, *Rational Monotonicity* is equivalent to several restricted forms of transitivity, in particular, *Weak Transitivity*, which says that transitivity for α, β, and γ holds provided α is compatible with β, i.e. provided β does not defeasibly entail α. A restricted form of contraposition, called *Weak Contraposition*, was also defined which says that contraposition between α and β, given γ as background information, holds provided that γ does not defeasibly entail β. *Weak Contraposition* was shown to be strictly weaker than *Rational Monotonicity*, again for preferential consequence relations. However, for (finite) preferential interpretations, which are ranked interpretations in which the ordering on the set of states is not a strict partial order but a strict modular partial order (or a total preorder), *Weak Contraposition* is equivalent to *Rational Monotonicity* when the labelling function is injective. The defeasible consequence relations induced by templated interpretations (which have been shown to be rational) satisfy both *Weak Transitivity* and *Weak Contraposition*.

Proposition 3.38 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha, \beta, \gamma \in L$. Then $\mid\!\sim_T$ satisfies the following properties:*

1. *if $\alpha \mid\!\sim_T \beta$ and $\beta \mid\!\sim_T \gamma$, and $\beta \not\mid\!\sim_T \neg\alpha$, then $\alpha \mid\!\sim_T \gamma$ (Weak Transitivity)*

2. *if $\alpha \wedge \gamma \mid\!\sim_T \beta$ and $\gamma \not\mid\!\sim_T \beta$, then $\neg\beta \wedge \gamma \mid\!\sim_T \neg\alpha$ (Weak Contraposition)*

Proof. 1. Let $\alpha \mid\!\sim_T \beta$ and $\beta \mid\!\sim_T \gamma$. So $Min_t(\alpha) \subseteq Mod_T(\beta)$ and $Min_t(\beta) \subseteq Mod_T(\gamma)$. Let $\beta \not\mid\!\sim_T \neg\alpha$. So $Min_t(\beta) \not\subseteq Mod_T(\neg\alpha)$. Choose any $s \in Min_t(\alpha)$. By lemma 3.8, $s \in Mod_T(\alpha)$ and $t(s) = dt(t, \alpha)$. Since $s \in Mod_T(\beta)$ it follows that either $dt(t, \beta) = t(s)$ or $dt(t, \beta) < t(s)$. If $dt(t, \beta) < t(s)$ then, by lemma 3.8, there must be some $s' \in Min_t(\beta)$ such that $dt(t, \beta) = t(s')$. So $s' \in Min_t(\beta)$ and $s' \notin Mod_T(\neg\alpha)$. But then $s' \in Mod_T(\alpha)$ from which it follows that $t(s') \geq t(s)$. Contradiction. So $dt(t, \beta) = t(s)$. But then, by lemma 3.8, $s \in Min_t(\beta)$. But $Min_t(\beta) \subseteq Mod_T(\gamma)$ and thus $s \in Mod_T(\gamma)$. Since s was chosen arbitrarily it follows that $Min_t(\alpha) \subseteq Mod_T(\gamma)$. But then $\alpha \mid\!\sim_T \gamma$.

2. Let $\alpha \wedge \gamma \mid\!\sim_T \beta$ and $\gamma \not\mid\!\sim_T \beta$. So $Min_t(\alpha \wedge \gamma) \subseteq Mod_T(\beta)$ and $Min_t(\gamma) \not\subseteq Mod_T(\beta)$. Choose any $s \in Min_t(\neg\beta \wedge \gamma)$. By lemma 3.8, $s \in Mod_T(\neg\beta \wedge \gamma)$ and $t(s) = dt(t, \neg\beta \wedge \gamma)$. So $s \in Mod_T(\neg\beta)$ and $s \in Mod_T(\gamma)$. But $Min_t(\gamma) \not\subseteq Mod_T(\beta)$. So $dt(t, \gamma) \geq dt(t, \neg\beta \wedge \gamma)$ otherwise there may be some $s' \in Min_t(\gamma)$ such that $s' \in Mod_T(\beta)$. Hence $t(s) \leq dt(t, \gamma)$. Suppose $s \notin Mod_T(\neg\alpha)$. Then $s \in Mod_T(\alpha)$. But $s \in Mod_T(\gamma)$ and thus $s \in Mod_T(\alpha \wedge \gamma)$. However, since $t(s) \leq dt(t, \gamma)$, it follows that $s \in Min_t(\gamma)$. But then, by lemma 3.9, $s \in Min_t(\alpha \wedge \gamma)$. But $Min_t(\alpha \wedge \gamma) \subseteq Mod_T(\beta)$ and thus $s \in Mod_T(\beta)$. Contradiction. So $s \in Mod_T(\neg\alpha)$. Since s was chosen arbitrarily it follows that $Min_t(\neg\beta \wedge \gamma) \subseteq Mod_T(\neg\alpha)$. But then $\neg\beta \wedge \gamma \mid\!\sim_T \neg\alpha$. ■

These results illustrate the role that the notions of plausibility and, in particular, distrust play in reasoning with t-orderings. As a practical example to reasoning with t-orderings, the lottery paradox is consider in the next section.

3.7 The lottery paradox

The *lottery paradox* is a famous problem (Kyburg, 1961) which arises when the conjunction of a number of likely propositions becomes very unlikely or even impossible. Poole (1989, 1991) has shown that the lottery paradox arises naturally in default reasoning systems, presenting difficulties for most.

The idea behind the lottery paradox is that there are a number of tickets, say m, and only one ticket that wins. The chance of any specific ticket, say ticket 117, winning is so small that $\neg Win(ticket_{117})$ is a likely proposition. The same applies for every other ticket. So it is possible to (defeasibly) infer that the conjunction $\neg Win(ticket_1) \wedge \neg Win(ticket_2) \wedge \ldots \wedge \neg Win(ticket_m)$ is true. But one ticket must win, in other words, the disjunction $Win(ticket_1) \vee Win(ticket_2) \vee \ldots \vee Win(ticket_m)$ must be true. The paradox is clear, it is possible to infer both the conjunction of a number of (likely) propositions and the negation thereof.

T-orderings offer a solution to the lottery paradox. In the example mentioned above, the definite information of one ticket winning, may be represented by assigning all states in which no ticket wins or more than one ticket wins to the top level of the t-ordering, i.e. by definitely excluding those states. The indefinite information of it being likely for a ticket not to win (but not most likely because some ticket has to win), may be represented by accommodating those states in which one ticket wins above the bottom level of the t-ordering, i.e. by tentatively excluding those states. How far above the bottom level is a matter of choice; a larger number of tickets warrants stronger exclusion, a smaller number of tickets weaker exclusion. For illustration purposes, assume that the states are assigned to level 1.

It is immediately apparent that the resulting t-ordering is weakly contradictory. So it is not unexpected that both $\neg Win(ticket_i)$ and $Win(ticket_i)$ may be inferred defeasibly. However, both are inferred with the lowest degree of confidence, both having a plausibility of 0 with respect to the t-ordering. Similarly, the conjunction $\neg Win(ticket_1) \wedge \neg Win(ticket_2) \wedge \ldots \wedge \neg Win(ticket_m)$ may only be inferred tentatively, as its plausibility is 0 too. However, the plausibility of the disjunction $Win(ticket_1) \vee Win(ticket_2) \vee \ldots \vee Win(ticket_m)$ with respect to the t-ordering is $n-1$, where n is the cardinality of the set of states, which implies that $Win(ticket_1) \vee Win(ticket_2) \vee \ldots \vee Win(ticket_m)$ may be inferred with the highest degree of confidence. It is interesting to note that the distrust in $\neg Win(ticket_1) \wedge \neg Win(ticket_2) \wedge \ldots \wedge \neg Win(ticket_m)$ with respect to the t-ordering is n (as would be expected) but that the distrust in $Win(ticket_1) \vee Win(ticket_2) \vee \ldots \vee Win(ticket_m)$ is 1, and not the lowest possible level of distrust. In essence, using t-orderings, it is possible to infer confidently that some ticket will win but only tentatively that no ticket will win. This resolves the paradox.

The lottery paradox provides an example of where a default rule cannot be represented by a t-ordering in normal form. It suggests a distinction between two different categories of default rules, those that can be represented by t-orderings in normal form and those that cannot. The latter leads to the kind of paradox encountered in the lottery example. In the context of control room agents, this is undesirable and hence the focus will be on default rules that can be represented by t-orderings in normal form.

Chapter 4

Knowledge and belief

IS THERE ANY KNOWLEDGE in the world which is so certain
that no reasonable man could doubt it?
The Problems of Philosophy (Bertrand Russel)

4.1 Kripke semantics

Hintikka (1962) originally formalised the informational attitudes of knowledge and belief using a modal logic (Hughes and Cresswell, 1968, 1996; Fitting, 1993) and a possible worlds semantics. This branch of modal logic is called *epistemic logic* after the Greek word 'episteme', meaning 'knowledge' (Fagin et al., 1995). The basic idea is that besides the actual world, there are a number of possible worlds and some of these worlds may be indistinguishable to the agent. Or, in other words, besides the actual state of the system, there are a number of states that, from the perspectives of the agent, are indistinguishable and equally good candidates for being the actual state. An agent is said to know a fact if the fact is true in all the states the agent considers possible candidates for being the actual state.

A modal language $\mathcal{L}$ for knowledge is constructed from the transparent propositional language L by adding a modal operator $\Box$ to the connectives of L. The intention is that $\Box$ will be read 'The agent knows that'. Every sentence of L is a sentence of $\mathcal{L}$. If α is a sentence of $\mathcal{L}$ then so too is $\Box\alpha$. $\mathcal{L}$ is the smallest set generated from *Atom* using the connectives as described.

Definition 4.1 *A **Kripke model** of $\mathcal{L}$ is a triple $\mathcal{M} = \langle S, R, l \rangle$ such that*

- *S is a non-empty set of states,*
- *$l : S \to U_T$ is a labelling function, and*
- *R is a binary relation on S (called an **accessibility** relation).*

If $(s, s') \in R$ then state $s' \in S$ is said to be *accessible* from state $s \in S$. The set of all states accessible from a state $s \in S$ is defined as $R(s) = \{s' \in S \mid (s, s') \in R\}$. A *frame* $\langle S, R \rangle$ is the class of all Kripke models sharing the same set S of states and the same accessibility relation R. A Kripke model $\mathcal{M} = \langle S, R, l \rangle$ is said to be based on the frame $\langle S, R \rangle$. Given a Kripke model $\mathcal{M} = \langle S, R, l \rangle$, the notion of a state $s \in S$ in $\mathcal{M}$ satisfying a sentence α of $\mathcal{L}$ may be extended from the possible worlds semantics of L to include the following case: $\mathcal{M}, s \Vdash \alpha$ iff

- $\alpha = \Box\beta$ and $\mathcal{M}, s' \Vdash \beta$ for every state $s' \in R(s)$.

A sentence $\Box\alpha \in \mathcal{L}$ is satisfied at a state $s \in S$ in $\mathcal{M}$ iff α is satisfied at every state that is accessible from s. The notion of a (local) model is defined in exactly the same way as before. A Kripke model $\mathcal{M} = \langle S, R, l \rangle$ is a (global) model of a sentence α of $\mathcal{L}$ iff α is $\mathcal{M}$-valid. A frame $\langle S, R \rangle$ is a *frame-model* of a sentence α of $\mathcal{L}$ iff every Kripke model based on $\langle S, R \rangle$ is a model of α.

To capture the intuition behind an agent knowing a fact if the fact is true in all the states the agent considers possible candidates for being the actual state, the accessibility relation on the set of states is taken to be an equivalence relation so that the set of candidate states constitutes an equivalence class. The class of Kripke models such that R is an equivalence relation on S can be characterised as the class of models of a set of sentences, which is called $KT45$. The set $KT45$ of sentences comprises all instances of the following schemas:

- $K : \Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$
- $T : \Box\alpha \rightarrow \alpha$
- $4 : \Box\alpha \rightarrow \Box\Box\alpha$
- $5 : \neg\Box\alpha \rightarrow \Box\neg\Box\alpha$

Schema K, which may equivalently be written as $(\Box(\alpha \rightarrow \beta) \wedge \Box\alpha) \rightarrow \Box\beta$, says that if an agent knows $\alpha \rightarrow \beta$ and the agent also knows α, then the agent will know β. It is named in honour of Saul Kripke for his contribution to the development of the possible worlds semantics for modal logic (Kripke, 1963). Schema K suggests that agents are ideal reasoners in the sense of knowing all the consequences of their knowledge. Every sentence which is an instance of schema K has as its models all Kripke models.

Schema T is sometimes referred to as the knowledge schema because it provides a distinction between knowledge and belief under the traditional philosophical assumption of knowledge as 'true justified belief'. Schema T says that if an agent knows α, then α is true. This schema characterises frames whose accessibility relations are reflexive, in other words, every sentence which is an instance of schema T has as its frame-models those frames whose accessibility relations are reflexive and, conversely, every frame whose accessibility relation is reflexive is a frame-model for every sentence which is an instance of schema T.

Schema 4 says that if an agent knows α then the agent knows that it knows α. Intuitively, it means that agents are capable of positive introspection. The positive introspection may be iterated without limit, suggesting that agents are ideal reasoners in the sense of having unlimited resources[1]. Schema 4 characterises frames whose accessibility relations are transitive.

Negative introspection is captured by schema 5, which says that if an agent does not know α then it knows that it does not know α. Again, negative introspection may be iterated without limit. This schema, when taken in conjunction with schemas T and 4, characterises frames whose accessibility relations are not only reflexive and transitive but also symmetric, and thus equivalence relations. Note that on its own, schema 5 characterises frames whose accessibility relations are euclidean, where by an *euclidean* relation R on X it is understood that for all elements $x, y, z \in X$, if $(x, y) \in R$ and $(x, z) \in R$, then $(y, z) \in R$. Frames whose accessibility relations are symmetric can be characterised as the class of models of a set of sentences that are the instances of a schema B, also called the Brouwerian schema, which is stated as

- $B : \alpha \rightarrow \Box\neg\Box\neg\alpha$.

Positive and negative introspection taken together lead to the *Plato principle*, stated as $\neg\alpha \rightarrow \Box\neg\Box\alpha$, which says that the mere falsity of a sentence is enough for the agent to know that it does not know the sentence. The Plato principle is often regarded as an idealistic and unrealistic feature (Girle, 1998). However, despite philosophical debate, it has become customary to regard schemas K, T, 4, and 5 as an appropriate set of

[1] One approach that considers resource boundedness is that of Konolige (1986), in which accessibility relations are replaced by pairs (Γ, Δ) where Γ is an initial set of beliefs and Δ is a deduction algorithm which may be rendered incomplete by constraints on time and memory.

properties of knowledge. Collectively, schemas K, T, 4, and 5 are referred to as the $S5$ properties. These properties have also proved useful in domains such as distributed systems (Halpern, 1987) and economics and game theory (Aumann, 1976; Parikh, 1985; van Benthem, 2001).

The modal language $\mathcal{L}$ defined earlier may be taken as a language of belief when $\Box$ is read as 'The agent believes that'. Assuming that an agent may hold a belief that is false, schema T would be excluded as a property of belief. A popular approach is to regard schemas K, D, 4, and 5 as the properties of belief, collectively referred to as the $KD45$ properties. Schema D is

- $D : \Box\alpha \to \neg\Box\neg\alpha$.

Schema D, which may equivalently be written as $\neg\Box(\alpha \wedge \neg\alpha)$, says that if an agent believes α, then the agent does not believe the negation of α. Intuitively, it means that the agent's beliefs are satisfiable. Schema D characterises frames whose accessibility relations are serial, where by a *serial* relation R on X it is understood that for every $x \in X$, there is some $y \in X$ such that $(x, y) \in R$.

The Kripke semantics of epistemic logic allows one to model either the knowledge or the beliefs of an agent. To model both knowledge and belief in the Kripke semantics, one would have to introduce two modal operators and an accessibility relation for each of the operators. In the absence of further provisions, the accessibility relations would be independent and thus counter intuitive to the notion that knowledge and belief are closely related.

4.2 Information-theoretic semantics

As a generalisation of the Kripke semantics of epistemic logic, Labuschagne and Ferguson (2002) define an information-theoretic semantics that assigns a t-ordering to each state of the system. The information-theoretic semantics is recalled in this section in the context of diagrammable systems with t-orderings defined as instances of a templated ordering ADT. As mentioned earlier, in the context of diagrammable systems, the states of the system correspond to a subset of the valuations of some finitely generated propositional language. The modal language to be constructed from such a propositional language will consist of a number of modal operators for belief, one for every element in the template B associated with the set S of states.

A modal language $\mathcal{L}'$ for belief is constructed from the transparent propositional language L by adding modal operators $[0], [1], \ldots, [n]$ to the connectives of L, where $n = card(S)$. The intention is that $[i]$ will be read 'The agent believes to degree i that'. Every sentence of L is a sentence of $\mathcal{L}'$. If α is a sentence of $\mathcal{L}'$ then so too is $[i]\,\alpha$. $\mathcal{L}'$ is the smallest set generated from $Atom$ using the connectives as described.

The intuition behind the different modal operators is that $[0]$ reflects the agent's most tentatively held defeasible beliefs. Modal operator $[n]$ simply reflects tautologies. Each modal operator $[i]$ in between reflects the agent's defeasible beliefs held to an increasing level of strength with $[n-1]$ reflecting the agent's most strongly held beliefs, or its knowledge. This intuition will be motivated information-theoretically in section 4.5 when the notion of epistemic entrenchment is examined.

An information-theoretic semantics for $\mathcal{L}'$ is defined by replacing the accessibility relation of the Kripke semantics on a (finite) set of states S with an accessibility function from S to T_S (Labuschagne and Ferguson, 2002).

Definition 4.2 *A* ***templated interpretation*** *of $\mathcal{L}'$ is a triple $\mathcal{T} = \langle S, F, l \rangle$ such that*

- *S is a finite non-empty set of states with $card(S) = n$,*
- *$l : S \to U_T$ is a labelling function, and*
- *$F : S \to T_S$ is an accessibility function.*

The accessibility function of a templated interpretation of $\mathcal{L}'$ associates with each state of S a t-ordering of T_S. The implication of using an accessibility function instead of an accessibility relation is that the agent will have available at each state of the system, not merely definite information, but indefinite information as well, because, as has been shown in the previous chapter, t-orderings allow for the combination of both definite and indefinite information in the same semantic structure. It is this capability to have indefinite information available at each state of the system that is the essential difference between accessibility functions and accessibility relations. The accessibility function in effect equips the agent, at every state, with a context-dependent default rule, in addition to recording the definite information available to the agent.

When restricting to definite information (and thus to definite t-orderings), the division of the finite set of states S into a set of included states C and a complementary set of excluded states $\overline{C}$ is directly equivalent to the Kripke semantics. Given an accessibility relation R and a state $s \in S$, the set of states considered possible candidates for being the actual state, may be formed by taking $C = R(s)$, where $R(s) = \{s' \in S \mid (s, s') \in R\}$, while the set of excluded states may be taken as the complement $\overline{C} = S - C$. In effect, an accessibility relation may be seen as associating with each state an ordered pair $(C, \overline{C})$ which, as shown earlier, may be represented by a definite t-ordering $t \in T_S$ where $bottom(t) = C$ and $top(t) = \overline{C}$. Conversely, given an accessibility function F and a state $s \in S$, the set of states considered possible candidates for being the actual state may be formed by taking $C = get_{\uparrow}(F(s), 0)$ while the set of excluded states may be taken as the complement $\overline{C} = get_{\rightarrow}(F(s), n)$. The set of states accessible from s is thus, in the case of a definite t-ordering, $get_{\uparrow}(F(s), 0)$.

If the restriction to definite information is removed, then the set of states considered possible candidates for being the actual state must be formed by taking $C = get_{\uparrow}(F(s), n{-}1)$ so that the set of excluded states may still be taken as $\overline{C} = get_{\rightarrow}(F(s), n)$, given an accessibility function F and a state $s \in S$. The set of states accessible from s will now become $get_{\uparrow}(F(s), n-1)$. Having established the correlation between accessibility relations and accessibility functions, it is now possible to define a refined notion of accessibility given an accessibility function.

Definition 4.3 *Let $\mathcal{T} = \langle S, F, l \rangle$ be a templated interpretation of $\mathcal{L}'$ and let $s, s' \in S$. Then state s' is **accessible** from state s iff $s' \in get_{\uparrow}(F(s), n-1)$. A state s' is **accessible to degree i** from state s iff $s' \in get_{\uparrow}(F(s), i)$ for $i < n$.*

The refined notion of accessibility given an accessibility function $F : S \rightarrow T_S$ suggests that it is possible to recover not merely one accessibility relation R on S from an accessibility function, but n, denoted by R_i where $i = 0, 1, \ldots, n-1$ and $n = card(S)$. This is one sense in which the information-theoretic semantics may be seen as a generalisation of the Kripke semantics. Note that in the process of recovering an accessibility relation from an accessibility function, some of the indefinite information available to the agent at each state of the system may be lost, although recovering all n relations simultaneously preserves all information. Formally, therefore, the information-theoretic semantics can be simulated by a multimodal version of epistemic logic in which one accessibility relation is designated the knowledge relation and the remainder are concerned with beliefs of varying degrees. The facilitation of visualisation by the t-ordering ADT is sacrificed by such a simulation, however.

Proposition 4.1 *Let S be a non-empty set of states with $card(S) = n$. Then every accessibility relation R on S induces an accessibility function $F : S \rightarrow T_S$ such that $R(s) = get_{\uparrow}(F(s), 0)$ for every $s \in S$ and conversely, from every accessibility function $F : S \rightarrow T_S$ an accessibility relation R_i on S may be recovered where $i < n$ such that $get_{\uparrow}(F(s), i) = R_i(s)$ for every $s \in S$.*

Proof. Let R be an accessibility relation on S. An accessibility function $F : S \rightarrow T_S$ is constructed by taking $F = \{(s,t) \mid \mathit{bottom}(t) = R(s) \text{ and } \mathit{top}(t) = S - R(s)\}$. It is trivial to shown that for every state $s \in S$, $R(s) = \mathit{get}_{\uparrow}(F(s), 0)$. Conversely, let $F : S \rightarrow T_S$ be an accessibility function. An accessibility relation R_i on S is constructed by taking $R_i = \{(s,s') \mid \ s' \in \mathit{get}_{\uparrow}(F(s), i)\}$ for $i < n$. Again, it is trivial to shown that for every state $s \in S$, $\mathit{get}_{\uparrow}(F(s), i) = R_i(s)$. ■

A *templated frame* $\langle S, F\rangle$ is the class of all templated interpretations sharing the same set S of states and the same accessibility function F. A templated interpretation $\mathcal{T} = \langle S, F, l\rangle$ is said to be based on the templated frame $\langle S, F\rangle$. Given a templated interpretation $\mathcal{T} = \langle S, F, l\rangle$, the notion of a state $s \in S$ in $\mathcal{T}$ satisfying a sentence α of $\mathcal{L}'$ may be extended from the possible worlds semantics of L to include the following case: $\mathcal{T}, s \Vdash \alpha$ iff

- $\alpha = [i]\beta$ and $\mathcal{T}, s' \Vdash \beta$ for every state $s' \in \mathit{get}_{\uparrow}(F(s), i)$.

A sentence $[i]\alpha \in \mathcal{L}'$ is satisfied at a state $s \in S$ in $\mathcal{T}$ iff α is satisfied at every state that is accessible to degree i from s. At one extreme, $[0]\alpha$ is satisfied at s iff α is satisfied at every state occurring in the bottom level of t-ordering $F(s)$, in other words, at those states considered most normal in t-ordering $F(s)$. At the other extreme, $[n]\alpha$ is satisfied at s iff α is $\mathcal{T}$-valid. More importantly, $[n-1]\alpha$ is satisfied at s iff α is satisfied at every state which is not definitely excluded in t-ordering $F(s)$. The notions of a (local) model, a model, and a frame-model of a sentence α of $\mathcal{L}'$ are defined in exactly the same way as for the Kripke semantics.

Proposition 4.2 *Let $\mathcal{T} = \langle S, F, l\rangle$ be a templated interpretation of $\mathcal{L}'$ and let $\mathcal{M} = \langle S, R_{n-1}, l\rangle$ be the Kripke model of $\mathcal{L}$ where R_{n-1} is the accessibility relation recovered from F. Let $s \in S$ and $\alpha \in L$. Let $\Box$ be taken as the knowledge operator of $\mathcal{L}$. Then $\mathcal{T}, s \Vdash [n-1]\alpha$ iff $\mathcal{M}, s \Vdash \Box\alpha$.*

Proof. Suppose s satisfies $[n-1]\alpha$ in $\mathcal{T}$. So s' satisfies α in $\mathcal{T}$ for every $s' \in get_{\uparrow}(F(s), n-1)$. But if $s' \in get_{\uparrow}(F(s), n-1)$, then $s' \in R_{n-1}(s)$ by proposition 4.1. Thus, since $\alpha \in L$, it follows that s' satisfies α in $\mathcal{M}$ for every $s' \in R_{n-1}(s)$. Hence, s satisfies $\Box\alpha$ in $\mathcal{M}$.

Conversely, suppose s satisfies $\Box\alpha$ in $\mathcal{M}$. So s' satisfies α in $\mathcal{M}$ for every $s' \in R_{n-1}(s)$. But if $s' \in R_{n-1}(s)$, then $s' \in bottom(F(s))$ by proposition 4.1. But $F(s)$ is definite and so $s' \in get_{\uparrow}(F(s), n-1)$. Thus, since $\alpha \in L$, it follows that s' satisfies α in $\mathcal{T}$ for every $s' \in get_{\uparrow}(F(s), n-1)$. Hence, s satisfies $[n-1]\alpha$ in $\mathcal{T}$. ∎

As shown in the previous result, modal operator $[n-1]$ of $\mathcal{L}'$ corresponds to modal operator $\Box$ of $\mathcal{L}$, both representing the agent's knowledge. The other modal operators of $\mathcal{L}'$ have no counterpart in $\mathcal{L}$. However, if modal operator $\Box$ of $\mathcal{L}$ is taken as the belief operator, then modal operator $[0]$ of $\mathcal{L}'$, representing the agent's most tentative defeasible beliefs, corresponds to $\Box$ given an appropriate accessibility relation. Again, the other modal operators of $\mathcal{L}'$ would now have no counterpart in $\mathcal{L}$. This is another sense in which the information-theoretic semantics may be seen as a generalisation of the Kripke semantics.

Proposition 4.3 *Let $\mathcal{T} = \langle S, F, l\rangle$ be a templated interpretation of $\mathcal{L}'$ and let $\mathcal{M} = \langle S, R_0, l\rangle$ be the Kripke model of $\mathcal{L}$ where R_0 is the accessibility relation recovered from F. Let $s \in S$ and $\alpha \in L$. Let $\Box$ be taken as the belief operator of $\mathcal{L}$. Then $\mathcal{T}, s \Vdash [0]\alpha$ iff $\mathcal{M}, s \Vdash \Box\alpha$.*

Proof. Suppose s satisfies $[0]\alpha$ in $\mathcal{T}$. So s' satisfies α in $\mathcal{T}$ for every $s' \in get_{\uparrow}(F(s), 0)$. But if $s' \in get_{\uparrow}(F(s), 0)$, then $s' \in R_0(s)$ by proposition 4.1. Thus, since $\alpha \in L$, it follows that s' satisfies α in $\mathcal{M}$ for every $s' \in R_0(s)$. Hence, s satisfies $\Box\alpha$ in $\mathcal{M}$.

Conversely, suppose s satisfies $\Box\alpha$ in $\mathcal{M}$. So s' satisfies α in $\mathcal{M}$ for every $s' \in R_0(s)$. But if $s' \in R_0(s)$, then $s' \in bottom(F(s))$ by proposition 4.1. But if $s' \in bottom(F(s))$

then $s' \in get_{\uparrow}(F(s), 0)$. Thus, since $\alpha \in L$, it follows that s' satisfies α in $\mathcal{T}$ for every $s' \in get_{\uparrow}(F(s), 0)$. Hence, s satisfies $[0]\alpha$ in $\mathcal{T}$. ■

However, the full extent of the generalisation of the Kripke semantics only becomes apparent when the properties of knowledge and belief supported by the information-theoretic semantics are revealed.

4.3 Properties of knowledge and belief

As mentioned earlier, it is customary to regard the schemas K, T, 4, and 5 as the properties of knowledge. Each of these schemas has a faithful counterpart (Labuschagne and Ferguson, 2002) in the information-theoretic setting of epistemic logic by the correspondence of modal operator $[n-1]$ of $\mathcal{L}'$ to modal operator $\Box$ of $\mathcal{L}$ (assuming that $\Box$ is read as 'The agent knows that').

- $K_{n-1} : [n-1](\alpha \to \beta) \to ([n-1]\alpha \to [n-1]\beta)$
- $T_{n-1} : [n-1]\alpha \to \alpha$
- $4_{n-1} : [n-1]\alpha \to [n-1][n-1]\alpha$
- $5_{n-1} : \neg[n-1]\alpha \to [n-1]\neg[n-1]\alpha$

Note that the operator $[n-1]$, like negation, governs the shortest well-formed sentence following it, so that $[n-1]\alpha \to \alpha$ should be read as $([n-1]\alpha) \to \alpha$, and not as $[n-1](\alpha \to \alpha)$.

One would expect every sentence which is an instance of schema K_{n-1} to have as its models every templated interpretation of $\mathcal{L}'$. Indeed, this will be shown to be the case. However, in the information-theoretic setting, there exists a more general counterpart to schema K with the property that if sentence $\gamma \in \mathcal{L}'$ is an instance of the schema, then every templated interpretation of $\mathcal{L}'$ is a model of γ.

- $K_i : [i](\alpha \to \beta) \to ([i]\alpha \to [i]\beta)$

The generalisation is that the value of i need not be fixed at $n-1$, but instead can be fixed at any value below n. Schema K_i says that if an agent defeasibly believes $\alpha \to \beta$ to degree i and the agent also believes α to degree i, then the agent will believe β to the same degree. When i is taken to be $n-1$, schema K_i is simply schema K_{n-1}.

Proposition 4.4 *If $\gamma \in \mathcal{L}'$ is of the form $[i](\alpha \to \beta) \to ([i]\alpha \to [i]\beta)$, then every templated interpretation $\mathcal{T} = \langle S, F, l \rangle$ of $\mathcal{L}'$ is a model of γ.*

Proof. Let γ be any sentence of $\mathcal{L}'$ of the form $[i](\alpha \to \beta) \to ([i]\alpha \to [i]\beta)$. Choose any templated interpretation $\mathcal{T} = \langle S, F, l \rangle$ of $\mathcal{L}'$ and any state $s \in S$. Assume that γ fails to be satisfied at s. So $[i](\alpha \to \beta)$ is satisfied at s but $([i]\alpha \to [i]\beta)$ is not satisfied at s, from which it follows that $[i]\alpha$ is satisfied at s but $[i]\beta$ not. Since $[i]\beta$ is not satisfied at s, there must be some state $s' \in get_{\uparrow}(F(s), i)$ such that β is not satisfied at s'. However, since $[i](\alpha \to \beta)$ is satisfied at s and $[i]\alpha$ is satisfied at s, it must be the case that both $\alpha \to \beta$ and α are satisfied at every state in $get_{\uparrow}(F(s), i)$, including state s'. But then β must be satisfied at s'. Contradiction. Thus γ is satisfied at the arbitrarily chosen s and therefore at every state in $\mathcal{T}$. But $\mathcal{T}$ itself was arbitrarily chosen and thus every templated interpretation is a model of γ. ∎

If an accessibility relation R on a set S of states is reflexive, then for every state $s \in S$, $s \in R(s)$, i.e. s is accessible from itself. Frames whose accessibility relations are reflexive are, as mentioned earlier, characterised as the class of Kripke models of the set of sentences which are the instances of schema T. In the information-theoretic setting, one would expect schema T_{n-1} to similarly characterise templated frames whose accessibility functions ensure that every state $s \in S$ is accessible from itself.

Proposition 4.5 *Let $\langle S, F\rangle$ be a templated frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of every sentence of $\mathcal{L}'$ which is an instance of schema T_{n-1} iff for every state $s \in S$, s is accessible from itself.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of every sentence which is an instance of schema T_{n-1}. It must be shown that for every state $s \in S$, $s \in get_{\uparrow}(F(s), n-1)$. Choose any $s \in S$. Suppose $s \notin get_{\uparrow}(F(s), n-1)$. Then a templated interpretation can be constructed based on $\langle S, F\rangle$ that is not a model of all the instances of schema T_{n-1}, in particular the sentence $[n-1]P(a) \rightarrow P(a)$. Take $\mathcal{T} = \langle S, F, l\rangle$ such that $v_{l(s)}(P(a)) = 0$ and $v_{l(s')}(P(a)) = 1$ for every state $s' \in get_{\uparrow}(F(s), n-1)$. So $P(a)$ is not satisfied at s but is satisfied at every $s' \in get_{\uparrow}(F(s), n-1)$. But then $[n-1]P(a)$ is satisfied at s and since $P(a)$ is not satisfied at s it follows that $[n-1]P(a) \rightarrow P(a)$ is not satisfied at s. But then $\langle S, F\rangle$ cannot be a frame-model of $[n-1]P(a) \rightarrow P(a)$. Contradiction. Thus $s \in get_{\uparrow}(F(s), n-1)$ and since s was chosen arbitrarily, every state $s \in S$ is accessible from itself.

Conversely, suppose that for every state $s \in S$, s is accessible from itself. So $s \in get_{\uparrow}(F(s), n-1)$. Choose any sentence $\gamma \in \mathcal{L}'$ of the form $[n-1]\alpha \rightarrow \alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $\mathcal{T} = \langle S, F, l\rangle$ of $\mathcal{L}'$ and some state $s \in S$ such that s fails to satisfy γ in $\mathcal{T}$. But then it must be the case that $[n-1]\alpha$ is satisfied at s and α is not satisfied at s. But if $[n-1]\alpha$ is satisfied at s then α is satisfied at every $s' \in get_{\uparrow}(F(s), n-1)$. But $s \in get_{\uparrow}(F(s), n-1)$ and hence α must be satisfied at s. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema T_{n-1}. ■

In the Kripke semantics, schemas K, T, 4, and 5 collectively characterise frames whose accessibility relations are equivalence relations. An equivalence relation captures the intuition that besides the actual state of the system, there are a number of states that, to the agent, are indistinguishable from the actual state and that this set of equally good candidates (for being the actual state) constitutes an equivalence class. In the information-theoretic semantics, schemas K_{n-1}, T_{n-1}, 4_{n-1}, and 5_{n-1} collectively characterise templated frames whose accessibility functions ensure that every state is accessible from itself (and thus a candidate for being the actual state, inasmuch as it has not been definitely excluded), and that mutually accessible states have t-orderings with exactly the same definite content.

Proposition 4.6 *Let $\langle S, F\rangle$ be a templated frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_{n-1}T_{n-1}4_{n-1}5_{n-1}$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_{n-1}, T_{n-1}, 4_{n-1}, or 5_{n-1} iff for every $s, s' \in S$,*

1. *s is accessible from itself, and*
2. *if s' is accessible from s, then $Cont_D(F(s')) = Cont_D(F(s))$.*

Proof. See proof in appendix B, section B.1. ■

The properties of belief are generally taken to be represented by the schemas K, D, 4, and 5, as mentioned earlier. Again, each of these schemas has a faithful counterpart in the information-theoretic setting by the correspondence of modal operator $[0]$ of $\mathcal{L}'$ to modal operator $\Box$ of $\mathcal{L}$ (assuming this time that $\Box$ is read as 'The agent believes that').

- $K_0 : [0](\alpha \rightarrow \beta) \rightarrow ([0]\alpha \rightarrow [0]\beta)$
- $D_0 : [0]\alpha \rightarrow \neg[0]\neg\alpha$
- $4_0 : [0]\alpha \rightarrow [0][0]\alpha$
- $5_0 : \neg[0]\alpha \rightarrow [0]\neg[0]\alpha$

In the information-theoretic semantics, schemas K_0, D_0, 4_0, and 5_0 collectively characterise templated frames whose accessibility functions ensure that every state is accessible to degree 0 from some state and that states which are mutually accessible to degree 0 have t-orderings with exactly the same indefinite content.

Proposition 4.7 *Let $\langle S, F\rangle$ be a templated frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_0D_04_05_0$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_0, D_0, 4_0, or 5_0 iff*

1. *for every $s \in S$ there exists some $s' \in S$ such that s is accessible to degree 0 from s' and*

2. *for every $s, s' \in S$, if s' is accessible to degree 0 from s, then $Cont_0(F(s')) = Cont_0(F(s))$.*

Proof. See proof in appendix B, section B.1. ■

Thus far, the focus has been primarily on properties of knowledge and belief that have faithful counterparts to the schemas familiar from the Kripke semantics of epistemic logic. However, the information-theoretic semantics allows for a wide range of counterparts to these properties (Ferguson and Labuschagne, 2001).

Schema T may be imposed over any of the defeasible belief operators $[0], [1], \ldots, [n-1]$, instead of only over the defeasible belief (or knowledge) operator $[n-1]$, thereby placing a stronger constraint on the capabilities of the agent.

- $T_i : [i]\alpha \rightarrow \alpha$

Schema T_i says that if an agent defeasibly believes α to degree i, then α is true. The class of agents with schema T_i as a property of knowledge and belief would, if i is taken to be less than $n-1$, have to have default rules that are infallible to degree i whereas, if i is taken to be $n-1$, would have to have sensors that deliver infallible definite information. The latter is the more feasible of the properties.

Various counterparts to schema 4 are possible, each of which imposes different constraints on the positive introspective capabilities of the agent.

- $4_i : [i]\alpha \rightarrow [i][i]\alpha$
- $4_{iK} : [n-1]\alpha \rightarrow [i][n-1]\alpha$
- $4_{Ki} : [i]\alpha \rightarrow [n-1][i]\alpha$

Schema 4_i says that if an agent defeasibly believes α to degree i, then the agent defeasibly believes to degree i that it defeasibly believes α (again, to degree i). The class of agents with schema 4_i as a property of knowledge and belief would, if i is taken to be $n-1$, be capable of knowing that they know whereas, if i is taken to be less than $n-1$, would be capable of believing that they believe. However, this class of agents would not be capable of believing that they know and would not be capable of knowing that they believe.

Schema 4_{iK} says that if an agent knows α, then the agent defeasibly believes to degree i that it knows α. Schema 4_{0K} reflects the weakest form of positive introspection in which the agent is only weakly aware of its knowledge. The strongest form of positive introspection is, of course, schema 4_{n-1}. The class of agents with schema 4_{iK} as a property of knowledge and belief would be capable of believing (to degree i) that they know.

Schema 4_{Ki} says that if an agent defeasibly believes α to degree i, then the agent knows that it defeasibly believes α to degree i. The class of agents with schema 4_{Ki} as a property of knowledge and belief would be capable of knowing that they believe (to degree i). In an approach such as ours, where knowledge is seen as a certain kind of belief, schema 4_{Ki} is intuitively more appealing than the other properties of positive introspection.

Proposition 4.8 *Let* $\langle S, F\rangle$ *be a templated frame of* $\mathcal{L}'$*. Then* $\langle S, F\rangle$ *is a frame-model of every sentence of* $\mathcal{L}'$ *which is an instance of schema* 4_{Ki} *iff for every* $s, s' \in S$*, if* s' *is accessible to degree* i *from* s*, then* $Cont_i(F(s)) \subseteq Cont_i(F(s'))$ *and* $F(s')(x) \geq F(s)(x)$ *for every* $x \in get_{\uparrow}(F(s'), i)$*.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of every sentence which is an instance of schema 4_{Ki}. Suppose that there are $s, s' \in S$ such that $s' \in get_{\uparrow}(F(s), i)$ but either $Cont_i(F(s)) \nsubseteq Cont_i(F(s'))$ or $F(s')(x) < F(s)(x)$ for some $x \in get_{\uparrow}(F(s'), i)$. Then, in both cases, a templated interpretation based on $\langle S, F\rangle$ can be constructed that is not a model of all the instances of schema 4_{Ki}, in particular the sentence $[i]P(a) \rightarrow [n-1][i]P(a)$. Assume $Cont_i(F(s)) \nsubseteq Cont_i(F(s'))$. Take $T = \langle S, F, l\rangle$ such that $v_{l(s'')}(P(a)) = 1$ for every state $s'' \in get_{\uparrow}(F(s), i)$ including state s', and $v_{l(s''')}(P(a)) = 0$ for some $s''' \in get_{\uparrow}(F(s'), i)$. So $P(a)$ is satisfied at every state $s'' \in get_{\uparrow}(F(s), i)$ and hence $[i]P(a)$ is satisfied at state s. However, $P(a)$ is not satisfied at state $s''' \in get_{\uparrow}(F(s'), i)$ and thus $[i]P(a)$ is not satisfied at state s'. But $s' \in get_{\uparrow}(F(s), i)$ and hence $s' \in get_{\uparrow}(F(s), n-1)$. But then $[n-1][i]P(a)$ is not satisfied at state s. But then s fails to satisfy $[i]P(a) \rightarrow [n-1][i]P(a)$ at state s. So $\langle S, F\rangle$ cannot be a frame-model of $[i]P(a) \rightarrow [n-1][i]P(a)$. Contradiction. Hence, if s' is accessible from s, then $Cont_i(F(s)) \subseteq Cont_i(F(s'))$. Now assume $F(s')(x) < F(s)(x)$ for some $x \in get_{\uparrow}(F(s'), i)$. Take $T = \langle S, F, l\rangle$ such that $v_{l(s'')}(P(a)) = 1$ for every state $s'' \in get_{\uparrow}(F(s), i)$ including state s', and $v_{l(s''')}(P(a)) = 0$ for some $s''' \in get_{\uparrow}(F(s'), i)$. Suppose $F(s')(s''') < F(s)(s''')$. So $P(a)$ is satisfied at every state $s'' \in get_{\uparrow}(F(s), i)$ and hence $[i]P(a)$ is satisfied at state s. However, $P(a)$ is not satisfied at state $s''' \in get_{\uparrow}(F(s'), i)$ and thus $[i]P(a)$ is not satisfied at state s'. But $s' \in get_{\uparrow}(F(s), i)$ and hence $s' \in get_{\uparrow}(F(s), n-1)$. But then $[n-1][i]P(a)$ is not satisfied at state s. So s fails to satisfy $[i]P(a) \rightarrow [n-1][i]P(a)$ at state s. So $\langle S, F\rangle$ cannot be a frame-model of $[i]P(a) \rightarrow [n-1][i]P(a)$. Contradiction. Hence, if s' is accessible from s, then $F(s')(x) \geq F(s)(x)$ for every $x \in get_{\uparrow}(F(s'), i)$.

Conversely, suppose that for every $s, s' \in S$, if s' is accessible to degree i from s, then $Cont_i(F(s)) \subseteq Cont_i(F(s'))$ and $F(s')(x) \geq F(s)(x)$ for every $x \in get_{\uparrow}(F(s'), i)$. To show that $\langle S, F\rangle$ is a a frame-model of every sentence that is an instance of schema 4_{Ki} choose any sentence $\gamma \in \mathcal{L}'$ of the form $[i]\alpha \rightarrow [n-1][i]\alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $\mathcal{T} = \langle S, F, l\rangle$ of $\mathcal{L}'$ and some state $s \in S$ such that s fails to satisfy γ in $\mathcal{T}$. So it must be the case that $[i]\alpha$ is satisfied at s but $[n-1][i]\alpha$ is not. If $[i]\alpha$ is satisfied at s then α is satisfied at every state $s' \in get_{\uparrow}(F(s), i)$. If $[n-1][i]\alpha$ is not satisfied at s then there must be some state $s'' \in get_{\uparrow}(F(s), n-1)$ at which $[i]\alpha$ is not satisfied. Hence there must be some state $s''' \in get_{\uparrow}(F(s''), i)$ that fails to satisfy α. But if $s''' \in get_{\uparrow}(F(s''), i)$ then $s''' \in get_{\uparrow}(F(s''), n-1)$. But $get_{\uparrow}(F(s''), n-1) \subseteq get_{\uparrow}(F(s), n-1)$ with $F(s'')(x) \geq F(s)(x)$ for every $x \in get_{\uparrow}(F(s''), n-1)$. Since $s''' \in get_{\uparrow}(F(s''), i)$ it follows that $F(s'')(s''') \leq i$ and hence $s''' \in get_{\uparrow}(F(s), i)$. But α is satisfied at every state $s' \in get_{\uparrow}(F(s), i)$ and thus α must be satisfied at s''. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema 4_{Ki}.

■

As with the properties of positive introspection, various counterparts to schema 5 are possible, each of which imposes different constraints on the negative introspective capabilities of the agent.

- $5_i : \neg[i]\alpha \rightarrow [i]\neg[i]\alpha$
- $5_{iK} : \neg[n-1]\alpha \rightarrow [i]\neg[n-1]\alpha$
- $5_{Ki} : \neg[i]\alpha \rightarrow [n-1]\neg[i]\alpha$

Schema 5_i says that if an agent does not defeasibly believe α to degree i, then the agent defeasibly believes to degree i that it does not defeasibly believe α (again, to degree

i). A weaker form of negative introspection is suggested by schema 5_{iK} which says that if an agent does not know α, then the agent defeasibly believes to degree i that it does not know α. Schema 5_{Ki} says that if an agent does not defeasibly believe α to degree i, then the agent knows that it does not defeasibly believe α to degree i.

The class of agents with schema 5_i as a property of knowledge and belief would, if i is taken to be $n-1$, be capable of knowing that they do not know whereas, if i is taken to be less than $n-1$, would be capable of believing that they do not believe. The class of agents with schema 5_{iK} as a property of knowledge and belief would be capable of believing to degree i that they do not know while the class of agents with schema 5_{Ki} as a property of knowledge and belief would be capable of knowing that they do not believe to degree i. Again, in the information-theoretic approach schema 5_{Ki} is intuitively more appealing than the other properties of negative introspection.

Proposition 4.9 *Let $\langle S, F\rangle$ be a templated frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of every sentence of $\mathcal{L}'$ which is an instance of schema 5_{Ki} iff for every $s, s' \in S$, if s' is accessible to degree i from s, then $Cont_i(F(s')) \subseteq Cont_i(F(s))$ and $F(s)(x) \geq F(s')(x)$ for every $x \in get_{\uparrow}(F(s), i)$.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of every sentence which is an instance of schema 5_{Ki}. Suppose that there are $s, s' \in S$ such that $s' \in get_{\uparrow}(F(s), i)$ but either $Cont_i(F(s')) \nsubseteq Cont_i(F(s))$ or $F(s)(x) < F(s')(x)$ for some $x \in get_{\uparrow}(F(s), i)$. Then a templated interpretation based on $\langle S, F\rangle$ can be constructed that is not a model of all the instances of schema 5_{Ki}, in particular the sentence $\neg[i]P(a) \rightarrow [n-1]\neg[i]P(a)$. Assume $Cont_i(F(s')) \nsubseteq Cont_i(F(s))$. Take $T = \langle S, F, l\rangle$ such that $v_{l(s')}(P(a)) = 0$ for state $s' \in get_{\uparrow}(F(s), i)$ and $v_{l(s'')}(P(a)) = 1$ for every $s'' \in get_{\uparrow}(F(s'), i)$. So $P(a)$ is not satisfied at state $s' \in get_{\uparrow}(F(s), i)$ and hence $[i]P(a)$ is not satisfied at state s from which it follows that $\neg[i]P(a)$ is satisfied at state s. However, $P(a)$ is satisfied at every state s''

$\in get_{\uparrow}(F(s'), i)$ and thus $[i]P(a)$ is satisfied at state s' from which it follows that $\neg[i]P(a)$ is not satisfied at state s'. But $s' \in get_{\uparrow}(F(s), n-1)$ and thus $[n-1]\neg[i]P(a)$ is not satisfied at state s. But then s fails to satisfy $\neg[i]P(a) \rightarrow [n-1]\neg[i]P(a)$ at state s. So $\langle S, F\rangle$ cannot be a frame-model of $\neg[i]P(a) \rightarrow [n-1]\neg[i]P(a)$. Contradiction. Hence, if s' is accessible from s, then $Cont_i(F(s')) \subseteq Cont_i(F(s))$. Now assume $F(s)(x) < F(s')(x)$ for some $x \in get_{\uparrow}(F(s), i)$. Take $T = \langle S, F, l\rangle$ such that $v_{l(s')}(P(a)) = 0$ for state $s' \in get_{\uparrow}(F(s), i)$ and $v_{l(s'')}(P(a)) = 1$ for every $s'' \in get_{\uparrow}(F(s'), i)$. Suppose $F(s)(s') < F(s')(s')$. So $P(a)$ is not satisfied at state $s' \in get_{\uparrow}(F(s), i)$ and hence $[i]P(a)$ is not satisfied at state s from which it follows that $\neg[i]P(a)$ is satisfied at state s. However, $P(a)$ is satisfied at every state $s'' \in get_{\uparrow}(F(s'), i)$ and thus $[i]P(a)$ is satisfied at state s', from which it follows that $\neg[i]P(a)$ is not satisfied at state s'. But $s' \in get_{\uparrow}(F(s), n-1)$ and thus $[n-1]\neg[i]P(a)$ is not satisfied at state s. But then s fails to satisfy $\neg[i]P(a) \rightarrow [n-1]\neg[i]P(a)$ at state s. So $\langle S, F\rangle$ cannot be a frame-model of $\neg[i]P(a) \rightarrow [n-1]\neg[i]P(a)$. Contradiction. Hence, if s' is accessible from s, then $F(s)(x) \geq F(s')(x)$ for every $x \in get_{\uparrow}(F(s), i)$.

Conversely, suppose that for every $s, s' \in S$, if s' is accessible to degree i from s, then $Cont_i(F(s')) \subseteq Cont_i(F(s))$ and $F(s)(x) \geq F(s')(x)$ for every $x \in get_{\uparrow}(F(s), i)$. To show that $\langle S, F\rangle$ is a a frame-model of every sentence that is an instance of schema 5_{Ki} choose any sentence $\gamma \in \mathcal{L}'$ of the form $\neg[i]\alpha \rightarrow [n-1]\neg[i]\alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $\mathcal{T} = \langle S, F, l\rangle$ of $\mathcal{L}'$ and some state $s \in S$ such that s fails to satisfy γ in $\mathcal{T}$. So it must be the case that $\neg[i]\alpha$ is satisfied at s but $[n-1]\neg[i]\alpha$ is not. If $[n-1]\neg[i]\alpha$ is not satisfied at s then there must be some state $s' \in get_{\uparrow}(F(s), n-1)$ at which $\neg[i]\alpha$ is not satisfied, i.e. at which $[i]\alpha$ is satisfied. So it must be the case that α is satisfied at every $s'' \in get_{\uparrow}(F(s'), i)$. Recall that $\neg[i]\alpha$ is satisfied at s and thus $[i]\alpha$ is not. So there must be some $s''' \in get_{\uparrow}(F(s), i)$ that fails to satisfy α. But since $s' \in get_{\uparrow}(F(s), n-1)$ it follows that $get_{\uparrow}(F(s), n-1) \subseteq get_{\uparrow}(F(s'), n-1)$ with

$F(s)(x) \geq F(s')(x)$ for every $x \in get_{\uparrow}(F(s), n-1)$. Since $s''' \in get_{\uparrow}(F(s), i)$ it follows that $F(s)(s''') \leq i$ and hence that $s''' \in get_{\uparrow}(F(s'), i)$. But then α must be satisfied at s'''. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema 5_{Ki}.

■

Collectively, schemas K_i, T_{n-1}, 4_{Ki}, and 5_{Ki} characterise templated frames whose accessibility functions ensure that every state is accessible from itself and that states which are mutually accessible to degree i have t-orderings which are identical up to level i and have exactly the same indefinite content (to degree i).

Proposition 4.10 *Let $\langle S, F\rangle$ be a templated frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_iT_{n-1}4_{Ki}5_{Ki}$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_i, T_{n-1}, 4_{Ki}, or 5_{Ki} iff for every $s, s' \in S$,*

1. *s is accessible from itself, and*
2. *if s' is accessible to degree i from s, then $Cont_i(F(s)) = Cont_i(F(s'))$ and $F(s')(x) = F(s)(x)$ for every $x \in get_{\uparrow}(F(s'), i)$.*

Proof. See proof in appendix B, section B.1. ■

In the information-theoretic semantics, it becomes possible to model the knowledge and beliefs of an agent as a hierarchy of defeasible beliefs, ranging from those held most tentatively by the agent to those, its knowledge, which are held most strongly. To capture this notion of a hierarchy of defeasible beliefs, a new schema, called schema H, is introduced.

- $H : [j]\alpha \to [i]\alpha \quad \text{for } j > i$

Schema H says that if an agent believes α to degree j, then the agent believes α to every lesser degree as well. In particular, it implies that $[n-1]\alpha \to [0]\alpha$, in other words, if the agent knows α, then the agent also believes α.

Proposition 4.11 *If $\gamma \in \mathcal{L}'$ is of the form $[j]\alpha \rightarrow [i]\alpha$ where $j > i$, then every templated interpretation $\mathcal{T} = \langle S, F, l \rangle$ of $\mathcal{L}'$ is a model of γ.*

Proof. Let γ be any sentence of $\mathcal{L}'$ of the form $[j]\alpha \rightarrow [i]\alpha$ where $j > i$. Choose any templated interpretation $\mathcal{T} = \langle S, F, l \rangle$ of $\mathcal{L}'$ and any state $s \in S$. Assume that γ fails to be satisfied at s. So $[j]\alpha$ is satisfied at s but $[i]\alpha$ is not. Since $[i]\alpha$ is not satisfied at s, there must be some state $s' \in get_{\uparrow}(F(s), i)$ such that α is not satisfied at s'. However, since $[j]\alpha$ is satisfied at s, it must be the case that α is satisfied at every state $s'' \in get_{\uparrow}(F(s), j)$. But since $j > i$, it follows that α is satisfied at every state s'' $\in get_{\uparrow}(F(s), i)$, including state s'. But then α must be satisfied at s'. Contradiction. Thus γ is satisfied at the arbitrarily chosen s and therefore at every state in $\mathcal{T}$. But $\mathcal{T}$ itself was arbitrarily chosen and thus every templated interpretation is a model of γ. ∎

Of all the properties of knowledge and belief, the entailment property (that knowledge entails belief as in schema H) is typically viewed as the least controversial. However, in an argument originally due to Lenzen (1978), it can be shown that if knowledge satisfies the $S5$ properties, belief satisfies the $KD45$ properties, and both the entailment and *positive certainty*[2] properties hold, then it is possible for an agent to have false beliefs. In a re-examination of the argument, Halpern (1996) shows that the entailment property can be viewed as the culprit, and, more importantly, that the original argument fails when the entailment property is weakened so that it applies only to *objective* (nonmodal) sentences, rather than to arbitrary sentences. Objective beliefs form the basis of an epistemic model for control room agents, as will be shown in sections 4.5 and 4.6.

The information-theoretic semantics allows for the modelling of both knowledge and belief in the same semantics so that, apart from being an ideal reasoner with perfect introspection who only has true knowledge and no false beliefs, an agent is capable of

[2]The positive certainty property says that if the agent believes α, then the agent believes that it knows α.

believing that it knows (and does not know) and of knowing that it believes (and does not believe).

4.4 Other models of knowledge and belief

Kraus and Lehmann (1988) also model both knowledge and belief, but in a Kripke semantics using two modal operators, one for knowledge and the other for belief, and two accessibility relations, one for each of the modal operators[3]. The modal operator for knowledge will be denoted by $\Box_K$ and the modal operator for belief by $\Box_B$ with the accessibility relations denoted by R_K and R_B respectively.

The Kripke semantics of Kraus and Lehmann requires that the accessibility relations R_K and R_B satisfy the following conditions:

(**R1**) R_K is an equivalence relation on S

(**R2**) R_B is a serial relation on S

(**R3**) R_B is contained in R_K (i.e. $R_B \subseteq R_K$)

(**R4**) for every $s, s', s'' \in S$, if $(s, s') \in R_K$ and $(s', s'') \in R_B$, then $(s, s'') \in R_B$

Conditions (**R3**) and (**R4**) describe the interrelationship between the two accessibility relations. The intuition behind condition (**R3**) is that states which are indistinguishable to an agent according to its beliefs will be indistinguishable according to its knowledge. However, the agent's beliefs may allow it to distinguish between states which are indistinguishable according to its knowledge. The intuition behind condition (**R4**) is that if states s and s' are indistinguishable according to the agent's knowledge and state s'' is

[3] The notion of common knowledge and common belief and their interrelationship are also considered by Kraus and Lehmann (1988).

considered possible according to the agent's beliefs when the system is in state s', then state s'' will also be considered possible when the system is in state s (since the agent cannot distinguish between states s and s'). Collectively, conditions (**R3**) and (**R4**) ensure that the accessibility relation R_B is euclidean and transitive.

In the approach of Kraus and Lehmann, schemas K, T, and 5 are selected as the properties of knowledge while schemas K and D are selected as the properties of belief. These properties are independent of each other. The interrelationship between knowledge and belief are, however, described by the following two schemas:

- $KL1 : \Box_B \alpha \rightarrow \Box_K \Box_B \alpha$
- $KL2 : \Box_K \alpha \rightarrow \Box_B \alpha$

Schema $KL1$ says that if the agent believes α, then the agent knows that it believes α while schemas $KL2$ says that if the agent knows α, then the agent also believes α. Note that from the collection of schemas, it is possible to derive schema 4 as a property of knowledge and schemas 4 and 5 as properties of belief.

The approach of Kraus and Lehmann is comparable to the information-theoretic approach when restricting to modal operators $[n-1]$ (for knowledge) and $[0]$ (for belief). Their properties of knowledge are simply schemas K_{n-1}, T_{n-1}, and 5_{n-1} while their properties of belief are schemas K_0 and D_0. The interrelationship between knowledge and belief, as captured by their schemas $KL1$ and $KL2$, is captured by schemas 4_{K0} and H (with $j = n-1$ and $i = 0$) respectively. In short, their schemas are a proper subset of the information-theoretic approach. Although the approaches are comparable, the underlying intuition differs. In their approach, knowledge and belief are independent, but interrelated. In the information-theoretic approach, knowledge and belief lie on opposite ends of the same spectrum of defeasible beliefs, with knowledge arising as a special case.

Turning to a very different approach, Moses and Shoham (1993) define three belief operators in terms of a standard knowledge operator. The basic idea is that an agent believes a fact if the agent knows the fact to be true relative to an assumption. Belief can thus be viewed as knowledge-relative-to-assumptions or, to use the terminology proposed by Moses and Shoham, as 'defeasible' knowledge. In making explicit the assumptions, the modal operators for belief become binary. The intuitive reading of $\Box_B^\delta \alpha$ is that the agent knows that, assuming δ holds, so does α. Given a modal operator $\Box_K$ for knowledge, the intuition is formally captured by the following definition:

- $\Box_B^\delta \alpha \stackrel{def}{=} \Box_K(\delta \rightarrow \alpha)$

In their approach, schemas K, T, 4, and 5 are taken as the properties of knowledge. From the definition of the belief operator $\Box_B^\delta$ and the properties of $\Box_K$, it follows that schemas K, 4, and 5 are also properties of belief. However, schema D is not a property of belief, in particular, when the agent knows the assumption to be false. The issue of 'assumptions known to be false' is what prompted Moses and Shoham to formulate two alternative definitions of belief. It is sufficient to note that the different formulations of belief result in different connections between knowledge and belief. The following connections result from the definition of belief stated earlier:

- $MS1 : \Box_K \alpha \rightarrow \Box_B^\delta \alpha$
- $MS2 : (\Box_B^\delta \alpha \leftrightarrow \Box_K \Box_B^\delta \alpha) \wedge (\neg \Box_B^\delta \alpha \leftrightarrow \Box_K \neg \Box_B^\delta \alpha)$
- $MS3 : \Box_B^\delta \Box_K \alpha \leftrightarrow \Box_K \alpha \vee \Box_K \neg \delta$
- $MS4 : \Box_B^\delta \neg \Box_K \alpha \leftrightarrow \neg \Box_K \alpha \vee \Box_K \neg \delta$

Schema $MS1$ says that the agent's knowledge entails its beliefs. Schema $MS2$ is concerned with the agent's knowledge about its beliefs and suggests that the agent knows what it believes and knows what it does not believe. The agent's beliefs about its knowledge is reflected by schemas $MS3$ and $MS4$. It suggests that if the agent believes (assuming δ holds) that it knows α, then the agent either knows α or knows that the assumption δ is false, and vice versa. Similarly, when the agent believes (assuming δ holds) that it does not know α, then the agent either does not know α or knows that the assumption δ is false, and vice versa.

The approach of Moses and Shoham (1993) is comparable to the information-theoretic approach when restricting, as before, to modal operators $[n-1]$ (for knowledge) and $[0]$ (for belief). Their properties of knowledge are simply schemas K_{n-1}, T_{n-1}, 4_{n-1}, and 5_{n-1} while their properties of belief are intuitively reflected by schemas K_0, 4_0, and 5_0. However, modal operators $[0]$ and $\Box_B^\delta$ for belief are different as manifested in the connections between knowledge and belief. In both approaches, knowledge entails belief, as reflected by schemas $MS1$ and H respectively. In terms of knowledge about belief, their schema $MS2$ is stronger than schemas 4_{K0} and 5_{K0}. But in terms of belief about knowledge, there are no counterparts to their schemas $MS3$ and $MS4$ since both schemas require explicit knowledge about the assumption δ.

We now turn to a third, and somewhat more quantitative, approach. Van der Hoek and Meyer (1992) describe a system of graded modalities based on a Kripke semantics consisting of modal operators $\Box_0, \Box_1, ...\Box_k$ (in the finite case) and a single accessibility relation R, which is an equivalence relation. The basic intuition is that knowledge is source dependent and that different sources may not necessarily agree on the truth of a sentence, leading to a notion of graded knowledge that is based on the number of 'disagreements' or exceptions amongst sources. The different sources are taken as the elements of the set S of possible worlds, with $card(S) = k$. The intuitive reading of $\Box_i\alpha$ is that an agent reckons with at most i exceptions for α meaning that $\Box_i\alpha$ is true at a state $s \in S$ iff α is false in at most i of the worlds accessible from s. Given a Kripke model $\mathcal{M} = \langle S, R, l \rangle$, the intuition may be captured by the following definition:

- $\mathcal{M}, s \Vdash \Box_i \alpha$ iff $card\{s' \in S \mid s' \in R(s) \text{ and } \mathcal{M}, s' \nVdash \alpha\} \leq i$

The greater the number of exceptions an agent has to reckon with for α, the less confidence expressed by α. So $\Box_i \alpha$ represents a form of 'uncertain' knowledge and $\Box_0 \alpha$ a form of 'certain' knowledge. As it turns out, $\Box_0$ is the standard knowledge operator since $card\{s' \in S \mid s' \in R(s) \text{ and } \mathcal{M}, s' \nVdash \alpha\} \leq 0$ iff $\mathcal{M}, s' \Vdash \alpha$ for every $s' \in R(s)$. The sentence $\Diamond_i \alpha$ will be used as an abbreviation for $\neg \Box_i \neg \alpha$ and $\Diamond_i ! \alpha$ as an abbreviation for $\Box_0 \neg \alpha$ if $i = 0$, and $\Diamond_{i-1} \alpha \wedge \neg \Diamond_i \alpha$ if $i > 0$. The intuitive reading of $\Diamond_i \alpha$ is that the agent considers more than i alternatives possible in which α is true[4] while the intuitive reading of $\Diamond_i ! \alpha$ is that the agent considers exactly i alternatives possible in which α is true. Van der Hoek and Meyer defined the following schemas for the graded modalities of knowledge for each $i \leq k$:

- $HM1 : \Box_0(\alpha \rightarrow \beta) \rightarrow (\Box_i \alpha \rightarrow \Box_i \beta)$
- $HM2 : \Box_i \alpha \rightarrow \Box_{i+1} \alpha$
- $HM3 : \Box_0 \neg(\alpha \wedge \beta) \rightarrow ((\Diamond_i ! \alpha \wedge \Diamond_j ! \beta) \rightarrow \Diamond_{i+j} !(\alpha \vee \beta))$
- $HM4 : \neg \Box_i \alpha \rightarrow \Box_0 \neg \Box_i \alpha$
- $HM5 : \Box_0 \alpha \rightarrow \alpha$

Schema $HM1$ is a generalisation of schema K. Schema $HM5$ is simply schema T. Although positive introspection is not considered, negative introspection is captured by schema $HM4$, which is a generalisation of schema 5. Schema $HM3$ captures the

[4] To see this note that $\neg \Box_i \neg \alpha$ says that it is not the case that at most i worlds fail to satisfy $\neg \alpha$, i.e. more than i worlds fail to satisfy $\neg \alpha$, i.e. fewer than $k - i$ worlds satisfy $\neg \alpha$, and thus strictly more than i worlds satisfy α.

notion that if the agent knows α and β to be mutually exclusive events, and considers i alternatives in which α is true and, at the same time, j alternatives in which β is true, then altogether the agent has to reckon with $i + j$ alternatives in which either α or β is true. Schema $HM2$ says that if the agent foresees at most i exceptions to α, then the agent also does so with at most $i + 1$ exceptions. It suggests a hierarchy of graded knowledge ranging from 'certain' knowledge through decreasing levels of certainty to 'most uncertain' knowledge as depicted in the sequence $\Box_0\alpha \rightarrow \Box_1\alpha \rightarrow \Box_2\alpha \rightarrow \ldots \rightarrow \Box_k\alpha$.

In the information-theoretic approach, there is a hierarchy of defeasible beliefs ranging from the 'most tentatively held' beliefs through increasing levels of strength to the 'most strongly held' beliefs (or knowledge). Based on schema H, this hierarchy of defeasible beliefs may be depicted in the sequence $[n-1]\alpha \rightarrow [n-2]\alpha \rightarrow [n-3]\alpha \rightarrow \ldots \rightarrow [0]\alpha$. The approach of Van der Hoek and Meyer is comparable to the information-theoretic approach when k is taken to be $n - 1$ and the modal operators are treated according to their relative position in the respective sequences so that, for example, modal operator $[n-1]$ corresponds to $\Box_0$ and modal operator $[n-2]$ corresponds to $\Box_1$. Keeping this in mind, it is clear to see that their schemas $HM2$, $HM4$, and $HM5$ are exactly schemas H, 5_{Ki}, and T_{n-1} respectively. If i is taken to be $n-1$, their schema $HM1$ is simply schema K_{n-1}. However, there is no counterpart to their schema $HM3$. The difference between the approach of Van der Hoek and Meyer and the information-theoretic approach lies in the motivation behind the modal operators. In their approach, the modal operators arise from the number of exceptions that may exist amongst different sources from which the agent may be getting input, in other words from the cardinalities of sets of worlds. In the information-theoretic approach, the modal operators arise from the definite and indefinite information available to the agent when the system is in a specific state, in other words from the set-theoretic inclusion of sets of states rather than from the cardinalities of these sets.

Finally, the approach of Friedman and Halpern (1997) models belief within a general framework of knowledge and plausibility (to which a notion of time is added). The approach extends the framework of Halpern and Fagin (1989) for modelling knowledge in multi-agent systems and is probabilistic in origin (Fagin and Halpern, 1994). It uses a standard modal operator for knowledge, denoted by $\Box_K$, and a standard modal operator for conditionals, denoted by $\rhd$. The intuitive reading of $\alpha \rhd \beta$ is that according to the agent's plausibility measure, α typically implies β. Belief is defined in terms of knowledge and plausibility.

A *plausibility space* is a pair (S, Pl) where Pl is a function, the *plausibility measure*, that maps subsets of S to elements in some arbitrary partially ordered set D[5]. The set D is assumed to contain the elements $\bot_D$ and $\top_D$ such that $\bot_D \leq d \leq \top_D$ for every $d \in D$ with $Pl(S) = \top_D$ and $Pl(\varnothing) = \bot_D$. If $A \subseteq B$ then $Pl(A) \leq Pl(B)$. A plausibility space is said to be *qualitative* if it satisfies the following two conditions:

- If A, B, and C are disjoint subsets of S, $Pl(A \cup B) > Pl(C)$, and $Pl(A \cup C) > Pl(B)$, then $Pl(A) > Pl(B \cup C)$
- If $Pl(A) = Pl(B) = \bot_D$, then $Pl(A \cup B) = \bot_D$

The notion of plausibility is captured by adding a *plausibility assignment* P, which is a function that maps worlds to plausibility spaces, to a Kripke model $\mathcal{M} = \langle S, R_K, l \rangle$ for knowledge. The accessibility relation R_K is an equivalence relation and $R_K(s) = \{s' \in S \mid (s, s') \in R_K\}$. Given a plausibility assignment $P(s) = (S_s, Pl_s)$ and letting $Mod_{\mathcal{M},S_s}(\alpha)$ denote the (local) models of α restricted to S_s, the semantics of $\alpha \rhd \beta$ can be defined as follows:

[5] This is the simplified definition of a plausibility space as used by Friedman and Halpern (1997). Note that a plausibility space is a direct generalisation of a probability space in which the plausibility measure replaces the probability measure.

- $\mathcal{M}, s \Vdash \alpha \rhd \beta$ if $\begin{cases} Pl_s(Mod_{\mathcal{M},S_s}(\alpha)) = \bot_D \text{ or} \\ Pl_s(Mod_{\mathcal{M},S_s}(\alpha \wedge \beta)) > Pl_s(Mod_{\mathcal{M},S_s}(\alpha \wedge \neg\beta)) \end{cases}$

An agent is said to believe α, denoted by $\Box_B\alpha$, if the agent knows that α is more plausible than $\neg\alpha$ in all the worlds that the agent considers possible. Formally, this notion of belief is defined as follows:

- $\Box_B\alpha \stackrel{def}{=} \Box_K(\top \rhd \alpha)$

In the approach of Friedman and Halpern, the properties of knowledge are, as usual, schemas K, T, 4, and 5. To reason about belief, the following conditions are imposed on the Kripke model $\mathcal{M} = \langle S, R_K, l, P\rangle$ for knowledge and plausibility:

- **QUAL** : $P(s)$ is qualitative for all $s \in S$
- **CONS** : $S_s \subseteq R_K(s)$ for all $s \in S$
- **NORM** : $\bot_s > \top_s$ for all $s \in S$

Condition $QUAL$ ensures that schema K is a property of belief. Condition $CONS$ ensures that if the agent knows α it will also believe α. This connection between knowledge and belief is exactly schema $KL2$ of Kraus and Lehmann. (Schema $KL1$ follows directly from the definition of $\Box_B$ and schema 4 for knowledge.) Condition $NORM$ ensures that the agent does not consider all sets to be completely implausible. Friedman and Halpern show that by imposing conditions $CONS$ and $NORM$, schemas D, 4, and 5 become properties of belief too.

The approach of Friedman and Halpern is very close to that of Kraus and Lehmann in the sense that the properties of knowledge and belief are the same as is the interrelationship between knowledge and belief. Their schemas are therefore also a proper subset of the information-theoretic approach. Although the approaches are comparable, the underlying intuition differs. In their approach, belief is probabilistic in nature and defined in terms of knowledge and plausibility. In the information-theoretic approach, knowledge and belief lie on opposite ends of the same spectrum of defeasible beliefs, with knowledge arising as a special case.

4.5 Epistemic entrenchment

The idea behind *epistemic entrenchment* is that some beliefs are more important than others and if an agent is forced to give up some of its beliefs, it will give up those with the lowest epistemic entrenchment. Recall that by language L we mean the nonmodal transparent propositional language and by language $\mathcal{L}'$ we mean the modal language based on L but containing all the belief operators $[i]$.

The formalisation of epistemic entrenchment as presented by Gärdenfors (1988) and Gärdenfors and Makinson (1988) applies to propositional languages under a traditional truth-value semantics. In the setting of epistemic logic, epistemic entrenchment applies to *objective* (nonmodal) beliefs as opposed to *subjective* (modal) beliefs. Objective beliefs are sentences of L. In the context of the modal language $\mathcal{L}'$, objective beliefs will always be prefaced with a modal operator $[i]$ but as befits sentences of the nonmodal language L will contain no subsequent occurrences of modal operators. For example, in the sentence $[0]P(c)$, the atom $P(c)$ is an objective belief. In the context of diagrammable systems, an *epistemic entrenchment ordering*, or *EE-ordering* for short, is a certain kind of binary relation on L, under the classical interpretation $M_0 = \langle S, l \rangle$. If α and β are sentences of L, then $\alpha \sqsubseteq \beta$ will be taken as a shorthand for 'β is at least as epistemically entrenched as α'. Note that an epistemic entrenchment ordering is defined only in relation to a specific belief set (or theory) of L.

Definition 4.4 *Let $\alpha, \beta \in L$ and Γ a theory of L. A binary relation $\sqsubseteq$ on L is an*

EE-ordering *(with respect to* Γ*) iff it satisfies the following postulates:*

(**EE1**) $\sqsubseteq$ *is transitive*

(**EE2**) *If* $\alpha \models \beta$ *then* $\alpha \sqsubseteq \beta$

(**EE3**) *For all* $\alpha, \beta \in \Gamma$, $\alpha \sqsubseteq \alpha \wedge \beta$ *or* $\beta \sqsubseteq \alpha \wedge \beta$

(**EE4**) *If* $\Gamma \neq L$ *then* $\alpha \notin \Gamma$ *iff* $\alpha \sqsubseteq \beta$ *for all* β

(**EE5**) *If* $\beta \sqsubseteq \alpha$ *for all* β *then* $\models \alpha$

Postulate (**EE1**) is a minimal requirement to impose on an ordering relation. If α entails β and either α or β must be retracted from Γ, then giving up α and retaining β is a smaller change than to give up β, because then α has to given up as well if Γ is to remain a theory. Hence the requirement by postulate (**EE2**) for β to be at least as epistemically entrenched as α. If the conjunction $\alpha \wedge \beta$ must be retracted from Γ, then either α or β has to be given up as stated by postulate (**EE3**). As a consequence of postulate (**EE3**) the ordering $\sqsubseteq$ has to be connected (and hence, a total preorder, since reflexivity follows from **EE2**). Postulate (**EE4**) requires that all sentences not contained in Γ are minimal in the ordering while postulate (**EE5**) requires that tautologies are maximal in the ordering.

An interesting result about the epistemic entrenchment orderings on sentences is that they are closely related to total preorders on sets of possible worlds (Meyer, Labuschagne, and Heidema, 2000a). The basic idea is to construct an EE-ordering from a *faithful* total preorder on the set of states by a suitable power construction.

Definition 4.5 *Let* Γ *be a theory of* L*. A total preorder* $\preceq$ *on* S *is* ***faithful*** *with respect to* Γ *iff*

1. $s \prec s'$ *for every* $s \in Mod(\Gamma)$ *and* $s' \notin Mod(\Gamma)$*, and*

2. $s \not\prec s'$ *for every* $s, s' \in Mod(\Gamma)$.[6]

The *power construction* lifts a faithful total preorder on states to an ordering on sets of states. Because every sentence is associated with a specific set of possible worlds - the set of states at which the sentence is true, or its set of models - the power ordering constructed on sets of states may be viewed as an ordering on sentences (or more accurately, as an ordering on propositions).

Definition 4.6 *Let* Γ *be a theory of* L. *Suppose* $\preceq$ *is a faithful total preorder on* S *with respect to* Γ. *Then the* ***power order*** $\sqsubseteq_{\preceq}$ *on* L *induced by* $\preceq$ *is defined, for all* $\alpha, \beta \in L$, *by* $\alpha \sqsubseteq_{\preceq} \beta$ *iff for every* $s' \in Mod(\neg\beta)$ *there is some* $s \in Mod(\neg\alpha)$ *such that* $s \preceq s'$.

The power order $\sqsubseteq_{\preceq}$ on L induced by $\preceq$ could have been defined information-theoretically as follows: $\alpha \sqsubseteq_{\preceq} \beta$ iff for every $s' \in Cont(\beta)$ there is an $s \in Cont(\alpha)$ such that $s \preceq s'$. (This is because $Cont(\alpha) = NMod(\alpha)$ and $Mod(\neg\alpha) = S - Mod(\alpha)$.) The definition states that β is at least as epistemically entrenched as α iff for every state s' there is some state s such that β excludes s' more strongly than α excludes s. More importantly however, is that every power order $\sqsubseteq_{\preceq}$ on L induced by a faithful total preorder $\preceq$ on S is an EE-ordering (and vice versa).

Theorem 4.1 (Meyer, Labuschagne, and Heidema, 2000a) *Let* Γ *be a theory of* L. *A binary relation* $\sqsubseteq$ *on* L *is an EE-ordering (with respect to* Γ*) iff it is the power order* $\sqsubseteq_{\preceq}$ *on* L *induced by some faithful total preorder* $\preceq$ *on* S *(with respect to* Γ*).*

Turning our attention to t-orderings, recall that, by proposition 3.14, every t-ordering $t \in T_S$ induces a total preorder on the set of states, defined as the relation $\preceq$ on S such

[6] *Due to the finiteness of* L, *the smoothness condition is not required. If* L *were infinite, then for every sentence* $\alpha \in L$, $\preceq$ *would need to be* X*-smooth, where* $X = Mod_{M_0}(\alpha)$ *and* $M_0 = \langle S, l\rangle$ *is the classical interpretation of* L.

that $s \preceq s'$ iff $t(s) \leq t(s')$. As shown below, the total preorder $\preceq$ induced by t is faithful with respect to the theory $\Gamma = Th_M(bottom(t))$ under a (finite) extensional interpretation $M = \langle S, l \rangle$.

Proposition 4.12 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t \in T_S$ be a t-ordering and $\Gamma = Th_M(bottom(t))$ a theory of L. Then the total preorder $\preceq$ on S induced by t is faithful (with respect to Γ).*

Proof. Since L is finite and M an extensional interpetation of L, it holds that for every $X \subseteq S$, $X = Mod_M(Th_M(X))$. So $Mod_M(\Gamma) = bottom(t)$. But then $t(s) = 0$ for every $s \in Mod_M(\Gamma)$ and $t(s') > 0$ for every $s' \notin Mod_M(\Gamma)$. Hence $s \prec s'$ for every $s \in Mod_M(\Gamma)$ and $s' \notin Mod_M(\Gamma)$ and secondly, $s \not\prec s''$ for every $s, s'' \in Mod_M(\Gamma)$. ∎

The connection between t-orderings and EE-orderings is via the faithful total preorders induced by t-orderings, from which the power orders are constructed. If $\preceq$ is the faithful total preorder on S induced by t-ordering t (with respect to $\Gamma = Th(bottom(t))$), then the power order $\sqsubseteq_{\preceq}$ on L induced by $\preceq$ is an EE-ordering on L (with respect to Γ) by theorem 4.1. However, the connection between t-orderings and EE-orderings can also be defined directly via the notion of plausibility, which was introduced in section 3.6. Recall that the plausibility of a sentence α (with respect to a t-ordering t) was defined as the greatest level in the t-ordering such that α is satisfied at every state (in the templated interpretation) lying at or below that level, or more succinctly, as the greatest j such that $get_{\uparrow}(t, j) \subseteq Mod_T(\alpha)$ (or -1 if there is no such j). A plausibility of -1 indicates that the sentence α is not plausible with respect to the t-ordering t, in other words, that $get_{\uparrow}(t, 0) \not\subseteq Mod_T(\alpha)$. The plausibility of a sentence (with respect to a t-ordering) is thus relative to the states occupying the lowest level in the t-ordering.

Definition 4.7 *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha, \beta \in L$. Then the **plausibility ordering** $\sqsubseteq_P$ on L (with respect to t) is defined as $\alpha \sqsubseteq_P \beta$ iff*

$pl(t, \alpha) \leq pl(t, \beta)$.

Proposition 4.13 *Let $T_0 = \langle S, t, l \rangle$ be a templated interpretation of L for $S = U_T$ and l the identity function. Let $\Gamma = Th_{T_0}(bottom(t))$ be a theory of L. Then the plausibility ordering $\sqsubseteq_P$ on L (with respect to Γ) is an EE-ordering on L (with respect to Γ).*

Proof. If a binary relation on L is the power order induced by some faithful total preorder $\preceq$ on S (with respect to Γ), then the binary relation is an EE-ordering on L (with respect to Γ), by theorem 4.1. Let $\preceq$ be the faithful total preorder on S (with respect to Γ) induced by t and let $\sqsubseteq_{\preceq}$ be the power order induced by $\preceq$. If it can be shown that $\sqsubseteq_P = \sqsubseteq_{\preceq}$, then $\sqsubseteq_P$ will be an EE-ordering (with respect to Γ). Choose any α and β in $\sqsubseteq_P$. So $pl(t, \alpha) \leq pl(t, \beta)$. Let $j = pl(t, \beta)$. So j is the greatest level in t such that $get_{\uparrow}(t, j) \subseteq Mod_{T_0}(\beta)$. But then for every every $s' \in Mod_{T_0}(\neg\beta)$, $t(s') > j$. Let $i = pl(t, \alpha)$. So i is the greatest level in t such that $get_{\uparrow}(t, i) \subseteq Mod_{T_0}(\alpha)$. But then there must be some $s \in Mod_{T_0}(\neg\alpha)$ such that $t(s) > i$. But $i \leq j$ and thus $t(s) \leq t(s')$, i.e. $s \preceq s'$. Hence $\alpha \sqsubseteq \beta$. Since α and β were chosen arbitrarily, it follows that $\sqsubseteq_P$ is a subset of $\sqsubseteq_{\preceq}$.

Choose any α and β in $\sqsubseteq_{\preceq}$. So for every $s' \in Mod_{T_0}(\neg\beta)$ there is an $s \in Mod_{T_0}(\neg\alpha)$ such that $s \preceq s'$, i.e. such that $t(s) \leq t(s')$. Let $i = t(s)$. So $s \in get_{\uparrow}(t, i)$ and s fails to satisfy α. But then $get_{\uparrow}(t, i) \not\subseteq Mod_{T_0}(\alpha)$, i.e. $pl(t, \alpha) < i$. Let $j = t(s')$. So $s' \in get_{\uparrow}(t, j)$ and s' fails to satisfy δ. But then $get_{\uparrow}(t, j) \not\subseteq Mod_{T_0}(\beta)$, i.e. $pl(t, \beta) < j$. But $i \leq j$ and thus $pl(t, \alpha) \leq pl(t, \beta)$. Hence $\alpha \sqsubseteq_P \beta$. Since α and β were chosen arbitrarily, it follows that $\sqsubseteq_{\preceq}$ is a subset of $\sqsubseteq_P$. But then $\sqsubseteq_P = \sqsubseteq_{\preceq}$, as required. ■

From the previous result it follows that if α and β are sentences of L, then β is at least as epistemically entrenched as α iff the plausibility of α is less than or equal to the plausibility of β (both with respect to t). The belief set in question is, of course, the set $\Gamma = Th_{T_0}(bottom(t))$.

Epistemic entrenchment is compatible with the modal operators of the language $\mathcal{L}'$. Let $s \in S$. Now suppose t is taken to be the t-ordering such that $t = F(s)$ where F is the accessibility function of some templated interpretation $\mathcal{T} = \langle S, F, l \rangle$ of $\mathcal{L}'$. Then the plausibility of a sentence $\alpha \in L$ (with respect to $F(s)$) is the greatest j such that for every state $s' \in get_{\uparrow}(F(s), j)$, s' satisfies α (under $\mathcal{T}$), in other words, the greatest j such that s satisfies $[j]\alpha$. The *entrenchment* of a sentence $\alpha \in L$ (at a state s) can be intuitively described as the greatest j in the t-ordering $F(s)$ such that the agent believes to degree j that α is true, which is nothing other than the plausibility of α with respect to $F(s)$. More formally, the connection between the epistemic entrenchment of sentences of L and the modal operators of the encompassing modal language $\mathcal{L}'$ is captured by the following result, due to Labuschagne and Ferguson (2002).

Proposition 4.14 (Labuschagne and Ferguson, 2002) *Let $\mathcal{T}_0 = \langle S, F, l \rangle$ be a templated interpretation of $\mathcal{L}'$ for $S = U_T$ and l the identity function . Let $\alpha, \beta \in L$ and $s \in S$. Let $\preceq$ be the faithful total preorder on S induced by t-ordering $F(s)$ and let $\sqsubseteq$ be the EE-ordering on L induced by $\preceq$. Then $\alpha \sqsubseteq \beta$ (at s) iff s satisfies $[i]\alpha \rightarrow [i]\beta$ for all i.*

Proof. Suppose that $\alpha \sqsubseteq \beta$. So for every $x \in Mod(\neg\beta)$ there is an $y \in Mod(\neg\alpha)$ such that $y \preceq x$, i.e. $F(s)(y) \leq F(s)(x)$. Suppose that $[j]\alpha \rightarrow [j]\beta$ fails to be satisfied at s. So $[j]\alpha$ is satisfied at s but $[j]\beta$ is not. Since $[j]\beta$ is not satisfied at s, there must be some state $s' \in get_{\uparrow}(F(s), j)$ such that β is not satisfied at s'. But then $s' \in Mod_{\mathcal{T}_0}(\neg\beta)$ and $F(s)(s') \leq j$. However, since $[j]\alpha$ is satisfied at s, it must be the case that α is satisfied at every state $s'' \in get_{\uparrow}(F(s), j)$. So $Mod_{\mathcal{T}_0}(\alpha) \subseteq get_{\uparrow}(F(s), j)$. But then $F(s)(s''') > j$ for every $s''' \in Mod_{\mathcal{T}_0}(\neg\alpha)$. So $F(s)(s') < F(s)(s''')$ for $s' \in Mod_{\mathcal{T}_0}(\neg\beta)$ and $s''' \in Mod_{\mathcal{T}_0}(\neg\alpha)$. Contradiction. So s satisfies $[i]\alpha \rightarrow [i]\beta$ for all i.

Conversely, suppose that s satisfies $[i]\alpha \rightarrow [i]\beta$ for all i. Assume it is not the case that $\alpha \sqsubseteq \beta$. So there must be some $x \in Mod_{\mathcal{T}_0}(\neg\beta)$ such that for every $y \in Mod_{\mathcal{T}_0}(\neg\alpha)$,

$x \prec y$, i.e. $F(s)(x) < F(s)(y)$. Let $j = F(s)(x)$. So $x \in get_{\uparrow}(F(s), j)$. But then there is some state $s' \in get_{\uparrow}(F(s), j)$, namely x, such that β is not satisfied at s'. Hence $[j]\beta$ is not satisfied at s. On the other hand, $F(s)(y) > j$ for every $y \in Mod_{\mathcal{T}_0}(\neg\alpha)$. But then α must be satisfied at every state $s'' \in get_{\uparrow}(F(s), j)$. Hence $[j]\alpha$ is satisfied at s. But then $[j]\alpha \rightarrow [j]\beta$ cannot be satisfied at s. Contradiction. So it must be case that $\alpha \sqsubseteq \beta$. ■

In terms of epistemic entrenchment, the agent's beliefs range from the least entrenched beliefs at one end of the spectrum to the most deeply entrenched beliefs at the other end of the spectrum. From an agent-oriented perspective to knowledge and belief (Labuschagne and Heidema, 2001), it makes sense to think of knowledge as consisting of beliefs that the agent is sufficiently reluctant to surrender, which, when modelled by an epistemic entrenchment relation, would be those beliefs that are sufficiently entrenched.

The most deeply entrenched beliefs (at a state s) are those sentences $\alpha \in L$ with the highest entrenchment value (or plausibility), in other words, those sentence α such that s satisfies $[n]\alpha$ - the tautologies. But, as suggested by Girle (1998), agents who know only true sentences are unrealistic and, thus, we shall take as knowledge the next-to-most deeply entrenched beliefs (which, of course, would include the tautologies). The next-to-most deeply entrenched beliefs (at a state s) are those sentences $\alpha \in L$ with an entrenchment value (or plausibility) of $n-1$. The agent's knowledge at state s would thus constitute those sentences $\alpha \in L$ such that s satisfies $[n-1]\alpha$. The least entrenched beliefs (at a state s) are those sentences $\beta \in L$ with an entrenchment value (or plausibility) of 0, and thus precisely the sentences β such that s satisfies $[0]\beta$.

As we have shown, the view that an agent's defeasible beliefs form a hierarchy ranging from its most strongly held beliefs (or knowledge) to its most tentatively held beliefs arises naturally from the notion of epistemic entrenchment and its connection with t-orderings. The use of a distinguished level in a t-ordering for definitely excluded states allows for a qualitative notion of 'sufficiently entrenched' to characterise knowledge.

4.6 T-orderings as epistemic states

In defining a model of epistemic states suitable for control room agents, who, as mentioned before, may be summarised as first-order intentional systems having specific informational and motivational attitudes, it suffices to consider only objective beliefs.

An epistemic state contains, in one form or another, the knowledge and beliefs of an agent at a specific point in time, the formation of which is dependent on the information available to the agent. At a specific point in time the system under consideration will be in a specific state and it is therefore reasonable to view an epistemic state as relative to a specific state (or possible world). The information-theoretic semantics of Labuschagne and Ferguson (2002) that was recalled earlier assigns a t-ordering to each state of the system. The connection between epistemic entrenchment and t-orderings has shown that the agent's knowledge at a state $s \in S$ comprises those sentences $\alpha \in L$ such that α is satisfied at every state $s' \in get_{\uparrow}(t, n-1)$ while the agent's (most tentative) beliefs (at s) consist of those sentences $\beta \in L$ such that β is satisfied at every state $s' \in get_{\uparrow}(t, 0)$, where t is the t-ordering assigned to s. This provides an information-theoretic justification for representing epistemic states by t-orderings.

Definition 4.8 *Let L be a finitely generated transparent propositional language and let S be a non-empty set of states with $card(S) = n$. An* ***epistemic state*** *E is represented by a regular t-ordering $t_E \in T_E$.*

In choosing to represent epistemic states by regular t-orderings, an information-theoretic model of epistemic states is adopted, which may be seen as a generalisation of the possible worlds model of epistemic states. Models of epistemic states based on possible worlds have been used by many authors (Harper, 1977; Grove, 1988; Spohn, 1988; Katsuno and Mendelzon, 1991; Darwiche and Pearl, 1997) and in subsequent chapters some of these, and other models, will be explored in more detail.

Definition 4.9 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state. Then associated with t_E are a* ***belief set*** *$Bel(t_E) \subseteq L$ and a* ***knowledge set*** *$Know(t_E) \subseteq L$ defined as $Bel(t_E) = Th_M(bottom(t_E))$ and $Know(t_E) = Th_M(get_{\uparrow}(t_E, n-1))$.*

The representation of epistemic states by regular t-orderings satisfies the principle of Duality in the sense that both knowledge and belief are represented with knowledge not restricted to tautologies. The belief set is the theory determined by the set of maximally preferred states while the knowledge set is the theory determined by the set of included states (= not definitely excluded). Given an extensional interpretation $M = \langle S, l \rangle$ of L, the belief set $Bel(t_E)$ is an axiomatisation of $bottom(t_E)$ and the knowledge set $Know(t_E)$ an axiomatisation of $get_{\uparrow}(t_E, n-1)$, in other words, $Mod_M(Bel(t_E)) = bottom(t_E)$ and $Mod_M(Know(t_E)) = get_{\uparrow}(t_E, n-1)$. Since L is finite, there exist a finite axiomatisation of $bottom(t_E)$ and of $get_{\uparrow}(t_E, n-1)$ under M, denoted by $bel(t_E)$ and $know(t_E)$ respectively. The sentence $bel(t_E)$ will be referred to as the *belief assertion* associated with epistemic state t_E (for reasons that will become clear in due course) while the sentence $know(t_E)$ will be referred to as the *knowledge assertion* associated with t_E.

By restricting the representation of epistemic states to regular t-orderings, the notion of relative distance, which can be expressed by unrestricted t-orderings, is removed and the approach becomes purely qualitative, thereby satisfying the principle of Qualitativeness. Since a regular t-ordering is a t-ordering in normal form that is not strongly contradictory, it follows that the bottom level of the t-ordering is non-empty, which ensures that the belief set and knowledge set associated with an epistemic state are satisfiable, thus satisfying the principle of Consistency. The principle of Logical Closure is satisfied by ensuring that both the belief set and knowledge set associated with an epistemic state are theories.

A sentence α is said to be *accepted* in an epistemic state t_E iff $bottom(t_E) \subseteq Mod_M(\alpha)$.

In other words, a sentence is accepted in an epistemic state iff it is an element of the belief set associated with the epistemic state. We think of 'accepted' as an abbreviation of 'accepted as a basis for action'. Some actions may entail greater risks for the agent than others, and consequently one would want to have some actions requiring more stringent preconditions than other actions may require. The notion of acceptance ought therefore to be nuanced. Sentences can be accepted with varying degrees of firmness. The degree of firmness with which a sentence is accepted in an epistemic state is defined as the plausibility of the sentence with respect to the (regular t-ordering representing the) epistemic state, thus reflecting the entrenchment of the sentence with respect to the belief set. The most strongly accepted sentences in an epistemic state (next to the tautologies) are, of course, the elements of the knowledge set associated with the epistemic state.

A sentence α is said to be *rejected* in an epistemic state t_E iff $bottom(t_E) \cap Mod_M(\alpha) = \varnothing$. Rejected sentences are disbeliefs. Using the notion of distrust, which was introduced in section 3.6, sentences can be rejected in an epistemic state with varying degrees of firmness. Recall that the distrust of a sentence α (with respect to a t-ordering t) was defined as the least level in the t-ordering such that α is satisfied at some state lying at or below that level, or alternatively, as the least j such that $get_{\uparrow}(t, j) \cap Mod_M(\alpha) \neq \varnothing$ (or n if there is no such j). The degree of firmness with which a sentence is rejected in an epistemic state is defined as the distrust of the sentence with respect to the (regular t-ordering representing the) epistemic state.

The information-theoretic model of epistemic states is, from a syntactic perspective, based on the *coherentist* approach, which takes the view that sentences constituting justifications for beliefs need not be part of an epistemic model; instead, the focus is on how coherent beliefs are with other beliefs that are accepted in the epistemic state. This is in contrast with the *foundationist* approach, which takes the view that an agent should keep track of the justifications for its beliefs and that sentences that have no justification should not be accepted as beliefs in the epistemic state. The distinction between the coherentist and foundationist approaches has resulted in ongoing debate amongst researchers (Sosa, 1980; Harman, 1986; Gärdenfors, 1990; Hansson and Olsson, 1999, Hansson, 2006). From our semantic perspective, the debate is based on a false dichotomy. Our own approach illustrates a third way, for, while technically coherentist, our approach does include semantic (as opposed to syntactic) justification for beliefs in the form of an epistemic state. (Recall that an epistemic state is a t-ordering, not merely a set of sentences.)

One aspect that neither the coherentist nor the foundationist approach considers seriously is how an agent arrives at an initial epistemic state. To arrive at an initial epistemic state, an agent must have available some information about the system under consideration. In the context of diagrammable systems this will take the form of *fixed information*, either given as a system diagram (or blueprint) or derived from perception. Perception produces images, from which facts and constraints in the form of sentences are produced by a psychological process of analog-to-discrete transformation (Harnad, 1990). This fixed information allows the agent to permanently rule out from consideration zero, one, or more of those states of the system that are unrealisable. On the other hand, observations, which provide state-dependent evidence, allow the agent to rule out some states definitely though not permanently. Additionally, an agent will usually learn or be given *default rules* about the system, which allow the agent to treat some of the realisable states as more normal (or typical) than others. To form the initial epistemic state of the agent, we shall single out the forms of information that might be more persistent than observational evidence; the agent's fixed information refined by its default rules about the system using a refinement operation on t-orderings in normal form. The refinement operation relies on a lexicographical ordering that is induced by t-orderings in normal form.

Definition 4.10 *Suppose X and Y are two sets on which there are order relations $\leq_X$ and $\leq_Y$ respectively. If A is any subset of the Cartesian product $X \times Y$, then the* ***lexicographic ordering*** *on A is the relation $\preceq$ such that $(x, y) \preceq (x', y')$ iff $x \leq_X x'$ and if $x = x'$ then $y \leq_Y y'$.*

Lemma 4.1 *Suppose X and Y are two finite sets on which there are linear orders $\leq_X$ and $\leq_Y$ respectively. Then the lexicographic ordering $\preceq$ on any subset A of the Cartesian product $X \times Y$ is a well-ordering on A.*

Proof. See proof in appendix B, section B.2. ■

Definition 4.11 *Let S be a non-empty set of states with $card(S) = n$ and let $B = \{0, 1, \ldots, n\}$ be the template for S with the usual linear order $\leq$ on B. Let $t_1, t_2 \in T_N$. Then the* ***index set*** *$A \subseteq B \times B$ induced by $\langle t_1, t_2 \rangle$ is defined as $A = \{(x, y) \mid x = t_1(s)$ and $y = t_2(s)$ for some $s \in S\}$.*

From lemma 4.1 it follows that if t_1 and t_2 are t-orderings in normal form, then the lexicographic ordering $\preceq$ on the index set A induced by $\langle t_1, t_2 \rangle$ is a well-ordering on A.

Definition 4.12 *Let $t_1, t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2 \rangle$ and well-ordered by the lexicographic ordering $\preceq$. The* ***refinement operation*** *$\boxminus$ is a binary operation on T_N where $t_1 \boxminus t_2$, which is the result of refining t_1 by t_2, is defined as follows*

- $(t_1 \boxminus t_2)(s) = \begin{cases} card(seg(t_1(s), t_2(s))) & \text{if } t_1(s) < n \\ n & \text{otherwise} \end{cases}$
  where $seg(t_1(s), t_2(s))$ is the initial segment of $(t_1(s), t_2(s)) \in A$.

The effect of the refinement operation is to push some of the states in the t-ordering being refined up some level(s). However, in the refinement of a t-ordering by another t-ordering, the ordering of the original t-ordering is respected in the sense that if some state $s \in S$ was at a level below state $s' \in S$, then s will be at a level below s' in the refined t-ordering too. In essence, the refinement takes place on states (below the top level) that occupy the same level. A refinement operation has no effect on the definite content of the original t-ordering but may increase its indefinite content.

Proposition 4.15 *Let S be a non-empty set of states with $card(S) = n$. Let $t_1, t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2 \rangle$ and well-ordered by the lexicographic ordering $\preceq$. Then the following holds for every $s, s' \in S$:*

1. $t_1(s) \leq (t_1 \boxminus t_2)(s)$
2. *if* $t_1(s) < t_1(s')$ *then* $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$
3. *if* $t_1(s) = t_1(s') \neq n$ *and* $t_2(s) < t_2(s')$ *then* $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$
4. $bottom(t_1 \boxminus t_2) \subseteq bottom(t_1)$
5. $Cont_D(t_1) = Cont_D(t_1 \boxminus t_2)$
6. $Cont_0(t_1) \subseteq Cont_0(t_1 \boxminus t_2)$

Proof. See proof in appendix B, section B.2. ■

The agent's fixed information reflects what the agent knows about the system specification, in other words, what the agent knows about states of the system that will never arise. The agent's default rules about the system, on the other hand, reflect what the agent has learned about the system, in other words, what the agent believes to be the most normal (or typical) states of the system. In the exceptional case where an agent has no fixed information and no default rules about the system, the initial epistemic state $t_I \in T_E$ of the agent will be taken to be tautological. Otherwise it would be arrived at in the manner defined below.

Definition 4.13 *Let $t_{SD} \in T_D$ (with t_{SD} regular) represent the agent's fixed information and let $t_{SI} \in T_I$ represent the agent's default rules. Then the agent's* ***initial epistemic state*** *$t_I \in T_E$ is determined by taking $t_I(s) = (t_{SD} \boxminus t_{SI})(s)$.*

Collectively, the information available to an agent provides the *epistemic framework* from which the agent's epistemic state is formed, and changed. For now, our focus has been on those components of the epistemic framework which allow the agent to arrive in a natural way at an initial epistemic state. However, in the next chapter, the focus will shift to those components of the epistemic framework which allow the agent to perform certain types of changes to its epistemic state based on state-dependent information.

Strictly speaking, if the idea is accepted that epistemic states should contain all the relevant information for determining how epistemic change operations should be performed, then the components of the agent's epistemic framework should form part of the agent's epistemic state. However, for notational convenience, the components of an agent's epistemic framework will be kept separate so that the agent's epistemic state is represented simply by a regular t-ordering (plus an associated belief set and knowledge set).

4.7 Ordinal conditional functions

The representation of epistemic states by regular t-orderings is closely related to the representation of epistemic states by the *ordinal conditional functions* (OCFs) of Spohn (1988). The notion of a OCF is a generalisation of the concept of a *well-ordered partition* (WOP) on the set of possible worlds S to the effect that such a partition may contain empty terms. In this sense, an OCF is very similar to the notion of a t-ordering as a distribution of a strict linearly ordered partition on S over a fixed construct of boxes (or template) where some of the boxes may be empty. However, the role of the top box in a t-ordering has no counterpart in an OCF. In choosing to represent epistemic states by OCFs, as opposed to WOPs, the notion of relative distance between possible worlds becomes as important as the ordering between worlds. This notion of *distance* is the key difference between representing epistemic states by OCFs versus regular t-orderings: OCFs are quantitative in nature whereas regular t-orderings are purely qualitative.

Formally, an *ordinal conditional function* (OCF) is a function k from a given set of possible worlds S into the class of ordinals such that $k(s) = 0$ for at least one $s \in S$. Intuitively, the ordinals represent degrees of plausibility; the smaller the ordinal, the more plausible the possible world. Spohn uses the traditional truth-value semantics of propositional logic, which of course, corresponds to the classical interpretation $M_0 = \langle S, l \rangle$. The belief set associated with epistemic state k, denoted by $Bel(k)$, is defined by $Mod(Bel(k)) = \{s \in S \mid k(s) = 0\}$. From the definition of an OCF it follows that every belief set $Bel(k)$ will be satisfiable.

Similarly to t-orderings, every OCF k induces a total preorder on the set of states, defined as the relation $\preceq$ on S such that $s \preceq s'$ iff $k(s) \leq k(s')$, which is faithful with respect to the theory $\Gamma = Th(\{s \in S \mid k(s) = 0\})$.

Proposition 4.16 *Let k be an OCF and let $\Gamma = Th(\{s \in S \mid k(s) = 0\})$ be a theory of L. Then k induces a total preorder $\preceq$ on S that is faithful (with respect to Γ).*

Proof. Let $\preceq$ be the relation on S induced by OCF k such that $s \preceq s'$ iff $k(s) \leq k(s')$. Clearly, the relation $\preceq$ is reflexive, transitive, and total. To see that $\preceq$ is faithful (with respect to Γ) note that for every $X \subseteq S$, $X = Mod(Th(X))$ (since L is finite and the underlying interpretation extensional). So $Mod(\Gamma) = \{s \in S \mid k(s) = 0\}$. But then $k(s) = 0$ for every $s \in Mod(\Gamma)$ and $k(s') > 0$ for every $s' \notin Mod(\Gamma)$. Hence $s \prec s'$ for every $s \in Mod(\Gamma)$ and $s' \notin Mod(\Gamma)$ and secondly, $s \not\prec s''$ for every $s, s'' \in Mod(\Gamma)$. ■

Spohn extends the definition of an OCF to sets of possible worlds, or sentences (because every sentence is associated with a specific set of possible worlds, its set of models), by associating every non-contradictory sentence with the smallest ordinal assigned to any of its models. An OCF $k : S \rightarrow Ord$ can be extended to a function k' whose domain includes all non-contradictory sentences by taking $k'(\alpha) = \min\{k(s) \mid s \in Mod(\alpha)\}$.

A sentence α is said to be accepted in epistemic state k iff $k'(\neg\alpha) > 0$, rejected in k iff $k'(\alpha) > 0$, and indetermined with respect to k iff $k'(\neg\alpha) = k'(\alpha) = 0$. The rich structure of an OCF allows a notion of *plausibility* to be defined. The degree of *firmness* with which a sentence α is *accepted* in k is measured by $k'(\neg\alpha)$; the higher the value, the more firmly believed, or the more plausible, the sentence is. The degree of firmness with which a sentence α is *rejected* in k is measured by $k'(\alpha)$; the higher the value, the more firmly disbelieved, or the less plausible, the sentence is. If α and β are sentences of L, then $\beta \sqsubset_k \alpha$, will be taken as a shorthand for 'α is more plausible than β' relative to k. The plausibility ordering $\sqsubset_k$ on L (relative to k) is defined as follows:

Definition 4.14 *Let k be an OCF and let $\alpha, \beta \in L$. Then the* ***plausibility ordering*** *$\sqsubset_k$ on L (relative to k) is defined as $\beta \sqsubset_k \alpha$ iff $k'(\neg\alpha) > k'(\neg\beta)$ or $k'(\alpha) < k'(\beta)$.*

The plausibility ordering relative to an OCF is very similar to the plausibility ordering relative to a t-ordering when the OCF and t-ordering are compatible. An OCF k and a t-ordering t are compatible if the underlying ordered partitions are identical, in other words, if the total preorder $\preceq_k$ induced by k and the total preorder $\preceq_t$ induced by t are the same. But in that case, it follows that for every $s, s' \in S$, $k(s) \leq k(s')$ iff $t(s) \leq t(s')$.

Proposition 4.17 *Let $t \in T_S$ be a t-ordering and k an OCF such that t and k are compatible. Then for every $\alpha, \beta \in L$, if $\beta \sqsubset_P \alpha$ then $\beta \sqsubset_k \alpha$ and if $\beta \sqsubset_k \alpha$ then $\beta \sqsubseteq_P \alpha$.*

Proof. Suppose that $\beta \sqsubset_P \alpha$. So $pl(t, \beta) < pl(t, \alpha)$. Let $i = pl(t, \beta)$. So i is the greatest level in t such that $get_{\uparrow}(t, i) \subseteq Mod(\beta)$. But then $\min\{t(s) \mid s \in Mod(\neg\beta)\} = i + 1$. Let $j = pl(t, \alpha)$. So j is the greatest level in t such that $get_{\uparrow}(t, j) \subseteq Mod(\alpha)$.

But then $\min\{t(s) \mid s \in Mod(\neg\alpha)\} = j + 1$. But $i < j$ and thus $\min\{t(s) \mid s \in Mod(\neg\beta)\} < \min\{t(s) \mid s \in Mod(\neg\alpha)\}$. Since t and k are compatible it follows that $\min\{k(s) \mid s \in Mod(\neg\beta)\} < \min\{k(s) \mid s \in Mod(\neg\alpha)\}$. But then $k'(\neg\beta) < k'(\neg\alpha)$, i.e. $\beta \sqsubset_k \alpha$.

Conversely, suppose that $\beta \sqsubset_k \alpha$. So $k'(\neg\alpha) > k'(\neg\beta)$ or $k'(\alpha) < k'(\beta)$. Suppose that $k'(\neg\alpha) > k'(\neg\beta)$. So $\min\{k(s) \mid s \in Mod(\neg\alpha)\} > \min\{k(s) \mid s \in Mod(\neg\beta)\}$. Since t and k are compatible it follows that $\min\{t(s) \mid s \in Mod(\neg\alpha)\} > \min\{t(s) \mid s \in Mod(\neg\beta)\}$. Let $\min\{t(s) \mid s \in Mod(\neg\alpha)\} = j$ and $\min\{t(s) \mid s \in Mod(\neg\beta)\} = i$. But then $get_\uparrow(t, j-1) \subseteq Mod(\alpha)$ and $get_\uparrow(t, i-1) \subseteq Mod(\beta)$, i.e. $pl(t,\alpha) = j - 1$ and $pl(t, \beta) = i - 1$. Since $j > i$ it follows that $pl(t, \alpha) > pl(t, \beta)$, i.e. $\beta \sqsubset_P \alpha$. Suppose that $k'(\alpha) < k'(\beta)$. So $\min\{k(s) \mid s \in Mod(\alpha)\} < \min\{k(s) \mid s \in Mod(\beta)\}$. Since t and k are compatible it follows that $\min\{t(s) \mid s \in Mod(\alpha)\} < \min\{t(s) \mid s \in Mod(\beta)\}$. Let $\min\{t(s) \mid s \in Mod(\alpha)\} = i$ and $\min\{t(s) \mid s \in Mod(\beta)\} = j$. So i is the least level in t such that $get_\uparrow(t, i) \cap Mod(\alpha) \neq \varnothing$ and j is the least level in t such that $get_\uparrow(t, j) \cap Mod(\beta) \neq \varnothing$. But then, since $i < j$, it follows that $dt(t, \alpha) < dt(t, \beta)$. Hence, by proposition 3.36(5), $pl(t, \alpha) \geq pl(t, \beta)$, i.e. $\beta \sqsubseteq_P \alpha$. ■

The main advantage of using t-orderings as a representation of epistemic states is that t-orderings are purely qualitative and allow for the representation of both knowledge and belief. Ordinal conditional functions, on the other hand, rely on the arithmetic of ordinals. This allows for more flexibility in the representation of epistemic states but, at the same time, violates the principle of Qualitativeness. The representation of knowledge is not supported by OCFs, largely because the information-theoretic notion of definitely excluded states is not an inherent part of OCFs.

Chapter 5

Revision and update

It is a capital mistake to theorize before one has data.
Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.
Sherlock Holmes in *The Adventures of Sherlock Holmes* (Sir Arthur Conan Doyle)

5.1 AGM belief revision

The AGM approach to belief revision is one of the most influential contributions to the theory of belief change and has become a standard against which to compare belief change operations. We shall begin this chapter by reviewing the AGM approach as well as several families of closely related approaches. In the latter sections of the chapter we shall provide a logical reconstruction of the ideas using t-orderings. The AGM approach is named after its three authors, Carlos Alchourrón, Peter Gärdenfors, and David Makinson after the publication of their seminal paper (Alchourrón, Gärdenfors, and Makinson, 1985). The approach had its origins in the philosophy of science (Gärdenfors, 1978, 1982, 1984) and the philosophy of law (Alchourrón and Makinson, 1981, 1982) and was strongly influenced by the earlier work of philosophers Harper (1976, 1977) and Levi (1977, 1980).

A historical perspective on the AGM approach may be found in Makinson (2003a).

In the AGM approach, beliefs are objective and represented by sentences of a propositional language under a traditional truth-value semantics. The language is equipped with an abstract (syntactic) consequence relation $\vdash$ in terms of which a consequence operation Cn is defined, which is assumed to be classical (i.e. reflexive, idempotent, and monotonic), to satisfy the deduction theorem and to be compact. In the context of diagrammable systems, the syntactic consequence relation $\vdash$ is replaced by a corresponding semantic consequence relation $\models_{M_0}$ on L under the classical interpretation $M_0 = \langle S, l \rangle$.

The epistemic state of an agent is taken in the AGM approach to be a belief set, or theory (Gärdenfors, 1988). In other words, the epistemic state of an agent is taken to be a set K of sentences of L such that $K = Cn(K)$. There is only one unsatisfiable belief set, the set of all sentences, which is denoted by $K_\perp$. If K is a satisfiable belief set, then a sentence α is said to be *accepted* in K iff $\alpha \in K$, *rejected* in K iff $\neg\alpha \in K$, and *indetermined* with respect to K iff $\alpha \notin K$ and $\neg\alpha \notin K$.

Three basic types of belief change operations are identified, namely, expansion, revision, and contraction.

- *Expansion* consists of adding new information, together with the logical consequences of the information, to the belief set without retracting any of the existing beliefs in the belief set. The resulting belief set may be unsatisfiable.
- *Revision* consists of adding new information, which is inconsistent with the belief set, to the belief set in such a way that the resulting belief set is satisfiable (unless, of course, the new information itself is unsatisfiable).
- *Contraction* consists of retracting information from the belief set, without adding any new information.

Of the three belief change operations, only expansion is defined in a unique way.

Definition 5.1 *Let K be a belief set of L. An* ***expansion operation*** $+$ *is an operation such that for every sentence $\alpha \in L$, the set $K + \alpha$, which is the result of expanding K by α, is defined as $K + \alpha = Cn(K \cup \{\alpha\})$.*

For revision and contraction, a number of *rationality postulates* are defined that every revision and contraction operation is expected to comply with. There are eight AGM postulates defined for revision.

Definition 5.2 *Let K be a belief set of L. A* ***revision operation*** $*$ *is an operation such that for every sentence $\alpha \in L$, the set $K * \alpha$, which is the result of revising K by α, satisfies the following set of postulates:*

($\mathbf{K * 1}$) $K * \alpha = Cn(K * \alpha)$

($\mathbf{K * 2}$) $\alpha \in K * \alpha$

($\mathbf{K * 3}$) $K * \alpha \subseteq K + \alpha$

($\mathbf{K * 4}$) *If* $\neg\alpha \notin K$, *then* $K + \alpha \subseteq K * \alpha$

($\mathbf{K * 5}$) $K * \alpha = K_{\perp}$ *iff* α *is a contradiction*

($\mathbf{K * 6}$) *If* $\alpha \equiv \beta$, *then* $K * \alpha = K * \beta$

($\mathbf{K * 7}$) $K * (\alpha \wedge \beta) \subseteq (K * \alpha) + \beta$

($\mathbf{K * 8}$) *If* $\neg\beta \notin K * \alpha$, *then* $(K * \alpha) + \beta \subseteq K * (\alpha \wedge \beta)$

Postulate ($\mathbf{K * 1}$) requires the result of revision to be a belief set while postulate ($\mathbf{K * 2}$) guarantees that the input sentence is accepted in the resulting belief set. Postulates ($\mathbf{K * 3}$) and ($\mathbf{K * 4}$) describe the relationship between revision and expansion; the belief set resulting from revision is a subset of the belief set resulting from expansion, given the same input sentence, but if the input sentence is consistent with the belief set, then revision coincides with expansion. Postulate ($\mathbf{K * 5}$) ensures that the belief set resulting from revision is satisfiable, unless the input sentence is a contradiction. To ensure that the resulting belief set is satisfiable, some of the existing beliefs may have to be given up. Postulate ($\mathbf{K} * \mathbf{6}$) says that revision by equivalent input sentences should result in equivalent belief sets, implying that belief revision should be analysed at the semantic level and not on the syntactic level. Postulates ($\mathbf{K} * \mathbf{7}$) and ($\mathbf{K} * \mathbf{8}$) together state that if a revised belief set is to be changed by a further input sentence, it should be done by expansion provided the additional input sentence is consistent with the already revised belief set. From postulates ($\mathbf{K} * \mathbf{2}$) and ($\mathbf{K} * \mathbf{3}$), together with postulates ($\mathbf{K} * \mathbf{7}$) and ($\mathbf{K} * \mathbf{8}$), the following property can be derived (Freund and Lehmann, 1994)

($\mathbf{K} * \mathbf{9}$) $(K * \alpha) * \beta = K * (\alpha \wedge \beta)$ if $\neg\beta \notin K * \alpha$.

Property ($\mathbf{K} * \mathbf{9}$) imposes some restrictions on iterated revision requiring that the belief set resulting from a revision by (the conjunction of) two input sentences be the same as the belief set resulting from a revision by one of the input sentences followed by a second revision by the other input sentence (provided the second input sentence is consistent with the already revised belief set). Iterated revision is considered in more detail in section 5.3. The first six postulates are referred to as the basic set of postulates for revision and the last two postulates as the supplementary set of postulates.

For contraction, eight AGM postulates are defined.

Definition 5.3 *Let K be a belief set of L. A **contraction operation** $-$ is an operation such that for every sentence $\alpha \in L$, the set $K - \alpha$, which is the result of contracting K with α, satisfies the following set of postulates:*

($\mathbf{K} - \mathbf{1}$) $K - \alpha = Cn(K - \alpha)$

($\mathbf{K} - \mathbf{2}$) $K - \alpha \subseteq K$

($\mathbf{K} - \mathbf{3}$) *If* $\alpha \notin K$, *then* $K - \alpha = K$

$(\mathbf{K-4})$ *If α is not a tautology, then $\alpha \notin K - \alpha$*

$(\mathbf{K-5})$ *If $\alpha \in K$, then $K \subseteq (K - \alpha) + \alpha$*

$(\mathbf{K-6})$ *If $\alpha \equiv \beta$, then $K - \alpha = K - \beta$*

$(\mathbf{K-7})$ $(K - \alpha) \cap (K - \beta) \subseteq K - (\alpha \wedge \beta)$

$(\mathbf{K-8})$ *If $\alpha \notin K - (\alpha \wedge \beta)$, then $K - (\alpha \wedge \beta) \subseteq K - \alpha$*

Postulate $(\mathbf{K-1})$ requires the result of contraction to be a belief set while postulate $(\mathbf{K-2})$ guarantees that contraction reduces the belief set. Postulate $(\mathbf{K-3})$ ensures that nothing is removed from the belief set unnecessarily while postulate $(\mathbf{K-4})$ ensures that the input sentence is removed from the belief set, unless it is a tautology. Postulate $(\mathbf{K-5})$ is known as the *Recovery postulate* and is the most controversial postulate (Makinson, 1987, 1997b). It requires contraction to be recoverable, that is, the original belief set should be recovered when the contracted belief set is expanded by the initial input sentence. Postulate $(\mathbf{K-6})$ requires input sentences which are semantically equivalent to lead to identical contractions, implying that contraction, similarly to revision, should be analysed at the semantic level and not on the syntactic level. Postulates $(\mathbf{K-7})$ and $(\mathbf{K-8})$ are less intuitive but may be interpreted as saying that contracting with $\alpha \wedge \beta$ must not remove more beliefs than when contracting with α and with β separately. Again, the first six postulates are referred to as the basic set of postulates (for contraction) and the last two postulates as the supplementary set of postulates.

Although the postulates for revision and contraction are independent in the sense that neither refer to the other, revision and contraction can be defined in terms of each via the Harper and Levi identities, as shown by Gärdenfors (1988).

- $K * \alpha = (K - \neg\alpha) + \alpha$ (Levi identity)
- $K - \alpha = (K * \neg\alpha) \cap K$ (Harper identity)

Theorem 5.1 (Gärdenfors, 1988) *If a contraction operation $-$ satisfies* $(\mathbf{K}-\mathbf{1})$ *to* $(\mathbf{K}-\mathbf{8})$*, then the revision operation* $*$ *defined via the* ***Levi identity*** *satisfies* $(\mathbf{K}*\mathbf{1})$ *to* $(\mathbf{K}*\mathbf{8})$*.*

Theorem 5.2 (Gärdenfors, 1988) *If a revision operation* $*$ *satisfies* $(\mathbf{K}*\mathbf{1})$ *to* $(\mathbf{K}*\mathbf{8})$*, then the contraction operation* $-$ *defined via the* ***Harper identity*** *satisfies* $(\mathbf{K}-\mathbf{1})$ *to* $(\mathbf{K}-\mathbf{8})$*.*

As mentioned earlier, the postulates for revision and contraction do not determine unique revision and contraction operations. *Partial meet contraction* is a specific construction method for contraction (Alchourrón, Gärdenfors, and Makinson, 1985) that is based on the intersection of a selected subfamily of the family of all maximal subsets of a belief set K not entailing the sentence α to be contracted. The method of *safe contraction*, due to Alchourrón and Makinson (1985), is another way of constructing contraction operations, based on the dual notion of minimal subsets of a belief set K that entail the sentence α to be contracted. The concept of *epistemic entrenchment*, which was encountered in section 4.5, provides an alternative construction method for contraction. In all of these classical syntactic methods, revision operations are obtained only indirectly via the Levi identity. It is not quite clear how often pure contractions occur in real life (Rott, 2001, pp.107) and as a result, in the context of diagrammable systems, our focus will be on revision.

The first direct semantic method for constructing revision operations is due to Grove (1988), who uses a system of 'spheres' that is similar to the sphere semantics for counterfactuals proposed by Lewis (1973). Grove's system of spheres is based on the notion of maximal satisfiable subsets of an underlying propositional language. These subsets may be viewed as possible worlds. Let $\mathbf{M}$ denote the set of all maximal satisfiable subsets of L. Every belief set K can be represented by the set $[K] \subseteq \mathbf{M}$ of maximal satisfiable subsets of K, given by $[K] = \{\Delta \in \mathbf{M} \mid K \subseteq \Delta\}$. (The same definition may be used to construct the set $[\alpha]$ for any sentence $\alpha \in L$.) When viewed as possible worlds, $[K] = Mod(K)$ (under the classical interpretation $M_0 = \langle S, l \rangle$). Conversely, for every subset $X \subseteq \mathbf{M}$, the set of sentences $th(X) = \{\alpha \in L \mid \alpha \in \cap\{\Delta \in X\}\}$ is a belief set. When viewed as possible worlds, the set $th(X)$ is the theory determined by the corresponding subset $X' \subseteq S$, i.e. $th(X) = Th(X')$.

Definition 5.4 *Let K be a belief set of L. A* ***system of spheres,*** *centered on $[K]$, is a collection $\mathcal{S}$ of subsets of $\mathbf{M}$ that satisfies the following conditions:*

(**S1**) *$\mathcal{S}$ is totally ordered by $\subseteq$*

(**S2**) *$[K]$ is the $\subseteq$-minimum of $\mathcal{S}$*

(**S3**) $\mathbf{M} \in \mathcal{S}$

(**S4**) *If any element of $\mathcal{S}$ intersects $[\alpha]$, then there is a smallest element of $\mathcal{S}$ intersecting $[\alpha]$*

For any sentence $\alpha \in L$, if $[\alpha]$ intersects any element of $\mathcal{S}$, then by condition (**S4**) there is a smallest element of $\mathcal{S}$ intersecting $[\alpha]$, say $c(\alpha)$. If $[\alpha]$ does not intersects any element of $\mathcal{S}$, then $c(\alpha)$ is taken to be $\mathbf{M}$, since, by condition (**S3**), it must be the case that $[\alpha] = \varnothing$. The 'closest' elements in $\mathbf{M}$ to $[K]$ in which α is an element can now be defined as the set $[\alpha] \cap c(\alpha)$. The key idea in Grove's approach to belief revision is that the revision of a belief set K by a sentence α can be represented by (the belief set determined by) the subset $[\alpha] \cap c(\alpha) \subseteq \mathbf{M}$ of 'worlds' closest to $[K]$.

Theorem 5.3 (Grove, 1988) *Let K be a belief set of L. A revision operation $*$ satisfies* ($\mathbf{K * 1}$) *to* ($\mathbf{K * 8}$) *iff there is some system of spheres $\mathcal{S}$ in $\mathbf{M}$, which is centered on $[K]$, such that for all $\alpha \in L$, $K * \alpha = th([\alpha] \cap c(\alpha))$.*

The result shows that revision can be characterised in terms of systems of spheres. As an alternative, Grove provides a characterisation of revision operations in terms of an ordering on the sentences of L (with respect to a belief set K) that satisfies a number of conditions.

Definition 5.5 *Let K be a belief set of L and let $\alpha, \beta \in L$. A binary relation $\sqsubseteq_G$ on L is a **G-ordering** (with respect to K) iff it satisfies the following conditions:*

- (**G1**) $\sqsubseteq_G$ *is connected*
- (**G2**) $\sqsubseteq_G$ *is transitive*
- (**G3**) *If* $\models \alpha \rightarrow \beta \vee \gamma$, *then either* $\beta \sqsubseteq_G \alpha$ *or* $\gamma \sqsubseteq_G \alpha$
- (**G4**) α *is* $\sqsubseteq_G$*-minimal iff* $\neg\alpha \notin K$
- (**G5**) α *is* $\sqsubseteq_G$*-maximal iff* α *is a contradiction*

The G-orderings can be induced by a system of spheres centered on $[K]$ by taking, for all $\alpha, \beta \in L$, $\alpha \sqsubseteq_G \beta$ iff $c(\alpha) \subseteq c(\beta)$ and hence $\alpha \sqsubset_G \beta$ iff $c(\alpha) \subset c(\beta)$. Grove's representation theorem for belief revision in terms of G-orderings is based on the idea that the revision of a belief set K by a sentence α can be represented by the set of all sentences $\beta \in L$ such that $(\alpha \wedge \beta) \sqsubset_G (\alpha \wedge \neg\beta)$.

Theorem 5.4 (Grove, 1988) *Let K be a belief set of L. A revision operation $*$ satisfies* $(\mathbf{K} * \mathbf{1})$ *to* $(\mathbf{K} * \mathbf{8})$ *iff there is some G-ordering $\sqsubseteq_G$ on L with respect to K, such that for all $\alpha \in L$, $K * \alpha = \{\beta \in L \mid (\alpha \wedge \beta) \sqsubset_G (\alpha \wedge \neg\beta)\}$.*

G-orderings can be used to construct EE-orderings, and vice versa (Gärdenfors, 1988, pp.95). The connection between the two orderings rests on the observation that an EE-ordering is used to determine which sentences should be retained in $K - \alpha$ whereas a G-ordering is used to determine which sentences should be included in $K * \alpha$. The intuitive idea in constructing an EE-ordering from a G-ordering is that, if $\beta \in K$ then, to determine whether β should be in $K - \alpha$, it suffices to (apply the Harper identify and) consider whether $\neg\beta$ is in $K * \alpha$, as characterised by some G-ordering (with respect to K).

Theorem 5.5 (Gärdenfors, 1988) *Let K be a belief set of L. A binary relation $\sqsubseteq$ on L is an EE-ordering (with respect to K) iff it can be defined in terms of a G-ordering $\sqsubseteq_G$ on L (with respect to K) using the definition, for all $\alpha, \beta \in L$, $\alpha \sqsubseteq \beta$ iff $\neg\alpha \sqsubseteq_G \neg\beta$.*

However, from every EE-ordering on L (with respect to some belief set K), a faithful total preorder on S (with respect to K) can be constructed as shown by theorem 4.1 in section 4.5. From the relationship between EE-orderings and G-orderings, the relationship between G-orderings and faithful total preorders on the set of states can be defined, again, with the aid of a suitable power construction (Meyer, 1999).

Definition 5.6 *Let K be a belief set of L. Suppose $\preceq$ is a faithful total preorder on S (with respect to K). Then the* ***power order*** *$\sqsubseteq_{GP}$ on L induced by $\preceq$ is defined, for all $\alpha, \beta \in L$, by $\alpha \sqsubseteq_{GP} \beta$ iff for every $s' \in Mod(\beta)$ there is some $s \in Mod(\alpha)$ such that $s \preceq s'$.*

Theorem 5.6 (Meyer, 1999) *Let K be a belief set of L. A binary relation $\sqsubseteq_G$ on L is a G-ordering (with respect to K) iff it is the power order $\sqsubseteq_{GP}$ on L induced by some faithful total preorder on S (with respect to K).*

The first semantic characterisation of revision operations in terms of faithful total preorders if due to Katsuno and Mendelzon (1989, 1991). Their approach, which is considered in section 5.2, uses a finitely generated propositional language under a traditional truth-value semantics, which allows a belief set K to be represented by a single sentence κ such that $K = Cn(\kappa)$. The key idea in their approach is that the revision of a belief set K by a sentence α can be represented by (the theory determined by) the models of α that are 'closest' to the models of K, i.e. the models of α that are minimal with respect to the faithful total preorder assigned to K. The following representation theorem is an adaptation of the characterisation of Katsuno and Mendelzon in terms of belief sets.

Theorem 5.7 *Let K be a belief set of L. A revision operation $*$ satisfies* $(\mathbf{K} * \mathbf{1})$ *to* $(\mathbf{K} * \mathbf{8})$ *iff there is some total preorder $\preceq$ on S, which is faithful with respect to K, such that for all $\alpha \in L$, $K * \alpha = Th(Min_{\preceq}(\alpha))$.*

The representation theorems presented have shown that a revision operation can be characterised in terms of a family of systems of spheres in $\mathbf{M}$, one for each belief set K; in terms of a family of G-orderings on L, one for each belief set K; and in terms of a family of faithful total preorders on S, one for each belief set K. All of these approaches describe how to produce, from a belief set and some kind of ordering a new belief set for any sentence, but give no indication as to what the new ordering should be. As pointed out by Friedman and Halpern (1999a), the epistemic state here is function not only of a belief set, but of an ordering too, and it is the ordering that determines how revision is performed and not the belief set, since there are many orderings for which the associated belief set would be the same. Since the revision process results in a new belief set only, and not a revised ordering as well, the representation does not fully support iterated revision. Iterated revision, the problem of dealing with a succession of changes to the epistemic state of an agent, is generally regarded as a fundamental limitation of the AGM approach to belief revision. This matter will be addressed in section 5.3. Another limitation of the AGM approach to belief revision is that it only deals with new information about a 'static' world, a matter that is taken up next.

5.2 Knowledge base update

Katsuno and Mendelzon (1992) make a fundamental distinction between the *revision* of a knowledge base and the *update* of a knowledge base. Revision is deemed appropriate when new information is obtained about a 'static' world (a world which persists in the same state) whereas update is deemed appropriate when the new information is about changes in a 'dynamic' world (a world which changes its state). The distinction was originally made by Keller and Winslett (1985) in the context of extended relational databases. The following example, due to Winslett (1988), illustrates that revision is not appropriate in cases involving a change of state.

Example 5.1 (Winslett, 1988) *Suppose that all we know in K about a particular room is that there is a table, a book and a magazine in it, and that either (β) the book is on the table, or (α) the magazine is on the table, but not both, i.e. the belief set K is essentially $Cn((\beta \wedge \neg\alpha) \vee (\alpha \wedge \neg\beta))$. A robot is then ordered to put the book on the table, and as a consequence, we learn that β. If we change our beliefs by revision we should, according to* ($\mathbf{K} * \mathbf{4}$) *end up with a belief set that contains ($\beta \wedge \neg\alpha$) since β is consistent with K. But why should we conclude that the magazine is not on the table?*

In the context of belief change, revision is appropriate when the agent obtains new information but the system has persisted in the same state whereas update is appropriate when the system has changed state.

In the approach of Katsuno and Mendelzon, beliefs are objective and represented by sentences of a finitely generated propositional language under a traditional truth-value semantics. The finiteness of the language allows the epistemic state of an agent to be taken as a single sentence κ, referred to as a *knowledge base.* Note that the term 'knowledge base' is not interchangeable with the term 'belief base'. A *belief base,* as a model of epistemic states, uses an arbitrary set of sentences as opposed to a set of sentences closed under semantic consequence (i.e. a belief set). It is motivated by the argument that some of the agent's beliefs have no independent standing but arise merely as inferences from the agent's more 'basic' beliefs (Fuhrmann, 1991; Hansson, 1992; Nebel, 1989, 1992). A belief base B is taken to consist of such basic beliefs and is said to be a *base* for a belief set K iff $Cn(B) = K$. Belief revision based on this model of epistemic states is known as *base revision* (Gärdenfors and Rott, 1995; Hansson, 1998, 1999a; Nebel, 1998).

For every knowledge base κ, a belief set K can be constructed by taking $K = Cn(\kappa)$ (under the classical interpretation $M_0 = \langle S, l \rangle$). The knowledge base κ is a finite axiomatisation of the models of the belief set K. If κ is a satisfiable knowledge base, then a sentence α is said to be accepted in κ iff $Mod(\kappa) \subseteq Mod(\alpha)$, rejected in κ iff $Mod(\kappa) \cap Mod(\alpha) = \varnothing$, and indetermined with respect to κ iff $Mod(\kappa) \not\subseteq Mod(\alpha)$ but $Mod(\kappa) \cap Mod(\alpha) \neq \varnothing$.

Definition 5.7 *Let κ be a knowledge base of L. An* ***expansion operation*** $+$ *is an operation such that for every sentence $\alpha \in L$, the knowledge base $\kappa + \alpha$, which is the result of expanding κ by α, is defined as $\kappa + \alpha \equiv \kappa \wedge \alpha$.*

There are six KM postulates defined for revision.

Definition 5.8 *Let κ be a knowledge base of L. A* ***revision operation*** $*$ *is an operation such that for every sentence $\alpha \in L$, the knowledge base $\kappa * \alpha$, which is the result of revising κ by α, satisfies the following set of postulates:*

(**KM** $*$ **1**) $\kappa * \alpha \models \alpha$

(**KM** $*$ **2**) *If $\kappa \wedge \alpha$ is satisfiable, then $\kappa * \alpha \equiv \kappa \wedge \alpha$*

(**KM** $*$ **3**) *If α is satisfiable, then $\kappa * \alpha$ is satisfiable*

(**KM** $*$ **4**) *If $\kappa \equiv \kappa'$ and $\alpha \equiv \alpha'$ then $\kappa * \alpha \equiv \kappa' * \alpha'$*

(KM $*$ 5) $(\kappa * \alpha) \wedge \beta \models \kappa * (\alpha \wedge \beta)$

(KM $*$ 6) *If* $(\kappa * \alpha) \wedge \beta$ *is satisfiable, then* $\kappa * (\alpha \wedge \beta) \models (\kappa * \alpha) \wedge \beta$

The KM postulates for revision of knowledge bases correspond directly to the AGM postulates for revision of belief sets when a belief set K is taken to be the set of consequences of a knowledge base κ, i.e. $K = Cn(\kappa)$. A belief set revision operation $*$ satisfies (**K** $*$ **1**) to (**K** $*$ **6**) iff the corresponding knowledge base revision operation $*$ satisfies (**KM** $*$ **1**) to (**KM** $*$ **4**), while (**K** $*$ **7**) and (**K** $*$ **8**) respectively correspond to (**KM** $*$ **5**) and (**KM** $*$ **6**).

As mentioned earlier, Katsuno and Mendelzon were the first to provide a semantic characterisation of revision operations in terms of (faithful) total preorders. A *faithful assignment* maps every knowledge base κ to a total pre-order $\preceq_\kappa$ on the set of worlds in such a way that the following conditions hold:

- $\preceq_\kappa$ is faithful with respect to $Cn(\kappa)$, and
- if $\kappa \equiv \kappa'$, then $\preceq_\kappa = \preceq_{\kappa'}$.

The faithful assignment ensures that the models of every knowledge base κ are strictly below its nonmodels in the total preorder assigned to κ.

Theorem 5.8 (Katsuno and Mendelzon, 1991) *A revision operation* $*$ *satisfies postulates* (**KM** $*$ **1**) *to* (**KM** $*$ **6**) *iff there exists a faithful assignment that maps each knowledge base* κ *to a total preorder* $\preceq_\kappa$ *on* S*, such that for all* $\alpha \in L$*,* $Mod(\kappa * \alpha) = Min_{\preceq_\kappa}(\alpha)$*.*

Turning the attention to update, Katsuno and Mendelzon (1992) provide two sets of postulates for update; one characterising update operations in terms of partial orders and the other characterising update operations in terms of total preorders. Our focus will primarily be on the latter, for which seven postulates are defined.[1]

[1] In the alternative formulation, postulate (**KM** $\diamond$ **6**) is replaced by postulates (**KM** $\diamond$ **8**): If $(\kappa \diamond \alpha) \models \beta$ and $(\kappa \diamond \beta) \models \alpha$, then $(\kappa \diamond \alpha) \equiv (\kappa \diamond \beta)$ and (**KM** $\diamond$ **9**): If κ is complete, then $(\kappa \diamond \alpha) \wedge (\kappa \diamond \beta) \models \kappa \diamond (\alpha \wedge \beta)$.

Definition 5.9 *Let κ be a knowledge base. An* **update operation** $\diamond$ *is an operation such that for every sentence $\alpha \in L$, the knowledge base $\kappa \diamond \alpha$, which is the result of updating κ with α, satisfies the following set of postulates:*

(**KM** $\diamond$ **1**) $\kappa \diamond \alpha \models \alpha$

(**KM** $\diamond$ **2**) *If $\kappa \models \alpha$, then $\kappa \diamond \alpha \equiv \kappa$*

(**KM** $\diamond$ **3**) *If both κ and α are satisfiable, then $\kappa \diamond \alpha$ is satisfiable*

(**KM** $\diamond$ **4**) *If $\kappa \equiv \kappa'$ and $\alpha \equiv \alpha'$, then $\kappa \diamond \alpha \equiv \kappa' \diamond \alpha'$*

(**KM** $\diamond$ **5**) $(\kappa \diamond \alpha) \wedge \beta \models \kappa \diamond (\alpha \wedge \beta)$

(**KM** $\diamond$ **6**) *If $(\kappa \diamond \alpha) \wedge \beta$ is satisfiable and κ is complete*[2]*, then $\kappa \diamond (\alpha \wedge \beta) \models (\kappa \diamond \alpha) \wedge \beta$*

(**KM** $\diamond$ **7**) $(\kappa \vee \kappa') \diamond \alpha \equiv (\kappa \diamond \alpha) \vee (\kappa' \diamond \alpha)$

Postulates (**KM** $\diamond$ **1**), (**KM** $\diamond$ **4**), and (**KM** $\diamond$ **5**) for update correspond directly to postulates (**KM** $*$ **1**), (**KM** $*$ **4**), and (**KM** $*$ **5**) for revision respectively. Postulate (**KM** $\diamond$ **2**) for update differs from its counterpart, postulate (**KM** $*$ **2**), for revision and asserts merely that if the input sentence α is a semantic consequence of κ, then updating by α does not influence the knowledge base κ. From postulate (**KM** $\diamond$ **3**) it follows that if the knowledge base is unsatisfiable, then updating it by a satisfiable input sentence α does not necessarily result in a satisfiable knowledge base, in contrast to revising it with α, as stated by postulate (**KM** $*$ **3**). Postulate (**KM** $\diamond$ **6**) corresponds to postulate (**KM** $*$ **6**) for revision but only applies to complete knowledge bases. Postulate (**KM** $\diamond$ **7**) is referred to as the *Disjunction Rule* and requires that the result of updating the disjunction of two knowledge bases with an input sentence be equivalent to the disjunction of updating each knowledge base with the input sentence.

[2] *A sentence α is said to be complete if for any sentence β, α entails β or α entails $\neg\beta$ (Katsuno and Mendelson, 1992).*

The characterisation of update operations (for knowledge bases) is based on the idea that the update of a knowledge base κ with a sentence α can be represented by selecting, for each model s of κ, the models of α that are 'closest' to s. (We shall have more to say about this idea in due course.) The intuitive notion of 'closeness' is captured by a function that assigns a total preorder to each world $s \in S$. Formally, a *faithful update assignment* maps every world $s \in S$ to a total preorder $\preceq_s$ on S in such a way that the following condition holds:

- for any $s' \in S$, if $s \neq s'$ then $s \prec_s s'$.

Using the notion of a faithful update assignment, Katsuno and Mendelzon provide a characterisation of update operations (for knowledge bases).

Theorem 5.9 (Katsuno and Mendelzon, 1992) *An update operation* $\diamond$ *satisfies postulates* $(\mathbf{KM} \diamond \mathbf{1})$ *to* $(\mathbf{KM} \diamond \mathbf{7})$ *iff there exists a faithful update assignment that maps each world* s *to a total preorder* $\preceq_s$ *on* S*, such that for all* $\alpha \in L$*,* $Mod(\kappa \diamond \alpha) = \bigcup_{s \in Mod(\kappa)} Min_{\preceq_s}(\alpha)$.

The class of update operations characterised in this way may be seen as a generalisation of the *possible models approach* of Winslett (1988, 1990). In the possible models approach, or PMA for short, a partial order $\leq_s$ is associated with every world $s \in S$ in such a way that $s' \leq_s s''$ iff $Dist(s, s') \subseteq Dist(s, s'')$ where $Dist(s, s')$ is the set of atoms that have different truth values under s and s'[3]. Intuitively, $s' \leq_s s''$ means that s' must be 'closer' to s than s'', since s and s' differ on fewer valuations than s and s''. The PMA update operation selects for each model s in a fixed knowledge base κ, the models of α that are 'closest' to s, in other words, the minimal models of α with respect to the partial

[3] If the new input sentence α is inconsistent with the knowledge base κ, then Winslett's PMA update operation coincides with the update operation of Borgida (1985, 1988).

order $\leq_s$ on S. The models of the updated knowledge base are then the union of these selected models. The PMA update operation satisfies postulates (**KM** $\diamond$ **1**) to (**KM** $\diamond$ **5**) and postulates (**KM** $\diamond$ **7**) to (**KM** $\diamond$ **9**). As such, the PMA update operations constitute a sub-class of the class of update operations characterised by Katsuno and Mendelzon in terms of a family of partial orders, one for each world $s \in S$.

The update operation of Forbus (1989), which is the update counterpart of the revision operation of Dalal (1988), is based on total preorders. In the approach of Forbus, a total preorder $\leq_s$ is associated with every world $s \in S$ in such a way that $s' \leq_s s''$ iff $card(Dist(s, s')) \leq card(Dist(s, s''))$ where $Dist(s, s')$ is, as before, the set of atoms that have different truth values under s and s'. As with the PMA update operation, the Forbus update operation selects for each model s in a fixed knowledge base κ, the minimal models of α with respect to the ordering $\leq_s$, which, in this case, is a total preorder on S. The models of the updated knowledge base are again the union of these selected models. In satisfying postulates (**KM** $\diamond$ **1**) to (**KM** $\diamond$ **7**), the Forbus update operations constitute a sub-class of the class of update operations characterised by Katsuno and Mendelzon in terms of a family of total preorders, again, one for each world $s \in S$.

The update operations of Winslett and Forbus are examples of a family of update operations, called *minimisation-based* updates, in which the distances between possible worlds are minimised. *Dependency-based* updates are another family of update operations in which the distances between possible worlds are constrained to be in some set Dep of exceptions determined by the input sentence α. In dependency-based updates, the point-wide update $s \diamond \alpha$ of models of κ can be defined as $s \diamond \alpha = \{s' \in Mod(\alpha) \mid Dist(s, s') \subseteq Dep(\alpha)\}$. The set $Dep(\alpha)$ can be defined in different ways. The basic idea behind dependence (Herzig, 1996) is that some of the atoms appearing in α should be exempted from change, specifically, those atoms upon which α is not dependent.

An alternative characterisation of dependency-based update operations, which is based on the principle 'forget, then expand', is given by Doherty, Lukaszewicz, and Madalińska-Bugaj (1998). The basic idea is that point-wide update $s \diamond \alpha$ of the models of κ proceeds by first 'forgetting' about atoms in $Dep(\alpha)$ and then 'expanding' the result with α. The notion of *independence* is key. A sentence α is said to be independent from $A \subseteq Atom$ iff there exists a sentence $\alpha' \equiv \alpha$ such that $Atm(\alpha') \cap A = \varnothing$. The sentence $Forget(\alpha, A)$ is the strongest semantic consequence of α that is independent from A. It can be obtained by transforming α into SDNF so that $\alpha = \sigma_1 \vee \sigma_2 \vee \ldots \vee \sigma_n$ and removing from each σ_i all occurrences of ρ_i and $\neg\rho_i$ for every $\rho_i \in A$. Taking the set $Dep(\alpha)$ to be the set of atoms in $Atm(\alpha)$ that α is dependent on, an alternative characterisation of dependency-based updates through the notion of forgetting is given as $\kappa \diamond \alpha \equiv Forget(\kappa, Dep(\alpha)) \wedge \alpha$.

In an examination of ten concrete update operations, all of which are defined in terms of the distance $Dist$ between worlds, Herzig and Rifi (1999) found, under a specified set of hypotheses, that only the update operations of Winslett and of Forbus satisfy the KM postulates for update. Based on the examination, they argue that postulates (**KM** $\diamond$ **2**), (**KM** $\diamond$ **5**) and (**KM** $\diamond$ **8**) are undesirable and that postulate (**KM** $\diamond$ **9**) is without importance. The strongest argument is against postulates (**KM** $\diamond$ **2**) and (**KM** $\diamond$ **5**). According to postulate (**KM** $\diamond$ **2**), an agent leaves unchanged its beliefs whenever the beliefs entail the new information, and when interpreted as saying that the agent prefers to consider the new information as noisy sensing in an unchanging world rather than correct sensing in a changing world, the criticism seems fair. The criticism against postulate (**KM** $\diamond$ **5**) derives from the result (Herzig and Rifi, 1999, Lemma 38) that postulate (**KM** $\diamond$ **5**), together with postulates (**KM** $\diamond$ **1**) and (**KM** $\diamond$ **4**), entail

(**Exor**) If $bel(t_E \diamond \alpha) \models_M \neg\beta$ and $bel(t_E \diamond \beta) \models_M \neg\alpha$ then $bel(t_E \diamond (\alpha \vee \beta) \models_M (\alpha \wedge \neg\beta) \vee (\neg\alpha \wedge \beta)$.

According to the condition (**Exor**), an update by an inclusive disjunction always leads to exclusive disjunction, something which is deemed undesirable[4]. Given that postulates ($\mathbf{KM} \diamond \mathbf{1}$), ($\mathbf{KM} \diamond \mathbf{4}$), and ($\mathbf{KM} \diamond \mathbf{5}$) have identical counterparts in revision, the criticism that disjunctive update should not be identified with exclusive disjunction could be raised against postulate ($\mathbf{KM} * \mathbf{5}$) for revision too, but seldom is.

An important difference between distance-based approaches to update and revision is that in the case of update, an ordering is required for every model of the knowledge base whereas, for revision, a single ordering is required for the knowledge base. Similarly to belief revision, a knowledge base does not uniquely determine the outcome of an update operation, rather, it is the orderings that determine how update is performed. The problem of iterated revision thus applies to update as well.

5.3 Iterated revision

The approach of Darwiche and Pearl (1994, 1997) is arguably the most influential approach to *iterated belief revision* with ideas initially proposed by Spohn (1988, 1990) serving as the main inspiration. In distinguishing between epistemic states and belief sets, they make the transition from revision as a function of belief sets to revision as a function of epistemic states (Darwiche and Pearl, 1997).

In the approach of Darwiche and Pearl, beliefs are objective and represented by sentences of a finitely generated propositional language under a traditional truth-value semantics. Although they do not give an explicit definition (or representation) of epistemic states, an epistemic state is taken to comprise not only a belief set, but all of the information needed for coherent reasoning and, in particular, the very 'strategy' for

[4] The criticims against this interpretation of disjunction are also raised in the context of *base update* (Herzig, 1998).

revising the belief set. In terms of the approaches described earlier, a strategy may be viewed as a kind of ordering, which may be a system of spheres in $\mathbf{M}$, a G-ordering on L, or a faithful total preorder on S, that determines how revision is performed.

Every epistemic state is assumed to have an associated belief set, but the belief set does not uniquely characterise the epistemic state, in other words, different epistemic states may have the same associated belief set. To accommodate the transition from revision as a function of belief sets to revision as a function of epistemic states, a modification of the postulates for revision is required. Darwiche and Pearl propose a modification to the KM postulates for revision, thereby taking a belief set to be represented by a knowledge base.[5]

Definition 5.10 *Let E be an epistemic state and $\kappa(E)$ the knowledge base associated with E. A **revision operation** $*$ is an operation such that for every sentence $\alpha \in L$, the epistemic state $E * \alpha$, which is the result of revising E by α, satisfies the following set of postulates:*

($\mathbf{DP * 1}$) $\kappa(E * \alpha) \models \alpha$

($\mathbf{DP * 2}$) *If $\kappa(E) \wedge \alpha$ is satisfiable, then $\kappa(E * \alpha) \equiv \kappa(E) \wedge \alpha$*

($\mathbf{DP * 3}$) *If α is satisfiable, then $\kappa(E * \alpha)$ is satisfiable*

($\mathbf{DP * 4}$) *If $E = E'$ and $\alpha \equiv \alpha'$, then $\kappa(E * \alpha) \equiv \kappa(E' * \alpha')$*

($\mathbf{DP * 5}$) $\kappa(E * \alpha) \wedge \beta \models \kappa(E * (\alpha \wedge \beta))$

($\mathbf{DP * 6}$) *If $\kappa(E * \alpha) \wedge \beta$ is satisfiable, then $\kappa(E * (\alpha \wedge \beta)) \models \kappa(E * \alpha) \wedge \beta$*

The only difference between these postulates and the KM postulates for revision is the reformulation of postulate ($\mathbf{KM * 4}$) into ($\mathbf{DP * 4}$), which makes belief revision a

[5] A similar modification to the AGM postulates for revision had been proposed independently by Friedman and Halpern (1996).

function of epistemic states rather than belief sets. Postulate ($\mathbf{KM} * \mathbf{4}$) effectively says that if the knowledge bases associated with epistemic states E and E' are equivalent, then revision by equivalent input sentences should result in equivalent knowledge bases. Postulate ($\mathbf{DP} * \mathbf{4}$) requires epistemic states E and E' to be identical for this to be the case. Note however, that postulate ($\mathbf{DP} * \mathbf{4}$) does not require the revised epistemic states to be identical and can therefore not be seen as a formal expression of the principle of Irrelevance of Syntax in the context of revision of epistemic states as is postulate ($\mathbf{KM} * \mathbf{4}$) in the context of revision of knowledge bases. The rationale for reformulating postulate ($\mathbf{KM} * \mathbf{4}$) into ($\mathbf{DP} * \mathbf{4}$) is clearly illustrated by the following example.

Example 5.2 (Darwiche and Pearl, 1997) *Two jurors in a murder trail possess different biases; juror-1 believes 'A is the murderer, B is a remote but unbelievable possibility while C is definitely innocent'. Juror-2 believes 'A is the murderer, C is a remote but unbelievable possibility while B is definitely innocent'. The two jurors share the same belief set* $\kappa(E) \equiv \kappa(E') =$ *'A is the only murderer'. A surprising evidence now obtains:* $\alpha =$ *'A is not the murderer' (A has produced a reliable alibi.) Clearly, any rational account of belief revision should allow juror-1 to uphold a different belief set than juror-2. Yet any approach based on a revision operator that satisfies postulate* ($\mathbf{KM} * \mathbf{4}$) *dictates that* $\kappa(E * \alpha) \equiv \kappa(E' * \alpha)$*, which is an indefensible position.*

Darwiche and Pearl provide a representation result in terms of (faithful) total pre-orders, which parallels theorem 5.8. It requires a *faithful assignment* that maps every epistemic state E to a total pre-order $\preceq_E$ on the set of worlds in such a way that the following conditions hold:

- $\preceq_E$ is faithful with respect to $Cn(\kappa(E))$, and
- if $E = E'$, then $\preceq_E = \preceq_{E'}$.

Theorem 5.10 (Darwiche and Pearl, 1997) *A revision operation $*$ satisfies* $(\mathbf{DP} * \mathbf{1})$ *to* $(\mathbf{DP} * \mathbf{6})$ *iff there exists a faithful assignment that maps each epistemic state E to a total preorder $\preceq_E$ on S, such that for all $\alpha \in L$, $Mod(\kappa(E * \alpha)) = Min_{\preceq_E}(\alpha)$.*

Darwiche and Pearl provide a number of convincing examples showing that a revision operation that satisfies the modified KM postulates $(\mathbf{DP} * \mathbf{1})$ to $(\mathbf{DP} * \mathbf{6})$ can nonetheless lead to counterintuitive forms of iterated revision. This leads them to propose an additional set of postulates, none of which is derivable from the modified KM postulates, to augment the modified KM postulates for revision. The plausibility of these proposed postulates for iterated revision is demonstrated by a number of concrete scenarios.

Definition 5.11 *Let E be an epistemic state and $\kappa(E)$ the knowledge base associated with E. A* ***revision operation*** *$*$ is an operation such that for every sentence $\alpha \in L$, the epistemic state $E * \alpha$, which is the result of revising E by α, satisfies, in addition to postulates* $(\mathbf{DP} * \mathbf{1})$ *to* $(\mathbf{DP} * \mathbf{6})$*, the following set of postulates:*

$(\mathbf{C1})$ *If $\beta \models \alpha$, then $\kappa((E * \alpha) * \beta) \equiv \kappa(E * \beta)$*

$(\mathbf{C2})$ *If $\beta \models \neg\alpha$, then $\kappa((E * \alpha) * \beta) \equiv \kappa(E * \beta)$*

$(\mathbf{C3})$ *If $\kappa(E * \beta) \models \alpha$, then $\kappa((E * \alpha) * \beta) \models \alpha$*

$(\mathbf{C4})$ *If $\kappa(E * \beta) \not\models \neg\alpha$, then $\kappa((E * \alpha) * \beta) \not\models \neg\alpha$*

Postulate $(\mathbf{C1})$ says that if the second input sentence is more specific (i.e. logically stronger) than the first input sentence, then the first input sentence is redundant. If the second input sentence contradicts the first input sentence, then postulate $(\mathbf{C2})$ requires the second (i.e. later) input sentence to prevail. Postulate $(\mathbf{C3})$ requires the first input sentence to be retained after revision by a second input sentence if revising (the original epistemic state) by the second input sentence would have entailed the first input sentence. Postulate $(\mathbf{C4})$ says that no input sentence can contribute to its own demise. In other words, if the first input sentence would not have been contradicted after revising (the original epistemic state) by the second input sentence, then postulate (**C4**) requires the first input sentence not to be contradicted after a subsequent revision by the second input sentence.

Each postulate for iterated revision represents a class of 'conditional beliefs'. The phrase 'conditional belief $\beta|\alpha$' is a shorthand for the phrase 'β will be accepted after revising the current epistemic state by α'. A conditional belief $\beta|\alpha$ is accepted in an epistemic state E precisely when $\kappa(E * \alpha)$ entails β.[6] Darwiche and Pearl show that E accepts the conditional belief $\beta|\alpha$ precisely when there exists a state $s \in S$ such that s satisfies $\alpha \wedge \beta$ and $s \prec_E s'$ for any state s' that satisfies $\alpha \wedge \neg\beta$. The conditional beliefs accepted by an epistemic state E are therefore encoded by the total preorder $\preceq_E$ associated with E and, similarly, the conditional beliefs accepted by $E * \alpha$ are encoded by the total preorder $\preceq_{E*\alpha}$. By making the total preorders $\preceq_E$ and $\preceq_{E*\alpha}$ as similar as (rationally) possible, the changes in conditional beliefs as a result of a revision can be minimised, thus adhering to the principle of Minimal Change. Postulates (**C1**) to (**C4**) constrain the relationship between $\preceq_E$ and $\preceq_{E*\alpha}$ as shown by the following representation theorem.

Theorem 5.11 (Darwiche and Pearl, 1997) *Suppose that a revision operation $*$ satisfies* (**DP $*$ 1**) *to* (**DP $*$ 6**). *Then $*$ satisfies* (**C1**) *to* (**C4**) *iff $*$ and its corresponding faithful assignment satisfy:*

(**CR1**) *If $s, s' \in Mod(\alpha)$, then $s \preceq_E s'$ iff $s \preceq_{E*\alpha} s'$*

(**CR2**) *If $s, s' \in NMod(\alpha)$, then $s \preceq_E s'$ iff $s \preceq_{E*\alpha} s'$*

(**CR3**) *If $s \in Mod(\alpha)$ and $s' \in NMod(\alpha)$, then $s \prec_E s'$ only if $s \prec_{E*\alpha} s'$*

[6] Darwiche and Pearl caution that this definition of conditional beliefs should not be viewed as an interpretation of 'conditionals'.

(**CR4**) *If* $s \in Mod(\alpha)$ *and* $s' \in NMod(\alpha)$, *then* $s \preceq_E s'$ *only if* $s \preceq_{E*\alpha} s'$

The representation theorem indicates how each of postulates (**C1**) to (**C4**) respectively ensures (through conditions (**CR1**) to (**CR4**)) that some part of $\preceq_E$ is preserved into $\preceq_{E*\alpha}$ after revising an epistemic state E by an input sentence α. Condition (**CR1**) ensures that the relative orderings of the models of α are preserved while condition (**CR2**) ensures that the relative orderings of the nonmodels of α are preserved. Conditions (**CR3**) and (**CR4**) together ensure that models of α that are below nonmodels of α will remain so. The preservation of some part of $\preceq_E$ into $\preceq_{E*\alpha}$ represents a form of minimal change, and hence, an application of the principle of Minimal Change.

A more drastic form of minimal change obtains when postulate (**CB**) is added to the AGM postulates for revision because it ensures that $\preceq_E$ is preserved as much as possible into $\preceq_{E*\alpha}$.

(**CB**) If $\kappa(E * \alpha)$ entails $\neg\beta$, then $\kappa((E * \alpha) * \beta) \equiv \kappa(E * \beta)$

Postulate (**CB**) says that accommodating a second input sentence should nullify the first input sentence if revision by the first input sentence would contradict the second input sentence. Note that postulate (**CB**) implies postulates (**C1**) to (**C4**) but that the converse does not hold. The following representation theorem shows how postulate (**CB**) ensures (through condition (**CRB**)) the maximum preservation of $\preceq_E$ into $\preceq_{E*\alpha}$.

Theorem 5.12 (Darwiche and Pearl, 1997) *Suppose that a revision operation* $*$ *satisfies* (**DP** $*$ **1**) *to* (**DP** $*$ **6**). *Then* $*$ *satisfies* (**CB**) *iff* $*$ *and its corresponding faithful assignment satisfy:*

(**CRB**) *If* $s, s' \in NMod(\kappa(E * \alpha))$, *then* $s \preceq_E s'$ *iff* $s \preceq_{E*\alpha} s'$

Condition (**CRB**) ensures that the relative orderings of the states in $NMod(\kappa(E * \alpha))$ are preserved, in other words, it ensures that the relative orderings of all states are preserved, except for the minimal models of α (with respect to $\preceq_E$). However, the faithfulness of $\preceq_{E*\alpha}$, which is required by the modified KM postulates (**DP** $*$ **1**) to (**DP** $*$ **6**), determines the order imposed on the minimal models of α (with respect to $\preceq_E$) in $\preceq_{E*\alpha}$ as follows:

- $s =_{E*\alpha} s'$ if both s and s' are minimal models of α (with respect to $\preceq_E$) and
- $s \prec_{E*\alpha} s'$ if s is a minimal model of α (with respect to $\preceq_E$) but s' is not.

Condition (**CRB**) is effectively Boutilier's definition of *natural revision* (Boutilier, 1993) and *minimal conditional revision* (Boutilier, 1996a). This form of iterated revision ensures absolute minimisation of changes in conditional beliefs but, as shown by the Red-Bird example, can lead to counterintuitive results.

Example 5.3 (Darwiche and Pearl, 1997) *We encounter a strange new animal and it appears to be a bird, so we believe the animal is a bird. As it comes closer to our hiding place, we see clearly that the animal is red, so we believe that it is a red bird. To remove further doubts about the animal birdness, we call in a bird expert who takes it for examination and concludes that it is not really a bird but some sort of mammal. The question now is whether we should still believe that the animal is red. Postulate* (**CB**) *tells us that we should no longer believe that the animal is red. This can be seen by substituting* $\kappa(E) \equiv \neg\beta =$ *'bird' and* $\alpha =$ *'red' in postulate* (**CB**)*, instructing us to totally ignore the color observation* α *as if it never took place.*

In this example, the original epistemic state reflects the (tentative) belief that the animal is a bird by taking $\kappa(E) =$ 'bird'. The first observation that the animal is red is reflected by taking $\alpha =$ 'red' while the subsequent (reliable) information that the animal is in fact not a bird is reflected by taking $\beta = \neg$'bird'. But then, by taking the set of atoms to be $\{bird, red\}$, it follows that $Mod(\kappa(E)) = \{11, 10\}$, $Mod(\alpha) = \{11, 01\}$, and $Mod(\beta) = \{01, 00\}$. Since $\kappa(E) \wedge \alpha$ is satisfiable, it follows by postulate (**KM** $*$ **2**) that $Mod(\kappa(E * \alpha)) = \{11\}$ and thus $\kappa(E * \alpha)$ entails $\neg\beta$ whence satisfying the if-part of postulate (**CB**). But then $\kappa((E * \alpha) * \beta) \equiv \kappa(E * \beta)$ (if postulate (**CB**) is to be satisfied). Under the assumption (by Darwiche and Pearl) that $\kappa(E * \beta) \equiv \beta$, which satisfies postulate (**KM** $*$ **1**), it then follows directly that $\kappa((E * \alpha) * \beta) \equiv \beta$, meaning that we can no longer tell whether the animal is red or not. So, in satisfying postulate (**CB**), a counterintuitive result is allowed. Darwiche and Pearl use the Red-Bird example as a counterexample to including postulate (**CB**) as a postulate for iterated revision.

In providing a concrete revision operation that satisfies the modified KM postulates (**DP** $*$ **1**) to (**DP** $*$ **6**) and the iterated revision postulates (**C1**) to (**C4**), Darwiche and Pearl provide further justification for their approach. The revision operation is based on Spohn's ordinal conditional functions (see section 4.7). Spohn (1988) proposed a construction for changing an OCF k that takes as input the pair (α, m) where α is the input sentence and m is the firmness with which α is to be accepted in the revised epistemic state (i.e. OCF). The revised OCF $k * (\alpha, m)$ is referred to as the (α, m)-*conditionalisation* of k.

Definition 5.12 *Let k be an OCF and $\alpha \in L$. The (α, m)-**conditionalisation** of k is defined as*

- $$(k * (\alpha, m))(s) = \begin{cases} k(s) - k'(\alpha) & \text{if } s \in Mod(\alpha) \\ k(s) - k'(\neg\alpha) + m & \text{otherwise} \end{cases}$$

The (α, m)-conditionalisation of k has the effect of shifting the models of α 'downwards' (whilst preserving their positions relative to one another) so that the minimal models of α are assigned the number 0, and shifting the nonmodels of α 'upwards' (whilst preserving their positions relative to one another) so that the minimal models of $\neg\alpha$ are assigned the number m. Conditionalisation served as the inspiration for many other approaches, for example, the κ-rankings of Goldszmidt and Pearl (1992) and the transmutations of Williams (1994).

Darwiche and Pearl relax the condition that $k(s) = 0$ for at least one $s \in S$ thus permitting a knowledge base $\kappa(E)$ to be unsatisfiable. They construct a concrete revision operation, based on Spohn's conditionalisation, in which m is effectively $k'(\neg\alpha)+1$. This ensures that the firmness with which an input sentence α is accepted in a revised epistemic state is one degree higher than its current plausibility, thus strengthening the belief in α.

Definition 5.13 *Let E be an epistemic state represented by OCF k. The revision operation $*_{DP}$ of Darwiche and Pearl is defined, for every sentence $\alpha \in L$, as*

- $(k *_{DP} \alpha)(s) = \begin{cases} k(s) - k'(\alpha) & \textit{if } s \in Mod(\alpha) \\ k(s) + 1 & \textit{otherwise} \end{cases}$

Note that if α is unsatisfiable, the knowledge base $\kappa(k *_{DP} \alpha)$ will be unsatisfiable too. The revision operation $*_{DP}$ of Darwiche and Pearl satisfies the modified KM postulates for revision, in which belief revision is a function of epistemic states rather than belief sets, as well as their postulates for iterated revision.

Theorem 5.13 (Darwiche and Pearl, 1997) *The revision operation $*_{DP}$ of Darwiche and Pearl satisfies postulates* $(\mathbf{DP} * \mathbf{1})$ *to* $(\mathbf{DP} * \mathbf{6})$ *and postulates* $(\mathbf{C1})$ *to* $(\mathbf{C4})$.

Another revision operation that satisfies the KM postulates for revision (reformulated for epistemic states) and the DP postulates for iterated revision is the observation-based revision operation $*_{\triangleright}$ of Papini (2001), which uses polynomials on natural numbers to assign to each epistemic state a weighting on the set of possible worlds. The revision operation $*_{\triangleright}$ modifies the weighting by giving preference to the most recent observation (in a sequence of observations) but takes into account the history of observations. Despite the compatibility of Papini's approach providing strong support for postulates $(\mathbf{C1})$ to

(**C4**), attention must be drawn to postulate (**C2**), which has received some criticism in the literature.

In the original formulation of postulate (**C2**) in terms of belief sets (Darwiche and Pearl, 1994), it has been pointed out by Freund and Lehmann (1994) that postulate (**C2**) is inconsistent with the original AGM postulates. The framework of Lehmann (1995), in which an epistemic state is represented by a finite sequence of revisions, has been shown to be incompatible with postulate (**C2**), but compatible with postulates (**C1**), (**C3**), and (**C4**). Similarly, the *choice-based revision* functions of Rott (2001) satisfy postulates (**C3**) and (**C4**) and, depending on the properties of the choice function involved, postulate (**C1**) too, but not postulate (**C2**).

An interesting result by Cantwell (1999) shows that a certain variant of the controversial recovery postulate for contraction is a derived property of any revision operation satisfying the AGM postulates for revision together with postulate (**C2**). More recently, it has been shown by Chopra, Meyer, and Wong (2006) that by weakening the (semantic version) of postulate (**C2**) so that the minimal models of $\neg\alpha$ retain their position after revision (as opposed to all the nonmodels of α retaining their position), a number of recovery-like postulates are satisfied allowing them to prove the result that if $\neg\alpha \in \kappa(E)$ then $\kappa((E * \alpha) - \alpha) \equiv \kappa(E - \alpha)$ and thereby establishing an unexpected connection between the recovery postulate and postulate (**C2**) for iterated revision.

Another criticism that has been raised against the Darwiche-Pearl approach is that it is overly permissive and too limited in scope (Nayak, Pagnucco, and Peppas, 2003). Using the Singing-Bird example of (Nayak et al., 1996) they argue that the Darwiche-Pearl account is too limited in that it does not guarantee seemingly reasonable behaviour.

Example 5.4 (Nayak et al., 1996) *Our agent believes that Tweety is a singing bird. However, since there is no strong correlation between singing and birdhood, the agent is prepared to retain the belief that Tweety sings even after accepting the information that*

Tweety is not a bird, and conversely, if the agent were to be informed that Tweety does not sing, she would still retain the belief that Tweety is a bird. Imagine that the agent first receives the information that Tweety is not a bird, then the information that Tweety does not sing. On such an occasion, it is reasonable to assume that the agent should believe that Tweety is a non-singing non-bird.

However, the postulates do not guarantee this. The reason, they claim, is that none of the DP postulates are applicable when $\kappa(E * \alpha) \not\models \beta$. In a counterexample, Jin and Thielscher (2007) argue that the behaviour which is claimed to be the reasonable one is not generally justified. In the case where the agent initially believes firmly that Tweety is a bird or sings, it would be reasonable, after the indicated successive revisions, for the agent to believe that Tweety is a bird after all. On the other hand, Jin and Thielscher agree with Nayak, Pagnucco, and Peppas that the DP postulates are too permissive. Both parties use the Red-Bird example, which was used as an counterexample by Darwiche and Pearl against natural revision, as a motivation for the claim and as a justification for proposing new postulates. The basic argument is that because the DP postulates are implied by postulate (**CB**), they do not block counterexamples against natural revision. The Red-Bird example will be re-examined in the information-theoretic approach to iterated revision, called templated revision.

5.4 Templated revision

Recall from section 4.6 that in the information-theoretic approach, the epistemic state of an agent is represented by a regular t-ordering $t_E \in T_E$ and associated with every epistemic state t_E are a belief assertion $bel(t_E) \in L$ and a knowledge assertion $know(t_E) \in L$ both of which are satisfiable. The reason for choosing the terms 'belief assertion' and 'knowledge assertion' was to avoid confusion with the terms 'belief base' and 'knowledge base' introduced earlier in the present chapter. Under an extensional interpretation $M = \langle S, l \rangle$ of L every sentence $\alpha \in L$ induces a definite t-ordering t_α such that $bottom(t_\alpha) = Mod_M(\alpha)$ and $top(t_\alpha) = NMod_M(\alpha)$ and conversely, every definite t-ordering $t_\alpha \in T_D$ has a syntactic expression in the form of a sentence $\alpha \in L$ that is a finite axiomatisation of $bottom(t_\alpha)$. By using the semantic representation of input sentences (that is, definite t-orderings) epistemic change operations effectively become operations on regular t-orderings.

Templated expansion is an expansion operation where the epistemic state of an agent is represented by a regular t-ordering. So too are templated revision and templated update[7]. An epistemic change operation under this scenario has to produce not only a new regular t-ordering and an associated belief assertion but a new knowledge assertion too. The knowledge assertion associated with an epistemic state together with the belief assertion play an important role in determining which epistemic change operation should be applied given a specific input sentence. Recall that in the information-theoretic model of epistemic states, the knowledge of an agent is taken to be the next-to-most deeply entrenched beliefs (which include the most deeply entrenched beliefs, namely, the tautologies). Furthermore, it is assumed that the new information is obtained in a reliable way such as making an observation or hearing from another agent who, by virtue of the cooperative nature of diagrammable systems, may be regarded as reliable. This ensures that the principle of Trustworthiness is satisfied. The new information should therefore not be contradictory and, in fact, will be taken to be satisfiable. If the new information is consistent with both the agent's knowledge and its beliefs, then expansion is appropriate. If the new information is consistent with the agent's knowledge but inconsistent with its beliefs, then revision is required. If the new information is inconsistent with the

[7]Some of the results of this section on templated revision appear in van der Westhuizen, van der Poll, and Labuschagne (2006).

agent's knowledge (and hence with the agent's beliefs as well), then the system must have changed state and hence update is called for.

Algorithm 5.1 *Let E be an epistemic state with $bel(E)$ the belief assertion and $know(E)$ the knowledge assertion associated with E. Let $\alpha \in L$ be a satisfiable new input sentence. The basic* ***epistemic change algorithm*** *is provided below:*

function $epistemicChange(E : epistemic\ state; \alpha : L) : epistemic\ state;$
  begin
  if *α is consistent with $know(E)$* ***then***
    if *α is consistent with $bel(E)$* ***then***
      $expand(E, \alpha);$
    else $revise(E, \alpha);$
  else $update(E, \alpha);$
  end;

The basic intuition underlying expansion is that the agent's previous beliefs are retained and expanded to include the given new information, even if it leads to inconsistency. In templated expansion, the idea of expansion, which embodies the principle of Success, is broadened to also include the agent's knowledge but without compromising the principle of Consistency. Templated expansion should therefore result in an epistemic state where the agent's knowledge and beliefs grow monotonically, in other words, templated expansion should result in a regular t-ordering $t_E + \alpha$ where the definite content and indefinite content of $t_E + \alpha$ are greater than or equal to the respective content of the original t-ordering t_E. The requirement for monotonic growth reflects a strong adherence to the principle of Informational Economy while the requirement for epistemic state $t_E + \alpha$ to be a regular t-ordering satisfies the principle of Categorical Matching.

In support of the principle of Minimal Change, a requirement is stated insisting that templated expansion should result in an epistemic state whereby the models of the input sentence retain the relative ordering of the original epistemic state (provided templated expansion is applied in accordance with the epistemic change algorithm). These ideas are formulated in the following semantic characterisation of templated expansion operations.

Definition 5.14 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. A **templated expansion operation** $+$ is an operation such that for every M-satisfiable sentence $\alpha \in L$ that is consistent with $bel(t_E)$, the epistemic state $t_E + \alpha \in T_E$, which is the result of expanding t_E with α, is defined as follows:*

1. $Mod_M(bel(t_E + \alpha)) = Mod_M(bel(t_E) \wedge \alpha)$
2. $Mod_M(know(t_E + \alpha)) = Mod_M(know(t_E) \wedge \alpha)$
3. *For every $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ iff $(t_E + \alpha)(s) \leq (t_E + \alpha)(s')$*

Templated expansion operations support the principle of Qualitativeness in the sense that the resulting epistemic states are regular t-orderings (in which the notion of relative distance is absent). As shown by the following proposition, templated expansion operations also support the principle of Irrelevance of Syntax.

Proposition 5.1 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E, t_{E'} \in T_E$ and let $\alpha, \alpha' \in L$ be consistent with $bel(t_E)$ and $bel(t_{E'})$ respectively. Then it holds that if $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$ then $t_E + \alpha = t_{E'} + \alpha'$.*

Proof. See proof in appendix C, section C.1. ■

Proposition 5.2 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be consistent with $bel(t_E)$. Then the following holds:*

1. $Cont_D(t_E) \subseteq Cont_D(t_E + \alpha)$

2. $Cont_0(t_E) \subseteq Cont_0(t_E + \alpha)$

Proof. See proof in appendix C, section C.1. ■

Proposition 5.3 *Let $T = \langle S, t_E, l \rangle$ and $T' = \langle S, t_E + \alpha, l \rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and consistent with $bel(t_E)$. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E + \alpha, \alpha)$

2. $dt(t_E, \alpha) = dt(t_E + \alpha, \alpha)$

Proof. See proof in appendix C, section C.1. ■

Templated expansion differs from templated revision in the sense that the input sentence triggering revision is not consistent with the agent's beliefs, but has to be consistent with the agent's knowledge. *Templated revision* operates at the level of epistemic states, the representation of which is by means of regular t-orderings. As was the case with templated expansion, a templated revision operation has to produce not only a new regular t-ordering and an associated belief assertion but a new knowledge assertion too. In doing so, the principles of Categorical Matching and Consistency are satisfied. The modified KM postulates ($\mathbf{DP*1}$) to ($\mathbf{DP*6}$) provide criteria that every belief assertion resulting from the revision of an epistemic state should comply with. These postulates have to be modified and augmented to also provide criteria that every knowledge assertion resulting from the revision of an epistemic state should comply with. The following rationality postulates are proposed for templated revision.

Definition 5.15 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. A* ***templated revision operation*** $*$ *is an operation such that for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$, the epistemic state $t_E * \alpha \in T_E$, which is the result of revising t_E by α, satisfies the following set of postulates*[8]*:*

($\mathbf{TR*1}$) $bel(t_E * \alpha) \models_M \alpha$

($\mathbf{TR*2}$) *If $bel(t_E) \wedge \alpha$ is M-satisfiable, then $bel(t_E * \alpha) \equiv_M bel(t_E) \wedge \alpha$*

($\mathbf{TR*3}$) *$bel(t_E * \alpha)$ and $know(t_E * \alpha)$ are M-satisfiable*

($\mathbf{TR*4}$) *If $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$ then $t_E * \alpha = t_{E'} * \alpha'$*

($\mathbf{TR*5}$) $bel(t_E * \alpha) \wedge \beta \models_M bel(t_E * (\alpha \wedge \beta))$

($\mathbf{TR*6}$) *If $bel(t_E * \alpha) \wedge \beta$ is M-satisfiable, then $bel(t_E * (\alpha \wedge \beta)) \models_M bel(t_E * \alpha) \wedge \beta$*

($\mathbf{TR*7}$) $know(t_E * \alpha) \equiv_M know(t_E) \wedge \alpha$

($\mathbf{TR*8}$) *If $(t_E * \alpha)(s) \leq (t_E * \alpha)(s') < n$, then $t_E(s) \leq t_E(s')$*

Postulates ($\mathbf{TR*1}$) to ($\mathbf{TR*6}$) correspond to postulates ($\mathbf{DP*1}$) to ($\mathbf{DP*6}$) with postulates ($\mathbf{TR*1}$), ($\mathbf{TR*2}$), ($\mathbf{TR*5}$), and ($\mathbf{TR*6}$) being identical to postulates ($\mathbf{DP*1}$), ($\mathbf{DP*2}$), ($\mathbf{DP*5}$), and ($\mathbf{DP*6}$) respectively. Note however that the modified KM postulates ($\mathbf{DP*1}$) to ($\mathbf{DP*6}$) are formulated under the classical interpretation $M_0 = \langle S, l\rangle$ whereas the postulates for templated revision are formulated under an extensional interpretation. Postulate ($\mathbf{TR*1}$) is a formal expression of the principle of Success. Postulate ($\mathbf{TR*3}$) is a strengthening of postulate ($\mathbf{DP*3}$) and ensures

[8] *In the spirit of the traditional definition of revision operations, this definition of templated revision operations makes provision for the case where α is consistent with $bel(t_E)$, even though the epistemic change algorithm would never allow it to be performed under those circumstances.*

that both the agent's knowledge and beliefs are satisfiable after performing a templated revision, given that the new information must be consistent with the agent's knowledge for revision to be applicable. Postulate $(\mathbf{TR}*\mathbf{4})$ is a reformulation of postulate $(\mathbf{DP}*\mathbf{4})$. It says that if two epistemic states are equivalent, then revision by equivalent input sentences should result in equivalent revised epistemic states (and consequently, in equivalent belief and knowledge assertions). In contrast to postulate $(\mathbf{DP}*\mathbf{4})$, postulate $(\mathbf{TR}*\mathbf{4})$ is a formal expression of the principle of Irrelevance of Syntax in the context of revision of epistemic states. Postulate $(\mathbf{TR}*\mathbf{7})$ guarantees that the agent's knowledge is expanded during a templated revision, thereby satisfying the principle of Informational Economy. Postulate $(\mathbf{TR}*\mathbf{8})$ ensures that templated revision retains as much as possible of the agent's epistemic state, in accordance with the principle of Minimal Change. To some extent, postulate $(\mathbf{TR}*\mathbf{8})$ reflects a condition present in the construction of the faithful total preorders of Darwiche and Pearl (and Katsuno and Mendelzon) whereby $s \preceq_E s' \stackrel{def}{=} s \in Mod(\kappa(E))$ or $s \in Mod(\kappa(E * \beta))$ for β any finite axiomatisation of $\{s, s'\}$.

This condition plays a key role in the representation theorem of Darwiche and Pearl, where a faithful assignment is required that maps every epistemic state E to a faithful total preorder $\preceq_E$ (with respect to $Cn(\kappa(E))$) in such a way that equivalent epistemic states are assigned the same preorder. The need for a faithful assignment is somewhat unexpected given their view that an epistemic state should comprise not only a belief set, but all of the information needed for changing the belief set. T-orderings alleviate the need for such an assignment because, as shown by proposition 4.12, every t-ordering $t \in T_S$ induces a total preorder $\preceq$ on the set of states that is faithful with respect to $Th_M(bottom(t))$ under an extensional interpretation $M = \langle S, l\rangle$. In other words, every epistemic state $t_E \in T_E$ induces a total preorder $\preceq_{t_E}$ on S that is faithful with respect to $Cn_M(bel(t_E))$. Moreover, for t-orderings that are equal, the faithful total preorders induced by these t-orderings are equal, as shown by the following proposition. On the other hand, different epistemic states may induce the same faithful total preorder.

Proposition 5.4 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t, t' \in T_S$ be t-orderings and let $\preceq_t$ and $\preceq_{t'}$ be the faithful total preorders (with respect to $Th_M(bottom(t))$ and $Th_M(bottom(t'))$) induced by t and t' respectively. If $t = t'$ then $\preceq_t = \preceq_{t'}$.*

Proof. Suppose $t = t'$. Choose any $(s, s') \in \preceq_t$. So $t(s) \leq t(s')$. But $t = t'$ and thus $t'(s) \leq t'(s')$, i.e. $(s, s') \in \preceq_{t'}$. Since (s, s') was chosen arbitrarily it follows that $\preceq_t \subseteq \preceq_{t'}$. Similarly, it can be shown that $\preceq_{t'} \subseteq \preceq_t$. Hence $\preceq_t = \preceq_{t'}$. ∎

In representing the epistemic states of agents by regular t-orderings, a semantic characterisation of templated revision operations can be given without the need for the existence of some additional faithful assignment mapping epistemic states to (faithful) total preorders. This is much closer to the idea that epistemic states should contain all the relevant information for determining how epistemic change operations, revision in this case, should be performed.

Theorem 5.14 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. A templated revision operation $* : T_E \times L \to T_E$ satisfies postulates* $(\mathbf{TR} * \mathbf{1})$ *to* $(\mathbf{TR} * \mathbf{8})$ *iff for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$ the following holds:*

1. $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$
2. $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha)$
3. *For every* $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ *iff* $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$

Proof. See proof in appendix C, section C.1. ∎

From the definition of postulates (**TR** $*$ **1**) to (**TR** $*$ **6**) and theorem 5.14, it follows that every templated revision operation satisfies postulates (**DP** $*$ **1**) to (**DP** $*$ **6**). However, as expected, not every revision operation that satisfies postulates (**DP** $*$ **1**) to (**DP** $*$ **6**) is a templated revision operation, as will be shown by using a concrete revision operation $*_{\boxminus}$ that is defined as an operation on regular t-orderings.

Definition 5.16 *Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. The revision operation $*_{\boxminus}$ is defined, for every sentence $\alpha \in L$, as $t_E *_{\boxminus} \alpha \stackrel{def}{=} t_\alpha \boxminus t_E$.*

Proposition 5.5 *Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. Then the revision operation $*_{\boxminus}$ satisfies postulates* (**DP** $*$ **1**) *to* (**DP** $*$ **6**) *under the classical interpretation $M_0 = \langle S, l \rangle$ of L.*

Proof. See proof in appendix C, section C.1. ∎

Proposition 5.5 demonstrates that the revision of (the belief assertion associated with) epistemic state t_E by input sentence α can be accomplished through the refinement of t-ordering t_α by t_E. This clearly illustrates the priority given to the new information in (an AGM) revision, as required by postulate (**DP** $*$ **1**), also called the *Success* postulate. The class of revision operations that do not come with a guarantee of success is referred to as *non-prioritised revision.* One approach to non-prioritised revision is to construct a revision in two steps, the first of which is to decide whether to accept or reject the new input sentence and the second of which is to perform the revision if the new input sentence is accepted. This is the basic idea behind *screened revision* (Makinson, 1997a) where the decision to accept an input sentence is based on the existence of a set of potential core beliefs that are immune to revision. For non-prioritised base revision, an alternative two-step approach involves as a first step, the expansion of the belief base by the new input sentence and as a second step, the consolidation of the belief base, an operation that makes an unsatisfiable belief base satisfiable. This is the approach taken in *semi-revision* (Hansson, 1997) where consolidation is defined as contraction by a contradiction. A generalisation of non-prioritised base revision is the framework of Ghose and Goebel (1998) which allows for disbeliefs as input and places a linear ordering of reliability on the inputs (including both beliefs and disbeliefs) of the 'information' state. The framework is extended by Chopra, Ghose, and Meyer (2003) where the linear ordering is replaced with a preference ranking. An early survey of non-prioritised revision may be found in the special issue of Erkenntnis on non-prioritised revision (Hansson, 1999b).

In the context of diagrammable systems, the success postulate is desirable. However, the epistemic change algorithm ensures that revision is not applied indiscriminately. It is important to note that the revision operation $*_{\boxminus}$ is not a templated revision operation because it does not satisfy, in particular, postulate ($\mathbf{TR} * \mathbf{7}$), as shown by the following counterexample.

Example 5.5 *For simplicity, we use the classical interpretation* $M_0 = \langle S, l \rangle$. *Let* $t_E = \{(11, 0), (10, 1), (01, 2), (00, 4)\}$ *and let* $t_\alpha = \{(11, 4), (10, 0), (01, 0), (00, 0)\}$.
So $Mod(know(t_E)) = get_{\uparrow}(t_E, n-1) = \{11, 10, 01\}$ *and* $Mod(\alpha) = bottom(t_\alpha) = \{10, 01, 00\}$. *Hence* $Mod(know(t_E) \wedge \alpha) = \{11, 10, 01\} \cap \{10, 01, 00\} = \{10, 01\}$. *Applying the revision operation* $*_{\boxminus}$ *yields* $t_\alpha \boxminus t_E = \{(11, 4), (10, 0), (01, 1), (00, 2)\}$. *So* $Mod(know(t_E *_{\boxminus} \alpha)) = get_{\uparrow}(t_\alpha \boxminus t_E, n-1) = \{10, 01, 00\}$. *But then* $Mod(know(t_E *_{\boxminus} \alpha)) \neq Mod(know(t_E) \wedge \alpha)$. *So* $*_{\boxminus}$ *does not satisfy postulate* ($\mathbf{TR} * \mathbf{7}$).

Postulate ($\mathbf{TR} * \mathbf{7}$) guarantees that the agent's knowledge grow monotonically during a templated revision operation, in other words, it guarantees that the definite content of t-ordering $t_E * \alpha \in T_E$ is greater than or equal to the definite content of the original t-ordering $t_E \in T_E$.

Proposition 5.6 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be consistent with $know(t_E)$. Then $Cont_D(t_E) \subseteq Cont_D(t_E{*}\alpha)$.*

Proof. See proof in appendix C, section C.1. ■

In the construction of the refinement operation $t_\alpha \boxminus t_E$, the definite content of t_E was not kept intact which meant that in the revised epistemic state $t_E *_\boxminus \alpha$, the agent's knowledge was not kept intact. However, by defining a partial refinement operation $t_\alpha \boxplus t_E$ that keeps the definite content of t_E intact, a concrete revision operation $*_\boxplus$ will be defined that is templated.

Definition 5.17 *Let $t_1 \in T_D$ and $t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2\rangle$ and well-ordered by the lexicographic ordering $\preceq$. The **partial refinement operation** $\boxplus$ is a binary operation $T_D \times T_N \to T_N$ where $t_1 \boxplus t_2$, which is the result of partially refining t_1 by t_2, is defined as follows:*

- $(t_1 \boxplus t_2)(s) = \begin{cases} card(seg(t_1(s), t_2(s))) & \text{if } t_1(s) < n \text{ and } t_2(s) < n \\ n & \text{otherwise} \end{cases}$
  where $seg(t_1(s), t_2(s))$ is the initial segment of $(t_1(s), t_2(s)) \in A$.

Proposition 5.7 *Let S be a non-empty set of states with $card(S) = n$. Let $t_1 \in T_D$ and $t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2\rangle$ and well-ordered by the lexicographic ordering $\preceq$. Then the following holds for every $s, s' \in S$:*

1. $t_1(s) \leq (t_1 \boxplus t_2)(s)$

2. *if* $t_1(s) < t_1(s')$ *then* $(t_1 \boxplus t_2)(s) \leq (t_1 \boxplus t_2)(s')$

3. *if* $t_1(s) = t_1(s') \neq n$ *and* $t_2(s) < t_2(s')$ *then* $(t_1 \boxplus t_2)(s) < (t_1 \boxplus t_2)(s')$

4. $bottom(t_1 \boxplus t_2) \subseteq bottom(t_1)$

5. $Cont_D(t_1) \subseteq Cont_D(t_1 \boxplus t_2)$

6. $Cont_0(t_1) \subseteq Cont_0(t_1 \boxplus t_2)$

7. $Cont_D(t_2) \subseteq Cont_D(t_1 \boxplus t_2)$

Proof. See proof in appendix C, section C.1. ∎

In contrast to the refinement operation that allows for the refinement of any t-ordering in normal form, the partial refinement operation only allows for the refinement of definite t-orderings. The primary difference between the two operations is reflected in property 7 of proposition 5.7, which ensures that the definite content of the refining t-ordering is kept intact. The addition of property 7 for partial refinement operations required adjustments to be made to properties 2 and 5 from the corresponding properties of refinement operations too.

Definition 5.18 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. The templated revision operation $*_\boxplus$ is defined, for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$, as $t_E *_\boxplus \alpha \stackrel{def}{=} t_\alpha \boxplus t_E$.*

Proposition 5.8 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. Then the templated revision operation $*_\boxplus$ satisfies postulates $(\mathbf{TR*1})$ to $(\mathbf{TR*8})$.*

Proof. See proof in appendix C, section C.1. ∎

In defining the templated revision operation $*_\boxplus$ in terms of the partial refinement operation $\boxplus$, the principle of Qualitativeness is satisfied for templated revision. To see this, note, firstly, that the index set A induced by $\langle t_\alpha, t_E\rangle$ depends only on the linear order $\leq$ on B (and not on any arithmetic operation on $\{0, 1, \ldots, n\}$), and secondly, that the same holds true for the construction of $(t_\alpha \boxplus t_E)(s)$ for every $s \in S$.

Returning to the Red-Bird example of Darwiche and Pearl (example 5.3 on page 174), note that the example is essentially modelled at the level of belief sets as opposed to at the level of epistemic states (since the total pre-order $\preceq_E$ associated with epistemic state E plays no role in the example). By remodelling the example at the level of epistemic states using t-orderings, it will be shown that it is possible to obtain an intuitively sensible result.

Example 5.6 *For simplicity, the classical interpretation* $M_0 = \langle S, l\rangle$ *is assumed. Suppose* L *is generated by* $Atom = \{P(a), P'(a)\}$ *where* $P(a)$ *represents the fact 'animal a is a bird' and* $P'(a)$ *represents the fact 'animal a is red'. The original epistemic state reflecting the agent's tentative belief that the strange new animal may be a bird is represented by* $t_E = \{(11, 0), (10, 0), (01, 1), (00, 1)\}$ *(where the atoms are considered in the given order, so that* 10 *corresponds to the valuation rendering* $P(a)$ *true but* $P'(a)$ *false). Observing clearly that the animal is red is represented by* $t_\alpha = \{(11, 0), (10, 4), (01, 0), (00, 4)\}$ *where* $\alpha = P'(a)$*. Hearing the reliable information that the animal is in fact a mammal is represented by* $t_\beta = \{(11, 4), (10, 4), (01, 0), (00, 0)\}$ *where* $\beta = \neg P(a)$*. Revising epistemic state* t_E *by* α *yields the epistemic state* $t_\alpha \boxplus t_E = \{(11, 0), (10, 4), (01, 1), (00, 4)\}$ *while subsequent revision of* $t_\alpha \boxplus t_E$ *by* β *yields* $t_\beta \boxplus (t_\alpha \boxplus t_E) = \{(11, 4), (10, 4), (01, 0), (00, 4)\}$*. So* $bottom(t_\beta \boxplus (t_\alpha \boxplus t_E)) = \{01\}$ *and hence the agent believes* $\neg P(a) \wedge P'(a)$*, that the animal is not a bird but nonetheless red. Also note that* $t_\beta \boxplus t_E = \{(11, 4), (10, 4), (01, 0), (00, 0)\}$*.*

The example shows, in contrast to example 5.3, that the templated revision operation $*_\boxplus$ gives an intuitively sensible result when using exactly the same 'conditions' as Darwiche and Pearl (i.e. $Mod(bel(t_E)) = \{11, 10\}$, $Mod(bel(t_\alpha)) = \{11, 01\}$, and

$Mod(bel(t_\beta)) = \{01, 00\}$) and under exactly the same assumption that $\kappa(E*\beta) \equiv \beta$ (since $Mod(bel(t_E *_\boxplus \beta)) = \{01, 00\} = Mod(bel(t_E)))$. Furthermore, postulate (**CB**) is not satisfied since $Mod(bel(t_E *_\boxplus \alpha)) \subseteq Mod(\neg\beta)$ while it is not the case that $Mod(bel((t_E *_\boxplus \alpha) *_\boxplus \beta)) = Mod(bel(t_E *_\boxplus \beta))$. (To see this, note that $Mod(bel(t_E *_\boxplus \alpha)) = bottom(t_\alpha \boxplus t_E) = \{11\}$ and $Mod(\neg\beta) = top(t_\beta) = \{11, 10\}$ with $\{11\} \subseteq \{11, 10\}$ whereas $Mod(bel((t_E *_\boxplus \alpha) *_\boxplus \beta)) = bottom(t_\beta \boxplus (t_\alpha \boxplus t_E)) = \{01\}$ and $Mod(bel(t_E *_\boxplus \beta)) = bottom(t_\beta \boxplus t_E) = \{01, 00\}$ with $\{01\} \neq \{01, 00\}$.) So, in effect, the templated revision operation $*_\boxplus$ blocks natural revision.

The reason for the intuitively sensible result in which natural revision has been blocked is that the Red-Bird example has been remodelled at the level of epistemic states using t-orderings, which allows for a clear distinction between knowledge and belief. In the example, the agent's original epistemic state is represented by an indefinite t-ordering, which accurately reflects the agent's tentative belief that the strange new animal may be a bird and, more importantly, the absence of any knowledge about the animal. If, however, the agent's original epistemic state were taken to be $t_E = \{(11, 0), (10, 0), (01, 4), (00, 4)\}$, then it would be a completely different scenario because the agent would now know that the animal is a bird, i.e. $Mod(bel(t_E)) = Mod(know(t_E)) = \{11, 10\}$. In this scenario, β would be inconsistent with $bel(t_E *_\boxplus \alpha)$ because $t_\alpha \boxplus t_E = \{(11, 0), (10, 4), (01, 4), (00, 4)\}$ and, since t_β is definite, with $know(t_E *_\boxplus \alpha)$ too. But in that case revision is not applicable, rather, update is called for. So, while it would still be the case that $Mod(bel(t_E *_\boxplus \alpha)) \subseteq Mod(\neg\beta)$, the agent's epistemic state after revision by α would not be subsequently revised by β (since β is inconsistent with $know(t_E *_\boxplus \alpha)$) and therefore natural revision would be (indirectly) blocked in this scenario too.

It is subsequently shown that every templated revision operation satisfies DP postulates (**C1**), (**C3**), and (**C4**) for iterated revision but, with the templated revision operation $*_\boxplus$ serving as a counterexample, that it is not that case that every templated revision operation satisfies the controversial DP postulate (**C2**). Since the DP postulates are implied by postulate (**CB**), this result is consistent with example 5.6 in which it has been shown that the templated revision operation $*_{\boxplus}$ does not satisfy postulate (**CB**).

Proposition 5.9 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Every templated revision operation $*$ satisfies postulates* (**C1**)*,* (**C3**)*, and* (**C4**)*, provided $*$ is defined for the input sentence.*

Proof. See proof in appendix C, section C.1. ■

Example 5.7 *For simplicity, the classical interpretation $M_0 = \langle S, l \rangle$ is assumed. Let $t_E = \{(11,0),(10,1),(01,2),(00,4)\}$ and let $t_\alpha = \{(11,4),(10,0),(01,0),(00,4)\}$. Then $t_\alpha \boxplus t_E = \{(11,4),(10,0),(01,1),(00,4)\}$. Let $t_\beta = \{(11,0),(10,4),(01,4),(00,4)\}$. Since $Mod(\beta) = bottom(t_\beta) = \{11\}$ and $NMod(\alpha) = top(t_\alpha) = \{11,00\}$, it follows that $Mod(\beta) \subseteq NMod(\alpha)$. But $t_\beta \boxplus t_E = \{(11,0),(10,4),(01,4),(00,4)\}$ and $t_\beta \boxplus (t_\alpha \boxplus t_E) = \{(11,4),(10,4),(01,4),(00,4)\}$. So $bottom(t_\beta \boxplus t_E) = \{11\}$ and $bottom(t_\beta \boxplus (t_\alpha \boxplus t_E)) = \varnothing$. But then $bottom(t_\beta \boxplus t_E) \neq bottom(t_\beta \boxplus (t_\alpha \boxplus t_E))$ and thus $bel(t_E *_{\boxplus} \beta) \not\equiv bel(t_E *_{\boxplus} \alpha) *_{\boxplus} \beta)$. Hence the templated revision operation $*_{\boxplus}$ does not satisfy postulate* (**C2**)*.*

Recall that postulate (**C2**) states that if $Mod(\beta) \subseteq NMod(\alpha)$, then $Mod(bel((t_E * \alpha) * \beta)) = Mod(bel(t_E * \beta))$. But if $Mod(\beta) \subseteq NMod(\alpha)$ then β is inconsistent with α and hence also inconsistent with $know(t_E) \wedge \alpha$. Since $know(t_E) \wedge \alpha \equiv know(t_E * \alpha)$ it means that β would be inconsistent with $know(t_E * \alpha)$ too. According to the basic epistemic change algorithm, the agent should in this case update $know(t_E * \alpha)$ with β rather than attempt to revise $know(t_E * \alpha)$ with β. Applying revision would yield a t-ordering in normal form that is strongly contradictory and hence both the belief assertion and knowledge assertion would be unsatisfiable. Moreover, applying revision in this instance would result in an increase of the distrust in β, something which is contrary to the normal behaviour of templated revision.

Proposition 5.10 *Let $T = \langle S, t_E, l \rangle$ and $T' = \langle S, t_E * \alpha, l \rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and consistent with $know(t_E)$. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E * \alpha, \alpha)$
2. $dt(t_E * \alpha, \alpha) \leq dt(t_E, \alpha)$

Proof. See proof in appendix C, section C.1. ■

The violation of postulate (**C2**) can be attributed to the lack of a clear criterion in (**C2**) by which the agent may choose between revision and update. In the context of diagrammable systems, the basic epistemic change algorithm suggests that if the new information, which is obtained in a reliable manner and hence not unsatisfiable, is inconsistent with the agent's knowledge, then the system must have changed state and therefore it would be more appropriate for the agent to choose update than to choose revision.

5.5 Templated update

One of the difficulties with the update semantics of Katsuno and Mendelzon, as pointed out by Boutilier (1998), lies in the interpretation of the orderings of 'closeness'. Intuitively, the 'closer' (or more 'similar') a possible world is to another, the smaller the change required to transition from one to the other. The key assumption underlying this interpretation is that the system changes smoothly and gradually, so that the worlds most similar to the present state are the likeliest new states. We feel obliged to point out that a discrete system doesn't have to work that way, and so, when dealing with discrete systems, we will need to add to our semantic architecture an extra device, namely, the *epistemic transition function.* But the case of a continuous system is not ignored - as will be shown, it is just the case in which the epistemic transition function may be taken to be the accessibility function.

For diagrammable systems, which are discrete, a change of state can result in a state of the system that differs radically from its predecessor. As an example of an action that radically changes the state, think about the Light-Fan system - or more elaborate versions thereof in which many components have been added - and imagine that in addition to each component having a switch to turn it on or off there is also a mains switch that, if it is turned off, will turn off every component. Now consider that the simple action of flicking the mains switch can change the system from state 111...1 to 000...0, which in terms of Hamming distance is the furthest state away from 111...1 that one can get.

Templated update operates at the level of epistemic states, the representation of which is by means of regular t-orderings. Operating at the level of epistemic states, a templated update operation has to produce, similarly to templated revision, not only a new regular t-ordering and an associated belief assertion but a new knowledge assertion too. This ensures that the principles of Categorical Matching and Consistency are satisfied.

To allow for comparison with the KM postulates for update, a modification of the postulates is required to accommodate the transition from update as a function of knowledge bases to update as a function of epistemic states.

Definition 5.19 *Let E be an epistemic state and $\kappa(E)$ the knowledge base associated with E. An* ***update operation*** $\diamond$ *is an operation such that for every sentence $\alpha \in L$, the epistemic state $E \diamond \alpha$, which is the result of updating E with α, satisfies the following set of postulates:*

$(\mathbf{KM} \diamond \mathbf{1}')$ $\kappa(E \diamond \alpha) \models \alpha$

$(\mathbf{KM} \diamond \mathbf{2}')$ *If* $\kappa(E) \models \alpha$, *then* $\kappa(E \diamond \alpha) \equiv \kappa(E)$

$(\mathbf{KM} \diamond \mathbf{3}')$ *If both* $\kappa(E)$ *and* α *are satisfiable, then* $\kappa(E \diamond \alpha)$ *is satisfiable*

(KM $\diamond$ 4′) *If $E = E'$ and $\alpha \equiv \alpha'$, then $\kappa(E \diamond \alpha) \equiv \kappa(E' \diamond \alpha')$*

(KM $\diamond$ 5′) $\kappa(E \diamond \alpha) \wedge \beta \models \kappa(E \diamond (\alpha \wedge \beta))$

(KM $\diamond$ 6′) *If $\kappa(E \diamond \alpha) \wedge \beta$ is satisfiable and $\kappa(E)$ is complete, then $\kappa(E \diamond (\alpha \wedge \beta)) \models \kappa(E \diamond \alpha) \wedge \beta$*

(KM $\diamond$ 7′) $\kappa((E \bigtriangledown E') \diamond \alpha) \equiv \kappa(E \diamond \alpha) \vee \kappa(E' \diamond \alpha)$

Similarly to revision, postulate (**KM $\diamond$ 4′**) is a reformulation of postulate (**KM $\diamond$ 4**). The only other difference between these postulates and the KM postulates for update is that postulate (**KM $\diamond$ 7′**) requires an operation to be defined on epistemic states so that the *Disjunction Rule* holds on the associated knowledge bases. For templated update, where the epistemic state of an agent is represented by a regular t-ordering, the required operation $\bigtriangledown$ must ensure that the belief assertion associated with the resulting t-ordering $t_E \bigtriangledown t_{E'}$ is semantically equivalent to the disjunction of the belief assertions associated with t-orderings t_E and $t_{E'}$ respectively.

Definition 5.20 *Let S be a non-empty set of states with $card(S) = n$. Let $t_1, t_2, \ldots, t_n \in T_N$. The* ***minimise operation*** *$\bigtriangledown$ is a n-ary operation on T_N defined as $\bigtriangledown_{i=1\ldots n} t_i = g \circ \vee(t_i)$ where $\vee(t_i)(s) = \min\{t_1(s), t_2(s), \ldots, t_n(s)\}$ and g is the normalise function for $ran(\vee(t_i))$.*

The effect of the minimise operation is to create a t-ordering in normal form in which the states are as low as possible given the input t-orderings. A minimise operation results in a t-ordering $t_1 \bigtriangledown t_2 \bigtriangledown \ldots \bigtriangledown t_n$ in normal form, the definite content of which is a subset of the definite content of every t_i and the indefinite content of which is, similarly, a subset of the indefinite content of every t_i.

Proposition 5.11 *Let S be a non-empty set of states with $card(S) = n$ and let $t_1, t_2, t_3 \in T_N$. Then the following holds:*

1. $t_1 \bigtriangledown t_2 = t_2 \bigtriangledown t_1$

2. $t_1 \bigtriangledown (t_2 \bigtriangledown t_3) = (t_1 \bigtriangledown t_2) \bigtriangledown t_3$

3. $bottom(t_1 \bigtriangledown t_2) = bottom(t_1) \cup bottom(t_2)$

4. $Cont_D(t_1 \bigtriangledown t_2) = Cont_D(t_1) \cap Cont_D(t_2)$

5. $Cont_I(t_1 \bigtriangledown t_2) = Cont_I(t_1) \cap Cont_I(t_2)$

Proof. See proof in appendix C, section C.2. ■

From proposition 5.11(3) it follows that the minimise operation $\bigtriangledown$ satisfies the requirement imposed by postulate ($\mathbf{KM} \diamond \mathbf{7'}$). Recall that $Mod(bel(t_E)) = bottom(t_E)$ under an extensional interpretation and hence also under the classical interpretation $M_0 = \langle S, l \rangle$. Then it is easy to see that the belief assertion associated with t-ordering $t_E \bigtriangledown t_{E'}$ is semantically equivalent to the disjunction of the belief assertions associated with t-orderings t_E and $t_{E'}$ respectively.

An *epistemic transition function* is a function that reflects the notion 'likelihood of transition' by mapping every state to a regular t-ordering. Intuitively, it represents the agent's knowledge and beliefs about the most likely transitions of the system from one state to the next. To ensure that the agent would not consider it likely for the system to transition to a state that it knows will never arise, an epistemic transition function would have to obey special restrictions with regards to the agent's fixed information.

Definition 5.21 *Let $t_I \in T_E$ be the agent's initial epistemic state with definite t-ordering $t_{SD} \in T_E$ representing the agent's fixed information. An* ***epistemic transition function*** *is a function $\vec{F} : S \rightarrow T_E$ such that for every $s \in S$, $\vec{F}(s)(s') = n$ if $s' \in top(t_{SD})$ and $\vec{F}(s) = t_I$ if $s \in top(t_{SD})$.*

In the case of a continuous system, the notion 'likelihood of transition' is proportional to the notion 'similarity to the present actual state'. Recall that the accessibility function reflects the agent's ability to distinguish between the actual state and other states. If we assume that the system changes 'continuously', then this agent-oriented kind of similarity is a guide to likelihood, 'as far as the agent can tell', of being a transition target. In this case, the epistemic transition function may be taken to be the accessibility function.

An epistemic transition function is a component of an agent's epistemic framework in the sense that it provides the agent with information about the behaviour of the system. Recall that collectively, the information available to the agent provides the epistemic framework from which the agent's epistemic state is formed, and changed. As a component of the agent's epistemic framework, epistemic transition functions play an important role in templated update. To illustrate, we shall use a concrete templated update operation $\diamond_{\boxminus}$ that is defined as an operation on regular t-orderings. The formal proof that $\diamond_{\boxminus}$ is a templated updated operation according to our (still to be defined) postulates for templated update will be deferred until proposition 5.13.

Definition 5.22 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. The* ***templated update operation*** *$\diamond_{\boxminus}$ is defined, for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, as $t_E \diamond_{\boxminus} \alpha \stackrel{def}{=} \bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))$.*

The templated update operation $\diamond_{\boxminus}$ operates on each model of the belief assertion associated with the agent's epistemic state independently, by refining the (definite t-ordering induced by the) input sentence with the regular t-ordering associated with each model through the epistemic transition function. This 'pointwise' refinement satisfies the principle of Qualitativeness. To see this, note, firstly, that the index set A induced by $\langle t_\alpha, t_E \rangle$ (according to definition 4.11) depends only on the linear order $\leq$ on B, and secondly, that the same holds true for the construction of $t_\alpha \boxminus \vec{F}(s)$ for every $s \in S$ (according to definition 4.12). The minimise operation $\bigtriangledown$ fulfils, at the level of epistemic states, the role fulfilled by set union at the level of belief sets. From definition 5.20 it is clear to see that the minimise operation satisfies the principle of Qualitativeness. The templated update operation $\diamond_\boxminus$ satisfies the modified KM postulates for update in which belief update is a function of epistemic states rather than belief sets.

Proposition 5.12 *Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. Then the templated update operation $\diamond_\boxminus$ satisfies postulates* $(\mathbf{KM} \diamond \mathbf{1'})$ *to* $(\mathbf{KM} \diamond \mathbf{7'})$ *for every satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, under the classical interpretation $M_0 = \langle S, l \rangle$ of L.*

Proof. See proof in appendix C, section C.2. ■

Since templated update is only defined for satisfiable sentences that are inconsistent with the knowledge assertion (and hence with the belief assertion) associated with an epistemic state, it follows that postulate $(\mathbf{KM} \diamond \mathbf{2'})$ is satisfied vacuously, in the sense that the hypothesis is never satisfied. Nonetheless, if a templated update operation $\diamond$ were applied to a sentence α that is entailed by $bel(t_E)$, then it would not always be the case that $Mod(bel(t_E \diamond \alpha)) = Mod(bel(t_E))$ and thus $\diamond$ would not satisfy postulate $(\mathbf{KM} \diamond \mathbf{2'})$. This is because postulate $(\mathbf{KM} \diamond \mathbf{2'})$ relies on each model s of $bel(t_E)$ to be minimal in $\vec{F}(s)$ with no other state $s' \neq s$ minimal in $\vec{F}(s)$ for it to be satisfied and the definition of an epistemic transition function $\vec{F}$ does not require that $\vec{F}(s)(s) = 0$ and that $\vec{F}(s)(s') > 0$ if $s' \neq s$, although it does not exclude such a possibility. Intuitively, it means that postulate $(\mathbf{KM} \diamond \mathbf{2'})$ is appropriate under the assumption that it is more normal for the system to persist in a state than to transition to another state. The *event-based* model and *generalized update (GU)* model of Boutilier (1996b, 1998) are examples of other approaches that do not make this assumption. Nonetheless, both can model scenarios under which the assumption is true (and postulate (**KM** $\diamond$ **2**$'$) satisfied) through the notion of a null event as the most plausible event in any situation. Scenarios under which the assumption is true are still subject to the criticism of Herzig and Rifi, provided that the hypothesis that observations may be unreliable is accepted, a hypothesis that we have rejected in the context of diagrammable systems.

The modified KM postulates for update provide criteria that every belief assertion resulting from the update of an epistemic state should comply with. But for templated update, it is not only the agent's beliefs that are updated but its knowledge too and since the new information is inconsistent with the agent's knowledge, templated update relies on input from the agent's epistemic framework to update the agent's current epistemic state. The following postulates are proposed for templated update.

Definition 5.23 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. A **templated update operation** $\diamond$ is an operation such that for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, the epistemic state $t_E \diamond \alpha \in T_E$, which is the result of updating t_E with α, satisfies the following set of postulates:*

(**TU** $\diamond$ **1**) $bel(t_E \diamond \alpha) \models_M \alpha$

(**TU** $\diamond$ **2**) $bel(t_E \diamond \alpha)$ *and* $know(t_E \diamond \alpha)$ *are M-satisfiable*

(**TU** $\diamond$ **3**) *If* $t_E = t_{E'}$ *and* $\alpha \equiv_M \alpha'$ *then* $t_E \diamond \alpha = t_{E'} \diamond \alpha'$

(**TU** $\diamond$ **4**) $bel(t_E \diamond \alpha) \wedge \beta \models_M bel(t_E \diamond (\alpha \wedge \beta))$

(**TU** $\diamond$ **5**) *If* $bel(t_E \diamond \alpha) \wedge \beta$ *is M-satisfiable and* $bel(t_E)$ *is complete, then* $bel(t_E \diamond (\alpha \wedge \beta)) \models_M bel(t_E \diamond \alpha) \wedge \beta$

$(\mathbf{TU} \diamond \mathbf{6})$ $bel((t_E \bigtriangledown t_{E'}) \diamond \alpha) \equiv_M bel(t_E \diamond \alpha) \vee bel(t_{E'} \diamond \alpha)$

$(\mathbf{TU} \diamond \mathbf{7})$ $know(t_E \diamond \alpha) \equiv_M \alpha$

$(\mathbf{TU} \diamond \mathbf{8})$ *If* $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2) < n$, *then* $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ *for some* $s \in Mod_M(bel(t_E))$

Postulates $(\mathbf{TU} \diamond \mathbf{1})$, $(\mathbf{TU} \diamond \mathbf{4})$, $(\mathbf{TU} \diamond \mathbf{5})$, and $(\mathbf{TU} \diamond \mathbf{6})$ are identical to postulates $(\mathbf{KM} \diamond \mathbf{1}')$, $(\mathbf{KM} \diamond \mathbf{5}')$, $(\mathbf{KM} \diamond \mathbf{6}')$, and $(\mathbf{KM} \diamond \mathbf{7}')$ respectively. Postulate $(\mathbf{TU} \diamond \mathbf{2})$ is a strengthening of postulate $(\mathbf{KM} \diamond \mathbf{3}')$ and ensures that the agent's knowledge and beliefs are satisfiable after performing a templated update, given the constraints on the input sentence. Postulate $(\mathbf{TU} \diamond \mathbf{3})$ is a reformulation of postulate $(\mathbf{KM} \diamond \mathbf{4}')$. It says that if two epistemic states are equivalent, then updating by equivalent input sentences should result in equivalent updated epistemic states (and consequently, in equivalent belief and knowledge assertions). In contrast to postulate $(\mathbf{KM} \diamond \mathbf{4}')$, postulate $(\mathbf{TU} \diamond \mathbf{3})$ is a formal expression of the principle of Irrelevance of Syntax in the context of update of epistemic states. Postulate $(\mathbf{TU} \diamond \mathbf{7})$ guarantees that the agent's knowledge is replaced by the input sentence. It represents the strongest possible formal expression of the principle of Success. Postulate $(\mathbf{TU} \diamond \mathbf{8})$ ensures that the agent's epistemic state is updated based on the most likely (or normal) transitions of the system from one state to the next. This gives a different context for applying the principles of Minimal Change and Informational Economy. It is reminiscent of a condition present in the construction of the faithful update assignment of Katsuno and Mendelzon whereby a total preorder $\preceq_s$ is mapped to each world s such that $s_1 \preceq_s s_2 \stackrel{def}{=} s_1 = s$ or $s_1 \in Mod(\kappa(E_s \diamond \beta))$ for $\kappa(E_s)$ any finite axiomatisation of $\{s\}$ and β any finite axiomatisation of $\{s_1, s_2\}$.

Albeit significantly different from the faithful update assignment, the epistemic transition function plays, similarly to the faithful update assignment, a key role in the representation theorem for templated update.

Theorem 5.15 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. A templated update operation $\diamond : T_E \times L \to T_E$ satisfies postulates $(\mathbf{TU} \diamond \mathbf{1})$ to $(\mathbf{TU} \diamond \mathbf{8})$ iff for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, the following holds:*

1. $Mod_M(bel(t_E \diamond \alpha)) = \bigcup_{s \in Mod(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$
2. $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$
3. *For every $s_1, s_2 \in Mod_M(\alpha)$, if $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$*

Proof. See proof in appendix C, section C.2. ■

Using the representation theorem for templated update, it will now be formally shown that the concrete templated update operation $\diamond_\boxminus$ introduced earlier is indeed a templated update operation.

Proposition 5.13 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. Then the templated update operation $\diamond_\boxminus$ satisfies postulates $(\mathbf{TU} \diamond \mathbf{1})$ to $(\mathbf{TU} \diamond \mathbf{8})$ for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$.*

Proof. See proof in appendix C, section C.2. ■

Templated update operations behave like templated revision operations in the sense that an update operation increases the plausibility of the input sentence with respect to the updated epistemic state (as compared to the original epistemic state) and, likewise, decreases the distrust in the input sentence.

Proposition 5.14 *Let $T = \langle S, t_E, l\rangle$ and $T' = \langle S, t_E \diamond \alpha, l\rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and inconsistent with $know(t_E)$. Let $\vec{F}$ be an epistemic transition function. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E \diamond \alpha, \alpha)$

2. $dt(t_E \diamond \alpha, \alpha) \leq dt(t_E, \alpha)$

Proof. See proof in appendix C, section C.2. ■

The essential difference between templated update and templated revision is that templated revision retains the agent's knowledge from the original epistemic state (as shown by proposition 5.6) whereas templated update does not.

Proposition 5.15 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be M-satisfiable and inconsistent with $know(t_E)$. Then $Cont_D(t_E) \not\subset Cont_D(t_E \diamond \alpha)$.*

Proof. See proof in appendix C, section C.2. ■

The other important difference is, of course, that whereas with templated revision the ordering is embedded in the representation of an epistemic state as a regular t-ordering, with templated update an epistemic transition function is required that maps each model of the belief assertion associated with the epistemic state to a regular t-ordering.

Another way to compare templated revision and templated update is through the respective links with nonmonotonic logic and counterfactuals. The connection between KM update and *counterfactuals* that has been shown to exist (Grahne, 1991; Makinson, 1993; Ryan, Schobbens, and Rodrigues, 1996; Ryan and Schobbens, 1997) relies on a notion of centeredness. An ordering $\preceq_s$ on S is *weakly centered* iff for any $s' \in S$, $s' \not\prec_s s$ and *fully centered* iff for any $s' \in S$, if $s \neq s'$ then $s \prec_s s'$. By definition, the faithful update assignments of KM update ensures that the orderings associated with each state are fully centered. As discussed earlier, epistemic transition functions are more general than the faithful update assignments of KM update. Epistemic transition functions do not satisfy the constraint of either full centering or weak centering. The implication is that the established connection between KM update and counterfactuals will not be maintained between templated update and counterfactuals. Taking heed of the cautionary remarks by Makinson (2003a) against an over-emphasis on formal translations that may obscure underlying differences of gestalt and intuition, the possible connections between templated update and counterfactuals will not be explored. However, because of the well-established relationship between revision and nonmonotonic logic (Makinson and Gärdenfors, 1991) and the link between t-orderings and strict modular partial orderings, some connections between templated revision and nonmonotonic logic will be explored in section 5.7.

Notwithstanding the differences and similarities between templated revision and templated update, the important feature is that both operations can be performed within the same framework and that the agent has a means (the epistemic change algorithm) to choose between these operations. As shown in the next section, this feature is generally absent in related frameworks of revision and update.

5.6 Other frameworks of revision and update

A framework for revision and update would be closely related to the *templated framework* if it allows an agent to select between revision and update in such a way that a revision of the agent's knowledge and beliefs may be followed by another revision or by an update, and similarly, that an update of the agent's knowledge and beliefs may be followed by another update or by a revision. In the templated framework the notion of iterated belief revision is broadened in the sense that an epistemic change operation may be followed by a different epistemic change operation and that an agent must be able to select the appropriate epistemic change operation to perform. This notion of *iterated epistemic change* is not prevalent in the literature. A notable exception is the recent investigation by Nayak et al. (2005) where, instead of using revision operations, different iterations of expansion and contraction operations (at the level of belief sets) are considered, in particular, the iteration where an expansion is followed by a contraction and the iteration where a contraction is followed by another contraction. However, the notion of iterated update, which has received little attention in the literature, is not considered.

One approach that pays attention to both iterated revision and iterated update is the *run-based framework* of Friedman and Halpern (1999b), which is the only approach from those modelling both knowledge and belief (see section 4.4) to consider the question of epistemic change. Their run-based framework, in which belief is modelled in terms of knowledge and plausibility, and time is modelled explicitly using the standard modal operators of temporal logic, captures both revision and update. The underlying system is characterised by describing it in terms of a state that changes over time, in other words, as sequences of state transitions over time, called *runs*. Formally, a run of the system is a function r from time to pairs (s_e, s_a), called global states, comprising the environment state s_e and the agent's local state s_a with $r_e(m) = s_e$ and $r_a(m) = s_a$. A *plausibility system* is a possible worlds interpretation $\mathcal{I} = \langle S, l, P \rangle$ where S is the set of runs, or rather *points* (r, m), and P is the plausibility assignment (mapping each point (r, m) to a plausibility space). The systems under consideration are assumed to be synchronous with agents having perfect recall. This ensures that any runs that an agent considers impossible at (r, m) are also considered impossible at $(r, m+1)$. A plausibility system is a *belief change system* (BCS) if it satisfies a number of conditions. Firstly, the language includes a set of atoms for reasoning about the environment, whose truth depend only on the environment state ($BCS1$) and a disjoint set of atoms for reasoning about observations ($BCS3$). Secondly, the agent's local (or epistemic) state s_a consists of a sequence of observations ($BCS2$), which are assumed to be reliable ($BCS4$). The belief set associated with the agent's epistemic state s_a is defined as follows:

- $Bel(\mathcal{I}, s_a) = \{\alpha \mid \mathcal{I}, (r,m) \Vdash \Box_B \alpha$ for some (r,m) such that $r_a(m) = s_a\}$.

Lastly, the agent has a prior plausibility measure over runs describing the agent's initial assessment on the possible executions of the systems ($BCS5$). The prior is constrained ($PRIOR$) to ensure that the agent's plausibility assessment at each point is determined by its prior.

Revision is captured by further restricting belief change systems. Environmental atoms do not change their truth value along a run ($REV1$) capturing the intuition that revision deals with 'static' worlds. The agent's prior is ranked ($REV2$), resulting in a total preorder over points. If a sentence α is satisfiable, then the prior over runs in which α is true at the initial state must be nontrivial ($REV3$). Finally, the observation of α does not provide any additional information ($REV4$). In a belief change system that satisfies conditions ($REV1$) to ($REV4$), the agent revises its epistemic state s_a by conditioning on its prior. The most plausible runs after revision by an observation α determine the agent's new epistemic state $s_a \cdot \alpha$. Friedman and Halpern show that, provided the belief set is satisfiable, revision by conditioning satisfies the AGM postulates for revision.

Update is captured by restricting belief change systems by a different set of requirements. There is a bijective correspondence between the set of valuations, which is assumed to be finite, and the states of the environment ($UPD1$). Similarly to ($REV2$), requirement ($UPD2$) puts constraints on the agent's prior. The prior is essentially a lexicographic ordering, based on a distance measure, that compares a set of runs sharing a common prefix of environment states. All sequences of satisfiable sentences have nontrivial plausibility ($UPD3$). As before, the agent obtains no additional information from observing α beyond the fact that α is true ($UPD4$). In a belief change system that satisfies conditions ($UPD1$) to ($UPD4$), the agent updates its epistemic state s_a by conditioning on its prior. Friedman and Halpern show that update by conditioning satisfies the KM postulates for update (which characterise update operations in terms of partial orders).

The important feature of the run-based framework is that both revision and update use conditioning to perform belief change. The essential difference lies in the allowable sequences of runs and in the agent's initial beliefs (as reflected in its prior).

The run-based framework of Friedman and Halpern is very different from the templated framework and not directly comparable. The class of agents that they consider can be viewed as second-order intentional systems, which focus exclusively on the informational attitudes of knowledge and belief; the templated framework, in contrast, regards an agent as a first-order intensional system, but with the same focus on informational attitudes. The class of objective systems (i.e. agent environments) differs too: in the run-based framework, systems are continuous, as implied by the assumption of synchronicity and requirement ($UPD2$), while, in the templated framework, systems are discrete. Another important difference is that the dynamics of systems in the run-based framework is modelled directly in the semantics of the underlying (modal) language whereas, in the templated framework, it is modelled as a component of the agent's epistemic framework. An interesting correlation between the systems is that the belief change system for update is essentially a diagrammable system, by virtue of requirement ($UPD1$). The representation of an agent's epistemic state as a sequence of (reliable) observations is not comparable with the representation of an agent's epistemic state as a regular t-ordering. Although both iterated revision and iterated update can be captured in the run-based framework, each requires a different belief change system and, in that sense, the notion of iterated epistemic change as described in the templated framework is not supported. Since knowledge in the run-based framework is related to the agent's capability to distinguish between states and not related to a notion of entrenchment, the need for revising and updating knowledge does not arise in the run-based framework.

The transmutation approach of Lang, Marquis, and Williams (2001), which is an extension of the transmutation approach of Williams (1994) for belief revision, focuses on iterated update at the level of epistemic states. Lang, Marquis, and Williams use a finitely generated propositional language under a traditional truth-value semantics, which of course, corresponds to the classical interpretation $M_0 = \langle S, l \rangle$. In their approach, the epistemic state of an agent is taken to be an OCF that is induced by a stratified belief base. A *stratified belief base* (SBB) B is a finite sequence $\langle \alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_\infty \rangle$ of sentences α_i of rank i. Fully certain beliefs are represented by α_∞ while α_n and α_1 represent respectively the most entrenched and the least entrenched of the uncertain beliefs. The OCF k_B induced by an SBB B is defined, for every $s \in S$, as follows:

- $k_B(s) = \begin{cases} \max\{i \mid s \in Mod(\neg\alpha_i)\} & \text{if such an } i \text{ exists} \\ 0 & \text{otherwise} \end{cases}$

Note that the range of an OCF is taken to be the set $\mathbb{N} \cup \{\infty\}$, as opposed to the class of ordinals, and that the condition that $k(s) = 0$ for at least one $s \in S$ is not mandated but instead used to define a *normalised* OCF. A OCF k_0 is said to be at least as specific as an OCF k_1, denoted by $k_1 \leq k_0$, iff for every $s \in S$, $k_1(s) \leq k_0(s)$.

A *transmutation* is any operation $\star$ that maps an OCF k and an input pair (α, m)[9], comprising a satisfiable nontautological sentence α and a rank m, to a new OCF $k \star (\alpha, m)$ such that

[9] To be precise, Williams (1994) uses pairs (Δ, m) comprising a nonempty set of consistent complete theories Δ (possible worlds X where $\varnothing \subset X \subset S$) and a degree of acceptance (or rank) m. However, for a finitely generated propositional language under a traditional truth-value semantics, every such X has a finite axiomatisation as a sentence α such that $Mod(\alpha) = X$.

1. $(k \star (\alpha, m))(\neg\alpha) = m$ and

2. $$Mod(Bel(k \star (\alpha, m))) = \begin{cases} \{s \in Mod(\alpha) \mid k(s) = k'(\alpha)\} & \text{if } m > 0 \\ \{s \in S \mid k(s) = 0 \text{ or} & \text{otherwise} \\ (s \in Mod(\neg\alpha) \text{ and } k(s) = k'(\neg\alpha)\} & \end{cases}$$

The OCF $k \star (\alpha, m)$ is said to be a (α, m)-transmutation of k. The effect of a transmutation, for $m > 0$, it to push the minimal nonmodels of α to level m and to pull the minimal models of α to level 0. It ensures that the input sentence α is accepted in the new OCF with firmness m. Spohn's (α, m)-conditionalisation of k for revision, as defined by definition 5.12, is one example of a (α, m)-transmutation of k. Another example of a transmutation is an *adjustment* (Williams, 1994) in which the OCF k is disturbed as little as necessary to accept α with firmness m. In the context of stratified belief bases, adjustment can be seen as a strategy for resolving conflict in stratified belief bases with *maxi-adjustment* (Williams, 1996) and *disjunctive-maxi-adjustment* (Benferhat, Kaci, Le Berre, and Williams, 2004) proposed as more refined conflict resolution strategies, both based on adjustment.

Lang, Marquis, and Williams use the alternative characterisation of dependency-based update operations to model an update operation. The principle of 'forget, then expand' is extended to OCFs in the following way[10]. An OCF k is said to be independent from $A \subseteq Atom$ iff for any nontautological sentence α such that $Atm(\alpha) \subseteq A$, $k'(\alpha) = 0$. The set $Dep(k)$ is defined as $Dep(k) = \{\rho_i \in Atom \mid k \text{ depends on } \{\rho_i\}\}$. The notion of forgetting is defined by taking $Forget(k, A)(s) = \min\{k(s') \mid s' \in S \text{ and } l(s) \approx_{Atom-A} l(s')\}$.

The *U-transmutation* of k by the pair (α, m) with respect to *Dep* and a transmutation operation $\star_U$ such that $(k \star_U (\alpha, m))(\alpha) = m$ is defined by $k \diamond_U (\alpha, m) =$

[10] Lang, Marquis, and Williams (2001) also provide an alternative formulation in terms of SBBs for the case where an OCF is defined implicitely by a SBB.

$Forget(k, Dep(\alpha)) \star_U (\alpha, m)$. By taking $\star_U$ to be conditionalisation, the (α, m)-*U-transmutation* of k (with respect to Dep) becomes:

- $(k \diamond_U (\alpha, m))(s) = \begin{cases} Forget(k, Dep(\alpha))(s) & \text{if } s \in Mod(\neg\alpha) \\ Forget(k, Dep(\alpha))(s) + m & \text{otherwise} \end{cases}$

The effect of 'forgetting' is reflected by the OCF $Forget(k, Dep(\alpha))$ in which selected models of α (and $\neg\alpha$) are pulled down to level 0 while the effect of 'expanding' is reflected by the U-transmutation of $Forget(k, Dep(\alpha))$ by (α, m) in which the minimal models of α are pushed up to level m. This implies that the input sentence α is rejected (as opposed to being accepted) in the new OCF with firmness m.

To model an update operation at the level of epistemic states, the OCFs k has to be transmuted by $(\neg\alpha, \infty)$ to ensure the acceptance of α in the updated OCF. An update operation $\diamond$ on OCFs can be defined as $k \diamond \alpha \stackrel{def}{=} Forget(k, Dep(\alpha)) \star_U (\neg\alpha, \infty)$. Although both the update and revision operations on OCFs make use of transmutation operations, the transmutation operations $\star_U$ and $\star$ have almost the opposite effect on the OCFs being transmuted.

The *transmutation framework* of Lang, Marquis, and Williams is comparable with the templated framework largely because of the close connection between OCFs and regular t-orderings as representations of epistemic states. Recall from section 4.7 that the essential differences between OCFs and regular t-orderings are that OCFs are quantitative in nature whereas regular t-orderings are purely qualitative and that the representation of knowledge is not supported by OCFs. An obvious difference between the transmutation update operations of Lang, Marquis, and Williams and templated update operations is that the former is dependency-based while the latter is minimisation-based, and both with different underlying semantic characterisations. The most important difference, however, is that their definition of dependence is based on a notion of 'closeness' between possible worlds, suggesting that the dynamics of the underlying system, which has been left undefined, is continuous rather than discrete. Although the notion of iterated epistemic change as described in the templated framework is supported in the sense that an epistemic change operation may be followed by a different epistemic change operation (in the same framework), it is not fully supported as the agent has not been equipped with a mechanism to select the appropriate epistemic change operation to perform.

5.7 Revision and nonmonotonic reasoning

Belief revision and nonmonotonic logic have a close relationship with each other. The key idea underlying this relationship, due to Makinson and Gärdenfors (1991), is that a belief set $K * \alpha$, the result of revising belief set K by input sentence α, can be seen as the set of defeasible consequences of α under the defeasible consequence relation determined by the revision operation $*$. Conversely, the set of defeasible consequences of α under some defeasible consequence relation can be seen, by fixing a belief set K in some appropriate way, as the revised belief set $K * \alpha$. Formally, this idea, which has become known as the Makinson-Gardenfors translation, can be expressed as $\alpha \mid\sim \beta$ iff $\beta \in K * \alpha$ for some fixed belief set K.

Using this translation scheme, Makinson and Gärdenfors (1991) shows that the AGM postulates for revision can be translated into postulates for nonmonotonic reasoning, and vice versa. The properties required to be satisfied by the rational consequence relations of Lehmann and Magidor, which will be referred to as the LM postulates for nonmonotonic logic, play an important role in these translations.[11] From the LM postulates, the following postulates may be derived:

- If $\alpha \mid\sim \beta$ for all $\beta \in \Gamma$ and $\Gamma \models \gamma$ then $\alpha \mid\sim \gamma$ *(Closure)*

[11] Using different sets of postulates, Rott (1998, 2001) shows that the theory of rational choice can be used to define nonmonotonic inferences which parallel semantic (and syntactic) constructions of belief change based on choice functions.

- If $\alpha \wedge \beta \mid\!\sim \gamma$ then $\alpha \mid\!\sim \beta \rightarrow \gamma$ *(Conditionalisation)*

Closure follows from *Right Weakening* and *And* using the compactness of the semantic consequence relation $\models$. In the presence of *Reflexivity*, *Right Weakening*, and *Left Logical Equivalence*, *Or* implies *Conditionalisation*[12] (Kraus, Lehmann, and Magidor, 1990). *Weak Conditionalisation* is a special case of *Conditionalisation* in which $\alpha = \top$. Similarly, *Weak Rational Monotonicity* is the special case of *Rational Monotonicity* in which $\alpha = \top$. It turns out that AGM postulates $(\mathbf{K} * \mathbf{1})$ to $(\mathbf{K} * \mathbf{4})$ correspond to *Closure*, *Reflexivity*, *Weak Conditionalisation*, and *Weak Rational Monotonicity* respectively while AGM postulates $(\mathbf{K} * \mathbf{6})$ to $(\mathbf{K} * \mathbf{8})$ correspond to *Left Logical Equivalence*, *Conditionalisation*, and *Rational Monotonicity* respectively. AGM postulate $(\mathbf{K} * \mathbf{5})$ corresponds to a postulate, called *Consistency Preservation*, which was introduced in Makinson and Gärdenfors (1991):

- If $\alpha \mid\!\sim \bot$ then $\alpha \models \bot$ *(Consistency Preservation)*

Consistency Preservation says that contradictions may only be defeasible consequences of contradictions. It is not satisfied by all rational consequence relations, in other words, given a ranked interpretation $P = \langle S, R, l \rangle$ of L, it is not necessarily the case that if $\alpha \mid\!\sim_P \bot$ then $\alpha \models \bot$. This is because it is possible for α to be P-unsatisfiable, i.e. $Mod_P(\alpha) = \varnothing$, so that $Min_R(\alpha) = \varnothing$ even though α is not a contradiction, i.e. α is not P_0-unsatisfiable where $P_0 = \langle S, R, l \rangle$ for $S = U_T$ and l the identity function. However, if *Consistency Preservation* is formulated so that instead of taking the entailment relation to be $\models$, that is $\models_{P_0}$, it is taken to be $\models_P$, then every rational consequence relation satisfies (this weakened form of) *Consistency Preservation*.

[12] *Conditionalisation is referred to as rule S by Kraus, Lehmann and Magidor (1990).*

Proposition 5.16 *Let $P = \langle S, R, l \rangle$ be a ranked interpretation of L and let $\mid\!\sim_P$ be the rational consequence relation on L induced by P. Then $\mid\!\sim_P$ satisfies the following property:*

- *if $\alpha \mid\!\sim_P \bot$ then $\alpha \models_P \bot$ (Weak Consistency Preservation)*

Proof. Let $\alpha \mid\!\sim_P \bot$. But then $Min_R(\alpha) = \varnothing$. Suppose that $\alpha \not\models_P \bot$. So $Mod_P(\alpha) \neq \varnothing$. But R is $Mod_P(\alpha)$-smooth, and thus for every $s \in Mod_P(\alpha)$ there is some $s' \in Mod_P(\alpha)$ such that $(s', s) \in R$ and s' is minimal in $Mod_P(\alpha)$. But then $Min_R(\alpha) \neq \varnothing$. Contradiction. So $\alpha \models_P \bot$. ■

Although not explicitly stated, the entailment relation used in *Left Logical Equivalence* and *Right Weakening* is the entailment relation $\models_P$ determined by the ranked interpretation under consideration. Bochman (2001) refers to such an entailment relation as the *internal logic* determined by a defeasible consequence relation (or *inference relation* in his terminology). The entailment relation $\models_P$ is a variation of the *pivotal-valuation* (monotonic) consequence relations of Makinson (2003b, 2005) that give rise to *default valuation* (nonmonotonic) consequence relations of which rational consequence relations are one example. Makinson calls the logic associated with pivotal-valuation consequence relations *paraclassical* and studies a number of such paraclassical logics (Makinson, 2003b, 2005) that act as natural bridges between classical logic and different kinds of nonmonotonic logics.

For rational consequence relations to satisfy *Consistency Preservation*, a special constraint has to be imposed on ranked interpretations. Following Gärdenfors and Makinson (1994), who used a slightly different formulation, these constrained ranked interpretations will be called *nice*.

Definition 5.24 *A ranked interpretation $P = \langle S, R, l \rangle$ of L is* ***nice*** *when the labelling* *function l is surjective*[13].

Definition 5.25 *An* ***expectation-based*** *consequence relation is a rational consequence relation $\mid\!\sim$ on L that satisfies Consistency Preservation.*

The fundamental idea underlying the notion of expectation-based consequence relations, provided by Gärdenfors and Makinson (1994), is that the reasoning of an agent is guided not only by its firm beliefs, but also by its *expectations*. These expectations come in difference forms, of which *expectation sets* and *expectation relations* are the most pertinent. The focus will be on the latter, which is expressed as an ordering $\sqsubseteq_E$ between sentences with $\alpha \sqsubseteq_E \beta$ interpreted as saying that 'β is at least as expected as α' or 'α is at least as surprising as β'.

Definition 5.26 *A binary relation $\sqsubseteq_E$ on L is an* ***expectation relation*** *iff it satisfies the following postulates:*

(**E1**) *$\sqsubseteq_E$ is transitive*

(**E2**) *If $\alpha \models \beta$ then $\alpha \sqsubseteq_E \beta$*

(**E3**) *For all $\alpha, \beta \in \Gamma$, $\alpha \sqsubseteq_E \alpha \wedge \beta$ or $\beta \sqsubseteq_E \alpha \wedge \beta$*

Expectation orderings can be used to determine defeasible consequence relations. Intuitively, $\alpha \mid\!\sim \beta$ iff β follows logically from α together with the sentences that are 'sufficiently well expected' in the light of α. Note that the postulates for expectation orderings are included in the postulates for epistemic entrenchment orderings (see section 4.5). In the context of belief revision, postulate (**EE4**) is required to relate the entrenchment ordering $\sqsubseteq$ to the specific belief set undergoing revision (or contraction)

[13] *The restriction of the labelling function to a surjective mapping from S to U_T corresponds to a constraint that Makinson and Gärdenfors (1994) call ample.*

while postulate (**EE5**) is required to generate an operation of contraction that satisfies the postulate of recovery. In the context of nonmonotonic reasoning, both of these postulates are superfluous. It turns out that an expectation ordering and an epistemic entrenchment ordering (with respect to some satisfiable belief set) generate the same class of defeasible consequence relations.

Definition 5.27 *A defeasible consequence relation $\mid\!\sim$ of L is a **comparative expectation** consequence relation iff there is an expectation ordering $\sqsubseteq_E$ on L such that $\alpha \mid\!\sim \beta$ iff either $\alpha \models \beta$ or there is some $\gamma \in L$ such that $\alpha \wedge \gamma \models \beta$ and $\neg\alpha \sqsubset_E \gamma$*

Theorem 5.16 (Gärdenfors and Makinson, 1994) *Let $\mid\!\sim$ be any binary relation on L. Then the following conditions are mutually equivalent:*

1. *$\mid\!\sim$ is the comparative expectation consequence relation determined by some expectation ordering $\sqsubseteq_E$ on L*
2. *$\mid\!\sim$ is the defeasible consequence relation induced by some ranked interpretation $P = \langle S, R, l \rangle$ that is nice*
3. *$\mid\!\sim$ is an expectation-based consequence relation*

The theorem establishes a formal relationship between belief revision and nonmonotonic reasoning based on the notion of expectation by viewing the relation of epistemic entrenchment as a kind of expectation ordering. It leads to the conclusion that belief revision and nonmonotonic reasoning are basically the same process, albeit used for different purposes. The process in belief revision is essentially a movement from new information (or evidence) to those models that are minimal with respect to the ordering representing the agent's beliefs (and expectations) and from there to a revised belief set that is true at each of the minimal models. In nonmonotonic reasoning, the process is essentially a movement from a hypothesis to those models that are minimal with respect to the ordering representing the agent's beliefs (and expectations) and from there to a defeasible consequence that is true at each of the minimal models, without changing the agent's beliefs or expectations. The difference is that in belief revision, beliefs are modified as necessary to maintain consistency with the new information, whereas in nonmonotonic reasoning, the hypothesis entertained need not lead to any loss of background beliefs or expectations. Belief revision can therefore be viewed as 'dynamic' and nonmonotonic reasoning as 'static'.

In the context of t-orderings, this view is confirmed, albeit with some subtle differences. The postulates for templated revision, when restricted to belief assertions, correspond precisely to the AGM postulates for revision (of belief sets) via the KM postulates for revision (of knowledge bases) with the exception of postulate (**TR** $*$ **4**). Before investigating a possible translation of postulate (**TR** $*$ **4**) into a postulate for nonmonotonic reasoning, and vice versa, the Makinson-Gärdenfors translation scheme will be adapted for t-orderings as representations of epistemic states.

Definition 5.28 *Let $T = \langle S, t_E, l \rangle$ be a templated interpretation of L such that l is injective and $card(S) = n$ and let $M = \langle S, l \rangle$ be the corresponding extensional interpretation of L. Let $t_E \in T_E$ be an epistemic state. Then for every sentence $\alpha \in L$ that is consistent with $know(t_E)$, $\alpha \mid\!\sim_T \beta$ iff $bel(t_E * \alpha) \models_M \beta$.*

Proposition 5.17 *Let $T = \langle S, t_E, l \rangle$ be a templated interpretation of L such that l is injective and $card(S) = n$ and let $M = \langle S, l \rangle$ be the corresponding extensional interpretation of L. Let $t_E \in T_E$ be an epistemic state. Then Postulate* (**TR** $*$ **4**) *implies Left Logical Equivalence.*

Proof. From postulate (**TR** $*$ **4**) it follows that if $\alpha \equiv_M \gamma$ then $bel(t_E * \alpha) \equiv_M bel(t_E * \gamma)$ given that α that is consistent with $know(t_E)$. Suppose that $bel(t_E * \alpha) \models_M \beta$

and $\alpha \equiv_M \gamma$. But then $bel(t_E * \gamma) \models_M \beta$. From the Makinson-Gärdenfors translation, it follows that if $\alpha \mid\!\sim_T \beta$ and $\alpha \equiv_M \gamma$ then $\gamma \mid\!\sim_T \beta$, which is the *Left Logical Equivalence* postulate. ■

The translation of *Left Logical Equivalence* into postulate ($\mathbf{TR} * \mathbf{4}$) does not hold because the equivalence of two belief assertions $bel(t_E * \alpha)$ and $bel(t_E * \gamma)$, given that $\alpha \equiv_M \gamma$, does not guarantee that the t-orderings $t_E * \alpha$ and $t_E * \gamma$ are equal as required by postulate ($\mathbf{TR} * \mathbf{4}$). The breakdown of the translation can be attributed to the fact that templated revision operates at the level of epistemic states rather than at the level of belief sets. For iterated revision, which typically operates at the level of epistemic states, the correspondence between iterated nonmonotonic reasoning and iterated revision appears to be unclear even at the level of belief sets, as pointed out by Makinson (2003a).

Example 5.8 (Makinson, 2003a) *An iterated revision (at the level of belief sets) is typically something like* $(K * \alpha) * \beta$ *while an iterated defeasible consequence is typically something like* $\beta \mid\!\sim (\alpha \mid\!\sim \gamma)$ *or* $(\beta \mid\!\sim \alpha) \mid\!\sim \gamma$. *The translation of* $\gamma \in (K * \alpha) * \beta$ *into the language of nonmonotonic reasoning would be* $\beta \mid\!\sim_J \gamma$ *where* $J = Cn(\alpha) = \{\delta \mid \alpha \mid\!\sim_K \delta\}$. *On the other hand, the translation of* $\beta \mid\!\sim (\alpha \mid\!\sim \gamma)$ *into the language of belief revision would be* $(\alpha \mid\!\sim \gamma) \in K * \beta$, *and the translation of* $(\beta \mid\!\sim \alpha) \mid\!\sim \gamma$ *would be* $\gamma \in K * (\beta \mid\!\sim \alpha)$. *It is not clear how these relate to each other - nor what they mean individually.*

Despite the formal map between belief revision and nonmonotonic reasoning, Makinson (2003a) cautions against an over-emphasis on formal translations that may obscure underlying differences of gestalt and intuition, such as illustrated by the case of iterated belief revision.

Another difference between belief revision and nonmonotonic reasoning, perhaps not of gestalt, is illustrated by proposition 5.10, which shows that in templated revision, the new information (or evidence) becomes more plausible in the revised epistemic state than it was in the original epistemic state. On the other hand, in nonmonotonic reasoning, the plausibility of the hypothesis does not change as the result of a defeasible consequence thus confirming the view that belief revision is ‘dynamic’ and nonmonotonic reasoning ‘static’.

Chapter 6

Merging

'On the Art of Combination'
De Arte Combinatoria (Gottfried Wilhelm Leibniz)

6.1 Knowledge base merging

Different proposals for *knowledge base merging* have been presented, many of which are based on different intuitions. The focus of this section will be on propositional knowledge bases as defined by Katsuno and Mendelzon (1991) and not on (first-order) knowledge bases described as logic programs (or deductive databases) with integrity constraints (Baral, Kraus, and Minker, 1991; Baral, et al., 1992; Subrahmanian, 1994). The proposals for knowledge base merging that will be reviewed are all based on finitely generated propositional languages under a traditional truth-value semantics. As mentioned before, the traditional truth-value semantics of propositional logic corresponds to the classical interpretation $M_0 = \langle S, l \rangle$, which will henceforth be assumed. The focus of the review is on the logical properties of merging.

The notion of *arbitration* is due to Revesz (1993), who defined arbitration as an alternative form of belief change. In the epistemic change operations considered thus far, the new information has been taken very seriously, to the extent that the agent is willing to sacrifice previously held beliefs or knowledge in order successfully to incorporate the new information. The key idea we now consider is that when the new information is considered neither better nor worse than (the information represented by) the existing belief set, then a form of 'merging' is called for. Using the formulation of Katsuno and Mendelzon for knowledge bases, arbitration is defined in terms of an auxiliary operation called *model-fitting*, which is a binary operation on L that 'fits' the knowledge base to the new information.

Revesz (1997) proposes the following seven postulates for model-fitting.

Definition 6.1 *Let κ be a knowledge base of L. A* ***model-fitting operation*** $\rhd$ *is an operation such that for every sentence $\alpha \in L$, the knowledge base $\kappa \rhd \alpha$, which is the result of model-fitting κ to α, satisfies the following set of postulates:*

($\mathbf{R} \rhd \mathbf{1}$) $\kappa \rhd \alpha \models \alpha$

($\mathbf{R} \rhd \mathbf{2}$) *If $\kappa \wedge \alpha$ is satisfiable, then $\kappa \rhd \alpha \equiv \kappa \wedge \alpha$*

($\mathbf{R} \rhd \mathbf{3}$) *If α is satisfiable, then $\kappa \rhd \alpha$ is satisfiable*

($\mathbf{R} \rhd \mathbf{4}$) *If $\kappa \equiv \kappa'$ and $\alpha \equiv \alpha'$ then $\kappa \rhd \alpha \equiv \kappa' \rhd \alpha'$*

($\mathbf{R} \rhd \mathbf{5}$) $(\kappa \rhd \alpha) \wedge \beta \models \kappa \rhd (\alpha \wedge \beta)$

($\mathbf{R} \rhd \mathbf{6}$) *If $(\kappa \rhd \alpha) \wedge \beta$ is satisfiable, then $\kappa \rhd (\alpha \wedge \beta) \models (\kappa \rhd \alpha) \wedge \beta$*

($\mathbf{R} \rhd \mathbf{7}$) $(\kappa \rhd \alpha) \wedge (\kappa' \rhd \alpha) \models (\kappa \vee \kappa') \rhd \alpha$

Postulates ($\mathbf{R} \rhd \mathbf{1}$) to ($\mathbf{R} \rhd \mathbf{6}$) correspond directly to postulates ($\mathbf{KM} * \mathbf{1}$) to ($\mathbf{KM} * \mathbf{6}$) for revision. Postulate ($\mathbf{R} \rhd \mathbf{7}$) says that any model that is closest to both κ in α and κ' in α must be a closest model to $\kappa \vee \kappa'$. The converse of postulate ($\mathbf{R} \rhd \mathbf{7}$) appears, amongst other differences, as an additional postulate in the original formulation of model-fitting operations (Revesz, 1993) and is stated here as postulate ($\mathbf{R} \rhd \mathbf{8}$).

($\mathbf{R} \rhd \mathbf{8}$) *if* $(\kappa \rhd \alpha) \wedge (\kappa' \rhd \alpha)$ *is satisfiable, then* $(\kappa \vee \kappa') \rhd \alpha \models (\kappa \rhd \alpha) \wedge (\kappa' \rhd \alpha)$

Together, postulates ($\mathbf{R} \rhd \mathbf{7}$) and ($\mathbf{R} \rhd \mathbf{8}$) say that the closest models to $\kappa \vee \kappa'$ are the intersection of the closest model to κ in α and to κ' in α, provided the intersection is nonempty. Arbitration is originally defined as a special case of model-fitting, namely, $\kappa \bigtriangleup \alpha \stackrel{def}{=} (\kappa \vee \alpha) \rhd \top$. However, in Revesz (1997) arbitration is described as a generalisation of weighted model-fitting. The intuition is that a knowledge base (comprising individual sentences, say κ and λ) can be tested, or 'model-fitted', against several possible hypotheses (represented by a single sentence, say α) without actually changing the knowledge base. In this scenario, arbitration resembles hypothetical querying rather than a form of belief change.

Liberatore and Schaerf (1998) give an alternative formulation for arbitration as a form of belief change. Their approach rests on the intuition that if there are two different sources of information, each having a different view of the situation and neither completely unreliable, then the best an agent can do is to 'merge' the two views into one that is consistent, whilst trying to preserve as much information as possible. They call this merging process *arbitration*. A knowledge base is defined as a finite set of sentences, represented by a single sentence κ. Arbitration is essentially a binary operation on L that merges two knowledge bases into a new knowledge base.

The following eight postulates for arbitration are proposed by Liberatore and Schaerf (1998).

Definition 6.2 *Let κ and λ be two knowledge bases of L. An* ***arbitration operation*** *$\bigtriangleup$ is an operation such that the knowledge base $\kappa \bigtriangleup \lambda$, which is the result of arbitrating between κ and λ, satisfies the following set of postulates:*

(**LS** $\triangle$ **1**) $\kappa \triangle \lambda \equiv \lambda \triangle \kappa$

(**LS** $\triangle$ **2**) $\kappa \wedge \lambda \models \kappa \triangle \lambda$

(**LS** $\triangle$ **3**) *If* $\kappa \wedge \lambda$ *is satisfiable, then* $\kappa \triangle \lambda \models \kappa \wedge \lambda$

(**LS** $\triangle$ **4**) $\kappa \triangle \lambda$ *is unsatisfiable iff both* κ *and* λ *are unsatisfiable*

(**LS** $\triangle$ **5**) *If* $\kappa \equiv \kappa'$ *and* $\lambda \equiv \lambda'$ *then* $\kappa \triangle \lambda \equiv \kappa' \triangle \lambda'$

(**LS** $\triangle$ **6**) $\kappa \triangle (\lambda \vee \mu) \equiv \begin{cases} \kappa \triangle \lambda & \textit{or} \\ \kappa \triangle \mu & \textit{or} \\ (\kappa \triangle \lambda) \vee (\kappa \triangle \mu) \end{cases}$

(**LS** $\triangle$ **7**) $\kappa \triangle \lambda \models \kappa \vee \lambda$

(**LS** $\triangle$ **8**) *If* κ *is satisfiable, then* $\kappa \wedge (\kappa \triangle \lambda)$ *is satisfiable*

Postulate (**LS** $\triangle$ **1**) ensures that arbitration is commutative. Postulate (**LS** $\triangle$ **2**) ensures that the result of an arbitration contains only information present in either of the knowledge bases while postulates (**LS** $\triangle$ **3**) and (**LS** $\triangle$ **7**) ensure that as much as possible of the information is retained. Postulate (**LS** $\triangle$ **4**) says that the result of an arbitration should be unsatisfiable if and only if both the knowledge bases are unsatisfiable. Postulate (**LS** $\triangle$ **5**) ensures that arbitration adheres to the Principle of Irrelevance of Syntax. Postulate (**LS** $\triangle$ **6**) guarantees that arbitration of composite knowledge bases can be obtained by composition of the arbitration of sub-knowledge bases. Postulate (**LS** $\triangle$ **8**) ensures that both knowledge bases contribute to the result of an arbitration. The basic properties that all arbitration operations should satisfy are postulates (**LS** $\triangle$ **1**) to (**LS** $\triangle$ **5**).

Arbitration is generally seen as one subclass of merging operations, the other important subclass being that of majority merging. Lin and Mendelzon (1999) were the first to formalise the notion of merging knowledge bases *by majority*, based on earlier work by Lin (1996). The intuition behind their approach is that the problem of merging the knowledge of multiple agents is the same as the problem of merging multiple knowledge bases. A knowledge base is defined as a finite set of sentences, which will be represented by a single sentence κ. Majority merging is an operation that takes a set of m knowledge bases and merges it into a single knowledge base in such a way that the view of the majority is reflected. Since the result of majority merging does not depend on any order of the knowledge bases in the set, majority merging is commutative by definition. To ease comparison with other approaches, majority merging will be restricted to two knowledge bases and treated as a binary operation on L. However, to ensure that majority merging as a binary operation is commutative, the following postulate is added to those proposed by Lin and Mendelzon.

(LM $\triangle$ 0) $\kappa \triangle \lambda \equiv \lambda \triangle \kappa$

The approach of Lin and Mendelzon relies on a notion of 'partial support'. A knowledge base κ is said to *support* a sentence α if $\kappa \models \alpha$ and to *oppose* α if $\kappa \models \neg\alpha$. The notion of partial support applies only to literals. A knowledge base κ is said to *partially support* a literal α, denoted by $\kappa \mathrel{|\!\!\approx} \alpha$, if there exists a sentence β, which mentions no atom appearing in α, such that $\kappa \models \alpha \vee \beta$ but $\kappa \not\models \alpha$ and $\kappa \not\models \beta$. Model-theoretically, $\kappa \mathrel{|\!\!\approx} \alpha$ iff $Mod(\kappa \wedge \neg\alpha) \neq \varnothing$ and there exists a state $s \in Mod(\kappa \wedge \alpha)$ such that $\bar{s} \notin Mod(\kappa)$, where $\bar{s}$ denotes the state that agrees with s on every atom except the one appearing in literal α.

Proposition 6.1 (Lin and Mendelzon, 1999) *For a knowledge base κ and a literal α, if $\kappa \models \alpha$ or $\kappa \models \neg\alpha$ or $\kappa \wedge \alpha$ is unsatisfiable, then $\kappa \mathrel{|\!\!\not\approx} \alpha$.*

Lin and Mendelzon (1999) propose the following four postulates for knowledge base merging by majority.

Definition 6.3 *Let κ and λ be knowledge bases of L. A* ***majority merging operation*** *$\triangle$ is an operation such that the knowledge base $\kappa \triangle \lambda$, which is the result of merging κ and λ by majority, satisfies the following set of postulates:*

(**LM** $\triangle$ **1**) *$\kappa \triangle \lambda$ is satisfiable*

(**LM** $\triangle$ **2**) *If $\kappa \wedge \lambda$ is satisfiable, then $\kappa \triangle \lambda \equiv \kappa \wedge \lambda$*

(**LM** $\triangle$ **3**) *If $\kappa \equiv \kappa'$ then $\kappa \triangle \lambda \equiv \kappa' \triangle \lambda$*

(**LM** $\triangle$ **4**) *For a literal $\alpha \in L$ and $\Gamma = \{\kappa, \lambda\}$, if $card(\{\beta \in \Gamma \mid \beta \models \alpha\}) > card(\{\beta \in \Gamma \mid \beta \models \neg\alpha\}) + card(\{\beta \in \Gamma \mid \beta \approx\!\!\!\!\models \neg\alpha\})$ then $\kappa \triangle \lambda \models \alpha$*

Postulate (**LM** $\triangle$ **1**) ensures that the result of merging is satisfiable (even if the knowledge bases being merged are unsatisfiable). Postulate (**LM** $\triangle$ **2**) says that if there is no conflict among the knowledge bases, then the result of merging is simply the conjunction of the knowledge bases. Postulate (**LM** $\triangle$ **3**) ensures that merging adheres to the Principle of Irrelevance of Syntax. Whilst postulates (**LM** $\triangle$ **1**) to (**LM** $\triangle$ **3**) apply to any merging operation, postulate (**LM** $\triangle$ **4**) applies only to merging operations that obey the majority principle. Postulate (**LM** $\triangle$ **4**) says that if the (combined) support for a literal is greater than the (combined) support and partial support for the negation of the literal, then the result of merging by majority should support the literal.

With their focus on the problem of merging the knowledge of multiple agents, Lin and Mendelzon review several protocols for resolving conflicts in a group of agents, in particular, the protocols of Borgida and Imielinski (1984) of which the 'democracy' protocol proved to be closest to the principle of majority merging.

Konieczny and Pino-Pérez (1998) make a clear distinction between majority merging and arbitration by defining both as subclasses of (pure) merging operations. A knowledge base is defined as a finite set of sentences, represented by a single sentence κ, and is assumed to be satisfiable. Merging is an operation that maps a multiset of m knowledge bases to a new knowledge base. When using multisets, the result of merging does not depend on any order of the knowledge bases in the multiset, so that merging is commutative by definition. As before, to ease comparison, a merging operation is restricted to two knowledge bases and treated as a binary operation on L. The following postulate is added to those proposed by Konieczny and Pino-Pérez to ensure that merging as a binary operation is commutative.

($\mathbf{KP} \triangle \mathbf{0}$) $\kappa \triangle \lambda \equiv \lambda \triangle \kappa$

The following six postulates are proposed for (pure) merging operations by Konieczny and Pino-Pérez (1998).

Definition 6.4 *Let κ and λ be satisfiable knowledge bases of L. A* ***merging operation*** *$\triangle$ is an operation such that the knowledge base $\kappa \triangle \lambda$, which is the result of merging κ and λ, satisfies the following set of postulates:*

($\mathbf{KP} \triangle \mathbf{1}$) $\kappa \triangle \lambda$ *is satisfiable*

($\mathbf{KP} \triangle \mathbf{2}$) *If $\kappa \wedge \lambda$ is satisfiable, then $\kappa \triangle \lambda \equiv \kappa \wedge \lambda$*

($\mathbf{KP} \triangle \mathbf{3}$) *If $\kappa \equiv \kappa'$ then $\kappa \triangle \lambda \equiv \kappa' \triangle \lambda$*

($\mathbf{KP} \triangle \mathbf{4}$) *If $\kappa \wedge \lambda$ is unsatisfiable, then $\kappa \triangle \lambda \not\models \kappa$*

($\mathbf{KP} \triangle \mathbf{5}$) $(\kappa \triangle \lambda) \wedge (\kappa' \triangle \lambda) \models (\kappa \vee \kappa') \triangle \lambda$

($\mathbf{KP} \triangle \mathbf{6}$) *If $(\kappa \triangle \lambda) \wedge (\kappa' \triangle \lambda)$ is satisfiable, then $(\kappa \vee \kappa') \triangle \lambda \models (\kappa \triangle \lambda) \wedge (\kappa' \triangle \lambda)$*

Postulate ($\mathbf{KP} \triangle \mathbf{1}$) says that the result of merging is satisfiable (under the assumption that the knowledge bases being merged are satisfiable). Postulates ($\mathbf{KP} \triangle \mathbf{2}$) and ($\mathbf{KP} \triangle \mathbf{3}$) are identical to postulates ($\mathbf{LM} \triangle \mathbf{2}$) and ($\mathbf{LM} \triangle \mathbf{3}$) of Lin and Mendelzon respectively. Postulate ($\mathbf{KP} \triangle \mathbf{4}$) is called the *Fairness* postulate and ensures that when two knowledge bases are merged, preference is given to neither. Postulates ($\mathbf{KP} \bigtriangleup \mathbf{5}$) and ($\mathbf{KP} \bigtriangleup \mathbf{6}$) correspond directly to postulates ($\mathbf{R} \rhd \mathbf{7}$) and ($\mathbf{R} \rhd \mathbf{8}$) for model-fitting.

Konieczny and Pino-Pérez propose two additional postulates to distinguish between majority merging and arbitration. The postulate for majority merging says that if a knowledge base is included enough times in a multiset, then the result of majority merging will entail the knowledge base while the postulate for arbitration says that the result of arbitration is, to a large extent, independent of the frequency with which a knowledge base appears in a multiset. These postulates cannot be translated in terms of binary operations but the formulation in terms of a multiset Γ of knowledge bases is included here as postulate ($\mathbf{KP} \bigtriangleup \mathbf{7}$) for majority merging and postulate ($\mathbf{KP} \bigtriangleup \mathbf{8}$) for arbitration.

($\mathbf{KP} \bigtriangleup \mathbf{7}$) $\forall \kappa \; \exists m \; \bigtriangleup(\Gamma \sqcup \kappa^m) \models \kappa$

($\mathbf{KP} \bigtriangleup \mathbf{8}$) $\forall \kappa' \; \exists \kappa \; \kappa' \not\models \kappa \; \forall m \; \bigtriangleup(\kappa' \sqcup \kappa^m) \equiv \bigtriangleup(\kappa' \sqcup \kappa)$

The original formulation for arbitration, which is stated below as postulate ($\mathbf{KP} \bigtriangleup \mathbf{8'}$), has been shown to be inconsistent with postulates ($\mathbf{KP} \bigtriangleup \mathbf{0}$) to ($\mathbf{KP} \bigtriangleup \mathbf{6}$) by Paolo Liberatore (Konieczny and Pino-Pérez, 1998).

($\mathbf{KP} \bigtriangleup \mathbf{8'}$) $\forall \kappa \; \forall m \; \bigtriangleup(\Gamma \sqcup \kappa^m) \equiv \bigtriangleup(\Gamma \sqcup \kappa)$

The notion of merging is extended by Konieczny and Pino-Pérez (1999) to merging with integrity constraints, also referred to as *IC merging*. When the integrity constraint is taken to be a tautology, IC merging is compatible with 'pure' merging (Konieczny and Pino-Pérez, 2002). The main differences are that postulates ($\mathbf{KP} \bigtriangleup 4$) and ($\mathbf{KP} \bigtriangleup \mathbf{7}$) are weaker than their counterparts for IC merging while the IC merging counterpart for postulate ($\mathbf{KP} \bigtriangleup \mathbf{8}$) is described as a rule expressing non-majority, as opposed to a rule characterising arbitration. However, the proposed postulate for IC arbitration is not expressible without integrity constraints.

The connection between IC merging operations and the operations of Revesz, Liberatore and Schaerf, and Lin and Mendelzon is studied to varying degree by Konieczny and Pino-Pérez (2002). The arbitration operations of Liberatore and Schaerf are investigated as a special case of IC merging operations where the integrity constraint is taken to be the disjunction of the knowledge bases. Provided certain properties hold, it is shown that an IC merging operation defined in this manner satisfies postulates ($\mathbf{LS} \triangle \mathbf{1}$) to ($\mathbf{LS} \triangle \mathbf{8}$). However, the proposed postulate for IC arbitration is not considered in the investigation. The model-fitting operations of Revesz, which are very close to IC merging operations, and the majority merging operations of Lin and Mendelzon are only briefly addressed.

In the next section, merging operations will be considered in more detail. However, it is useful to compare the four proposals for merging operations within the context of a uniform representation, accepting the assumption of Konieczny and Pino-Pérez that merging only applies to knowledge bases that are satisfiable. In contrast to Konieczny and Pino-Pérez (2002), our comparison focusses on 'pure' merging (as opposed to IC merging) and on those postulates which are regarded as 'basic'.

Lin and Mendelzon's proposal that postulates ($\mathbf{LM} \triangle \mathbf{1}$) to ($\mathbf{LM} \triangle \mathbf{3}$) should apply to any merging operation appears to be justified. Postulate ($\mathbf{LM} \triangle \mathbf{1}$) correspond directly to postulates ($\mathbf{KP} \triangle \mathbf{1}$), ($\mathbf{LS} \triangle \mathbf{4}$), and ($\mathbf{R} \rhd \mathbf{3}$); while postulate ($\mathbf{LM} \triangle \mathbf{2}$) correspond directly to postulates ($\mathbf{KP} \triangle \mathbf{2}$), ($\mathbf{LS} \triangle \mathbf{2}$) together with ($\mathbf{LS} \triangle \mathbf{3}$), and ($\mathbf{R} \rhd \mathbf{2}$); with postulate ($\mathbf{LM} \triangle \mathbf{3}$) corresponding directly to postulates ($\mathbf{KP} \triangle \mathbf{3}$), ($\mathbf{LS} \triangle \mathbf{5}$), and ($\mathbf{R} \rhd \mathbf{4}$).

Commutativity is also strongly supported. Postulate ($\mathbf{LM} \triangle \mathbf{0}$) corresponds directly to postulates ($\mathbf{KP} \triangle \mathbf{0}$) and ($\mathbf{LS} \triangle \mathbf{1}$) and is satisfied by every model-fitting operation of Revesz if $\kappa \wedge \lambda$ is satisfiable, a result which follows directly by postulate ($\mathbf{R} \rhd \mathbf{2}$) and the commutativity of conjunction. However, if $\kappa \wedge \lambda$ is unsatisfiable, then postulate ($\mathbf{LM} \triangle \mathbf{0}$) is not satisfied by every model-fitting operation of Revesz, as shown by the following example.

Example 6.1 *Suppose L is generated by $Atom = \{P(a), P'(a)\}$. Let $\triangleright$ be any model-fitting operation of Revesz, i.e. $\triangleright$ satisfies postulates* $(\mathbf{R} \triangleright \mathbf{1})$ *to* $(\mathbf{R} \triangleright \mathbf{7})$*. Let $Mod(\kappa) = \{11\}$ and $Mod(\lambda) = \{10\}$. By postulate* $(\mathbf{R} \triangleright \mathbf{1})$ *it follows that $Mod(\kappa \triangleright \lambda) \subseteq \{10\}$ and $Mod(\lambda \triangleright \kappa) \subseteq \{11\}$. Since neither may be empty by postulate* $(\mathbf{R} \triangleright \mathbf{3})$*, it follows that $Mod(\kappa \triangleright \lambda) = \{10\}$ and $Mod(\lambda \triangleright \kappa) = \{11\}$ and hence $Mod(\kappa \triangleright \lambda) \neq Mod(\lambda \triangleright \kappa)$.*

The only other 'basic' postulates are postulates $(\mathbf{KP} \triangle \mathbf{4})$, $(\mathbf{KP} \triangle \mathbf{5})$, and $(\mathbf{KP} \triangle \mathbf{6})$ of Konieczny and Pino-Pérez. They are compared with the model-fitting operations of Revesz, the arbitration operations of Liberatore and Schaerf, and the majority merging operations of Lin and Mendelzon in the following results.

Proposition 6.2 *Let κ and λ be satisfiable knowledge bases of L. Postulate* $(\mathbf{KP} \triangle \mathbf{4})$ *is satisfied by every model-fitting operation of Revesz, not satisfied by every majority merging operation of Lin and Mendelzon, and not satisfied by every arbitration operation of Liberatore and Schaerf.*

Proof. See proof in appendix D, section D.1. ■

Proposition 6.3 *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* $(\mathbf{KP} \triangle \mathbf{5})$ *is satisfied by every arbitration operation of Liberatore and Schaerf but is not satisfied by every majority merging operation of Lin and Mendelzon.*

Proof. See proof in appendix D, section D.1. ■

Postulate $(\mathbf{KP} \triangle \mathbf{5})$ corresponds directly to postulate $(\mathbf{R} \triangleright \mathbf{7})$. Although not satisfied by every majority merging operation of Lin and Mendelzon, postulate $(\mathbf{KP} \triangle \mathbf{5})$ is satisfied if $\kappa \wedge \lambda$ or $\kappa' \wedge \lambda$ is satisfiable, as shown by the following result.

Proposition 6.4 *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* (**KP** $\triangle$ **5**) *is satisfied by every majority merging operation of Lin and Mendelzon, provided $\kappa \wedge \lambda$ or $\kappa' \wedge \lambda$ is satisfiable.*

Proof. See proof in appendix D, section D.1. ■

Proposition 6.5 *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* (**KP** $\triangle$ **6**) *is not satisfied by every model-fitting operation of Revesz, not satisfied by every arbitration operation of Liberatore and Schaerf, and not satisfied by every majority merging operation of Lin and Mendelzon.*

Proof. See proof in appendix D, section D.1. ■

The comparison shows that postulates (**KP** $\triangle$ **1**) to (**KP** $\triangle$ **3**) are uncontroversial as logical properties of merging. There is strong support for postulates (**KP** $\triangle$ **0**) and (**KP** $\triangle$ **5**) as basic properties of merging but less so for postulates (**KP** $\triangle$ **6**) and (**KP** $\triangle$ **4**).

In all of these approaches a semantic characterisation of merging operations is provided in terms of preorders on the set of possible worlds S (or $\wp(S)$ in the case of Liberatore and Schaerf) satisfying different constraints. This leads to semantic methods for constructing merging operations whereby the result of merging is a knowledge base the models of which are minimal in some sense with respect to the particular preorder.

6.2 Some merging operations

Most of the semantic methods for constructing (knowledge base) merging operations rely on a notion of distance between possible worlds. A popular choice is the distance function of Dalal (1988), also known as the Hamming distance, which is a function $d : S \times S \to \mathbb{N}$ such that $d(s, s')$ is the number of atoms that have different truth values under s and s'.

More abstractly, Konieczny and Pino-Pérez (2002) define the distance between possible worlds as a function $d : S \times S \to \mathbb{N}$ such that $d(s, s') = d(s', s)$, and $d(s, s') = 0$ iff $s = s'$.

The distance between possible worlds is used to define a distance between a possible world s and a knowledge base κ as the minimal distance between s and the models of κ as $d(s, \kappa) = \min\{d(s, s') \mid s' \in Mod(\kappa)\}$. From the distance between a possible world and an individual knowledge base, an *overall distance* between a possible world s and a multiset Γ of knowledge bases to be merged can be defined using different functions $\odot$ as $d_{\odot}(s, \Gamma) = \odot\{d(s, \kappa) \mid \kappa \in \Gamma\}$.

Merging operations based on the notion of distance between possible worlds differ primarily in their choice of the function $\odot$. Irrespective of the choice of $\odot$, the overall distance between a possible world s and the multiset Γ of knowledge bases to be merged can be used to induce a pre-order $\leq_{\Gamma,\odot}$ on the set of possible worlds S (with respect to Γ) as $s \leq_{\Gamma,\odot} s'$ iff $d_{\odot}(s, \Gamma) \leq d_{\odot}(s', \Gamma)$. This ensures that there is an assignment that maps each multiset Γ of knowledge bases to a preorder $\leq_{\Gamma,\odot}$ on S. The result of merging a multiset Γ of knowledge bases can then be obtained by taking as the models of $\triangle_{\odot}(\Gamma)$ the minimal models of $\top$ with respect to $\leq_{\Gamma,\odot}$, i.e. $Mod(\triangle_{\odot}(\Gamma)) = Min_{\leq_{\Gamma,\odot}}(\top)$.

Provided the assignment satisfies the conditions associated with a specific set of postulates for merging operations, the corresponding representation theorem can then be used to prove that the merging operation $\triangle_{\odot}$ satisfies the set of postulates.

Three families of merging operations are considered. The first family of merging operations, called Σ operations, uses the sum of the distances between a possible world and the individual knowledge bases as their function $\odot$.

Definition 6.5 *Let Γ be a multiset of satisfiable knowledge bases of L and let $s, s' \in S$. Then the merging operation $\triangle_{\Sigma}$ is defined by taking $d_{\Sigma}(s, \Gamma) = \sum_{\kappa \in \Gamma} d(s, \kappa)$ and $s \leq_{\Gamma,\Sigma} s'$ iff $d_{\Sigma}(s, \Gamma) \leq d_{\Sigma}(s', \Gamma)$ so that $Mod(\triangle_{\Sigma}(\Gamma)) = Min_{\leq_{\Gamma,\Sigma}}(\top)$.*

The merging operation $\triangle_{\Sigma}$ is given by Lin and Mendelzon (1999) as an example of a majority merging operation and independently, in a slightly different format, by Revesz (1993) as an example of a weighted model-fitting operation. Both examples are based on the Hamming distance between possible worlds. $\triangle_{\Sigma}$ is also a majority merging operation in the sense of Konieczny and Pino-Pérez (1998).

Proposition 6.6 (Lin and Mendelzon, 1999) $\triangle_{\Sigma}$ *satisfies postulates* $(\mathbf{LM} \triangle \mathbf{1})$ *to* $(\mathbf{LM} \triangle \mathbf{4})$*.*

Proposition 6.7 (Konieczny and Pino-Pérez, 1998) $\triangle_{\Sigma}$ *satisfies postulates* $(\mathbf{KP} \triangle \mathbf{1})$ *to* $(\mathbf{KP} \triangle \mathbf{6})$ *and* $(\mathbf{KP} \triangle \mathbf{7})$*.*

The second family of merging operations, called Max operations, uses the maximum of the distances between a possible world and the individual knowledge bases as their function $\odot$.

Definition 6.6 *Let* Γ *be a multiset of satisfiable knowledge bases of* L *and let* $s, s' \in S$*. Then the merging operation* $\triangle_{Max}$ *is defined by taking* $d_{Max}(s, \Gamma) = \max_{\kappa \in \Gamma} d(s, \kappa)$ *and* $s \leq_{\Gamma, Max} s'$ *iff* $d_{Max}(s, \Gamma) \leq d_{Max}(s', \Gamma)$ *so that* $Mod(\triangle_{Max}(\Gamma)) = Min_{\leq_{\Gamma, Max}}(\top)$*.*

The merging operation $\triangle_{Max}$ is given by Revesz (1993) as an example of a model-fitting operation, again based on the Hamming distance between possible worlds. $\triangle_{Max}$ is regarded as a quasi-merging operation by Konieczny and Pino-Pérez (2002) because it does not satisfy postulate $(\mathbf{KP} \triangle \mathbf{6})$. Although it satisfies postulate $(\mathbf{KP} \triangle \mathbf{8})$, Konieczny and Pino-Pérez's postulate for arbitration, it is not an arbitration operation in the sense of Liberatore and Schaerf.

Proposition 6.8 (Konieczny and Pino-Pérez, 1998) $\triangle_{Max}$ *satisfies postulates* $(\mathbf{KP} \triangle \mathbf{1})$ *to* $(\mathbf{KP} \triangle \mathbf{5})$ *and* $(\mathbf{KP} \triangle \mathbf{8})$*, but not postulate* $(\mathbf{KP} \triangle \mathbf{6})$*.*

Proposition 6.9 *Let κ, λ, and μ be satisfiable knowledge bases of L. $\triangle_{Max}$ satisfies postulates* (**LS $\triangle$ 1**) *to* (**LS $\triangle$ 6**)*, but fails to satisfy postulates* (**LS $\triangle$ 7**) *and* (**LS $\triangle$ 8**)*.*

Proof. Postulates (**LS $\triangle$ 1**) to (**LS $\triangle$ 5**) are satisfied through the correspondence with postulates (**KP $\triangle$ 0**) to (**KP $\triangle$ 3**) and proposition 6.8.

(**LS $\triangle$ 6**) It must be shown that $Mod(\kappa\triangle_{Max}(\lambda\vee\mu)) = Mod(\kappa\triangle_{Max}\lambda)$ or $Mod(\kappa\triangle_{Max}\mu)$ or $Mod(\kappa \triangle_{Max} \lambda)\cup Mod(\kappa \triangle_{Max} \mu)$. Choose any $s \in Mod(\kappa \triangle_{Max} (\lambda \vee \mu))$. So there is no $s' \in Mod(\top)$ such that $s' \leq_{\{\kappa,\lambda\vee\mu\},Max} s$ unless $s \leq_{\{\kappa,\lambda\vee\mu\},Max} s'$. In other words, for every $s' \in Mod(\top)$ it holds that $d_{Max}(s, \{\kappa, \lambda \vee \mu\}) \leq d_{Max}(s', \{\kappa, \lambda \vee \mu\})$, i.e. $\max\{d(s, \kappa), d(s, \lambda \vee \mu)\} \leq \max\{d(s', \kappa), d(s', \lambda \vee \mu)\}$. Suppose that $s \notin Mod(\kappa \triangle_{Max} \lambda)$ and $s \notin Mod(\kappa \triangle_{Max} \mu)$ and therefore also $s \notin (Mod(\kappa \triangle_{Max} \lambda) \cup Mod(\kappa \triangle_{Max} \mu)$. Since $s \notin Mod(\kappa \triangle_{Max} \lambda)$ it follows that there is some $s'' \in Mod(\top)$ such that $\max\{d(s'', \kappa), d(s'', \lambda)\} < \max\{d(s, \kappa), d(s, \lambda)\}$. Suppose that $\max\{d(s, \kappa), d(s, \lambda)\} = d(s, \kappa)$. So $d(s'', \kappa) < d(s, \kappa)$ and $d(s'', \lambda) < d(s, \kappa)$ and $d(s, \lambda) \leq d(s, \kappa)$. But $d(s'', \lambda \vee \mu) \leq d(s'', \lambda)$ and $d(s, \lambda\vee\mu) \leq d(s, \lambda)$. So $\max\{d(s'', \kappa), d(s'', \lambda\vee\mu)\} < \max\{d(s, \kappa), d(s, \lambda\vee \mu)\}$. Contradiction. Now suppose that $\max\{d(s, \kappa), d(s, \lambda)\} = d(s, \lambda)$. So $d(s'', \kappa) \leq d(s, \lambda)$ and $d(s'', \lambda) \leq d(s, \lambda)$ and $d(s, \kappa) \leq d(s, \lambda)$. If $d(s, \lambda \vee \mu) = d(s, \lambda)$ then $\max\{d(s'', \kappa), d(s'', \lambda \vee \mu)\} \leq \max\{d(s'', \kappa), d(s'', \lambda)\} < \max\{d(s, \kappa), d(s, \lambda)\} =$ $\max\{d(s, \kappa), d(s, \lambda \vee \mu)\}$, resulting in a contradiction. Otherwise $d(s, \lambda \vee \mu) < d(s, \lambda)$ and thus $d(s, \lambda \vee \mu) = d(s, \mu)$. Since $s \notin Mod(\kappa \triangle_{Max} \mu)$ it follows that there is some $s''' \in Mod(\top)$ such that $\max\{d(s''', \kappa), d(s''', \mu)\} \leq \max\{d(s, \kappa), d(s, \mu)\}$. Since $d(s''', \lambda \vee \mu) \leq d(s''', \mu)$ and $d(s, \lambda \vee \mu) = d(s, \mu)$ it follows that $\max\{d(s''', \kappa), d(s''', \lambda \vee \mu)\} \leq \max\{d(s, \kappa), d(s, \lambda \vee \mu)\}$. Contradiction. But then it must be the case that $s \in Mod(\kappa \triangle_{Max} \lambda)$ or $s \in Mod(\kappa \triangle_{Max} \mu)$ or $s \in (Mod(\kappa \triangle_{Max} \lambda) \cup Mod(\kappa \triangle_{Max} \mu)$. Since s was chosen arbitrarily, the result holds in one direction.

Conversely, choose any $s \in Mod(\kappa\triangle_{Max}\lambda)$ or $s \in Mod(\kappa\triangle_{Max}\mu)$ or $s \in (Mod(\kappa\triangle_{Max} \lambda) \cup Mod(\kappa \triangle_{Max} \mu)$. So for every $s' \in Mod(\top)$ it holds that $\max\{d(s, \kappa), d(s, \lambda)\} \leq$

$\max\{d(s', \kappa), d(s', \lambda)\}$ or $\max\{d(s, \kappa), d(s, \mu)\} \leq \max\{d(s', \kappa), d(s', \mu)\}$ or both. Suppose that $s \notin Mod(\kappa \triangle_{Max} (\lambda \vee \mu))$. So there is some $s'' \in Mod(\top)$ such that $\max\{d(s'', \kappa), d(s'', \lambda \vee \mu)\} \leq \max\{d(s, \kappa), d(s, \lambda \vee \mu)\}$. If $d(s'', \lambda \vee \mu)\} = d(s'', \lambda)$ then $\max\{d(s'', \kappa), d(s'', \lambda\} \leq \max\{d(s, \kappa), d(s, \lambda \vee \mu)\}$. But $d(s, \lambda \vee \mu) \leq d(s, \lambda)$ and so $\max\{d(s'', \kappa), d(s'', \lambda\} \leq \max\{d(s, \kappa), d(s, \lambda\}$. Otherwise $d(s'', \lambda \vee \mu) < d(s'', \lambda)$ and thus $d(s'', \lambda \vee \mu) = d(s'', \mu)$. But then $\max\{d(s'', \kappa), d(s'', \mu\} \leq \max\{d(s, \kappa), d(s, \lambda \vee \mu)\}$. But $d(s, \lambda \vee \mu) \leq d(s, \mu)$ and thus $\max\{d(s'', \kappa), d(s'', \mu\} \leq \max\{d(s, \kappa), d(s, \mu\}$. So it cannot be the case that for every $s' \in Mod(\top)$ it holds that $\max\{d(s, \kappa), d(s, \lambda)\} \leq \max\{d(s', \kappa), d(s', \lambda)\}$ or $\max\{d(s, \kappa), d(s, \mu)\} \leq \max\{d(s', \kappa), d(s', \mu)\}$ or both. Contradiction. But then $s \in Mod(\kappa \triangle_{Max} (\lambda \vee \mu))$. Since s was chosen arbitrarily, the result holds in the other direction as well. But then $\triangle_{Max}$ satisfies postulate (**LS** $\triangle$ **6**).

(**LS** $\triangle$ **7**) It will be shown that it not always the case that $Mod(\kappa \triangle_{Max} \lambda) \subseteq Mod(\kappa \vee \lambda)$. The proof is by counterexample. Suppose L is generated by $Atom = \{P(a), P'(a)\}$. Let $Mod(\kappa) = \{00\}$ and $Mod(\lambda) = \{11\}$. Abusing notation, $d(\kappa) = \{(11, 2), (10, 1), (01, 1), (00, 0)\}$ and $d(\lambda) = \{(11, 0), (10, 1), (01, 1), (00, 2)\}$. But then $d_{Max}(\{\kappa, \lambda\}) = \{(11, 2), (10, 1), (01, 1), (00, 2)\}$ and thus $Mod(\kappa \triangle_{Max} \lambda) = \{10, 01\}$. But $Mod(\kappa \vee \lambda) = \{00, 11\}$. So $Mod(\kappa \triangle_{Max} \lambda) \nsubseteq Mod(\kappa \vee \lambda)$. So postulate (**LS** $\triangle$ **7**) is not satisfied.

(**LS** $\triangle$ **8**) It will be shown that it not always the case that $Mod(\kappa) \cap Mod(\kappa \triangle_{Max} \lambda) \neq \varnothing$. Using the same example as for postulate (**LS** $\triangle$ **7**), it follows that $Mod(\kappa) \cap Mod(\kappa \triangle_{Max} \lambda) = \{00\} \cap \{10, 01\} = \varnothing$. So postulate (**LS** $\triangle$ **8**) is not satisfied. ■

The third family of merging operations, called $GMax$ operations, are a generalisation of Max operations. A $GMax$ operation orders the distances between a possible world and the individual knowledge bases in descending order and uses the lexicographical order $\leq_{lex}$ on sequences of $\mathbb{N}$ to induce a pre-order on S.

Definition 6.7 *Let Γ be a multiset of satisfiable knowledge bases of L and let $s, s' \in S$.*

Then the merging operation $\triangle_{GMax}$ *is defined by taking* $d_{GMax}(s,\Gamma) = \langle d(s,\kappa) \mid \kappa \in \Gamma\rangle$ *in descending order and* $s \leq_{\Gamma,GMax} s'$ *iff* $d_{GMax}(s,\Gamma) \leq_{lex} d_{GMax}(s',\Gamma)$ *so that* $Mod(\triangle_{GMax}(\Gamma)) = Min_{\leq_{\Gamma,GMax}}(\top)$.

The merging operation $\triangle_{GMax}$ is given by Konieczny and Pino-Pérez (1998) as an example of an arbitration operation. As with the other two merging operations, it is based on Dalal's distance function. However, $\triangle_{GMax}$ is not an arbitration operation in the sense of Liberatore and Schaerf.

Proposition 6.10 (Konieczny and Pino-Pérez, 1998) $\triangle_{GMax}$ *satisfies postulates* $(\mathbf{KP} \triangle \mathbf{1})$ *to* $(\mathbf{KP} \triangle \mathbf{6})$ *and satisfies postulate* $(\mathbf{KP} \triangle \mathbf{8})$ *iff* $card(Atom) > 1$.

Proposition 6.11 *Let* κ, λ, *and* μ *be satisfiable knowledge bases of* L. $\triangle_{GMax}$ *satisfies postulates* $(\mathbf{LS} \triangle \mathbf{1})$ *to* $(\mathbf{LS} \triangle \mathbf{6})$, *but fails to satisfy postulates* $(\mathbf{LS} \triangle \mathbf{7})$ *and* $(\mathbf{LS} \triangle \mathbf{8})$.

Proof. Postulates $(\mathbf{LS} \triangle \mathbf{1})$ to $(\mathbf{LS} \triangle \mathbf{5})$ are satisfied through the correspondence with postulates $(\mathbf{KP} \triangle \mathbf{0})$ to $(\mathbf{KP} \triangle \mathbf{3})$ and proposition 6.10.

$(\mathbf{LS} \triangle \mathbf{6})$ It must be shown that $Mod(\kappa \triangle_{GMax} (\lambda \vee \mu)) = Mod(\kappa \triangle_{GMax} \lambda)$ or $Mod(\kappa \triangle_{GMax} \mu)$ or $Mod(\kappa \triangle_{GMax} \lambda) \cup Mod(\kappa \triangle_{GMax} \mu)$. Choose any $s \in Mod(\kappa \triangle_{GMax} (\lambda \vee \mu))$. So there is no $s' \in Mod(\top)$ such that $s' \leq_{\{\kappa,\lambda\vee\mu\},GMax} s$ unless $s \leq_{\{\kappa,\lambda\vee\mu\},GMax} s'$. In other words, for every $s' \in Mod(\top)$ it holds that $d_{GMax}(s, \{\kappa, \lambda \vee \mu\}) \leq_{lex} d_{GMax}(s', \{\kappa, \lambda \vee \mu\})$, i.e. $\langle d(s,\alpha) \mid \alpha \in \{\kappa, \lambda \vee \mu\}\rangle \leq_{lex} \langle d(s',\alpha) \mid \alpha \in \{\kappa, \lambda \vee \mu\}\rangle$. Suppose that $s \notin Mod(\kappa \triangle_{GMax} \lambda)$ and $s \notin Mod(\kappa \triangle_{GMax} \mu)$ and therefore also $s \notin (Mod(\kappa \triangle_{GMax} \lambda) \cup Mod(\kappa \triangle_{GMax} \mu)$. Since $s \notin Mod(\kappa \triangle_{GMax} \lambda)$ it follows that there is some $s'' \in Mod(\top)$ such that $\langle d(s'',\alpha) \mid \alpha \in \{\kappa, \lambda\}\rangle <_{lex} \langle d(s,\alpha) \mid \alpha \in \{\kappa, \lambda\}\rangle$. Suppose that $\max\{d(s,\kappa), d(s,\lambda)\} = d(s,\kappa)$. So $d(s'',\kappa) < d(s,\kappa)$ and $d(s'',\lambda) < d(s,\kappa)$ and $d(s,\lambda) \leq d(s,\kappa)$. But $d(s'',\lambda \vee \mu) \leq d(s'',\lambda)$ and $d(s,\lambda \vee \mu) \leq d(s,\lambda)$. So $\langle d(s'',\alpha) \mid \alpha \in \{\kappa, \lambda \vee \mu\}\rangle$

$<_{lex}$ $\langle d(s,\kappa), d(s,\lambda\vee\mu)\rangle$. Contradiction. Now suppose that $\max\{d(s,\kappa), d(s,\lambda)\} = d(s,\lambda)$. So $d(s'',\kappa) \leq d(s,\lambda)$ and $d(s'',\lambda) \leq d(s,\lambda)$ and $d(s,\kappa) \leq d(s,\lambda)$. If $d(s,\lambda\vee\mu) = d(s,\lambda)$ then $\langle d(s'',\alpha) \mid \alpha \in \{\kappa, \lambda\vee\mu\}\rangle \leq_{lex} \langle d(s,\lambda\vee\mu), d(s,\kappa)\rangle$ resulting in a contradiction. Otherwise $d(s,\lambda\vee\mu) < d(s,\lambda)$ and thus $d(s,\lambda\vee\mu) = d(s,\mu)$. Since $s \notin Mod(\kappa\triangle_{GMax}\mu)$ it follows that there is some $s''' \in Mod(\top)$ such that $\langle d(s''',\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$ $\leq_{lex}$ $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$. Since $d(s''',\lambda\vee\mu) \leq d(s''',\mu)$ and $d(s,\lambda\vee\mu) = d(s,\mu)$ it follows that $\langle d(s''',\alpha) \mid \alpha \in \{\kappa,\lambda\vee\mu\}\rangle \leq_{lex} \langle d(s,\lambda\vee\mu), d(s,\kappa)\rangle$. Contradiction. But then it must be the case that $s \in Mod(\kappa \triangle_{GMax} \lambda)$ or $s \in Mod(\kappa \triangle_{GMax} \mu)$ or $s \in (Mod(\kappa \triangle_{GMax} \lambda) \cup Mod(\kappa \triangle_{GMax} \mu)$. Since s was chosen arbitrarily, the result holds in one direction.

Conversely, choose any $s \in Mod(\kappa \triangle_{GMax} \lambda)$ or $s \in Mod(\kappa \triangle_{GMax} \mu)$ or $s \in (Mod(\kappa \triangle_{GMax} \lambda) \cup Mod(\kappa \triangle_{GMax} \mu)$. So for every $s' \in Mod(\top)$ it holds that $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle \leq_{lex} \langle d(s',\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle$ or $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$ $\leq_{lex}$ $\langle d(s',\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$ or both. Suppose that $s \notin Mod(\kappa \triangle_{GMax} (\lambda\vee\mu))$. So there is some $s'' \in Mod(\top)$ such that $\langle d(s'',\alpha) \mid \alpha \in \{\kappa,\lambda\vee\mu\}\rangle \leq_{lex} \langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\vee\mu\}\rangle$. If $d(s'',\lambda\vee\mu)\} = d(s'',\lambda)$ then $\langle d(s'',\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle \leq_{lex} \langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\vee\mu\}\rangle$. But $d(s,\lambda\vee\mu) \leq d(s,\lambda)$ and thus $\langle d(s'',\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle \leq_{lex} \langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle$. Otherwise $d(s'',\lambda\vee\mu) < d(s'',\lambda)$ and thus $d(s'',\lambda\vee\mu) = d(s'',\mu)$. But then $\langle d(s'',\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle \leq_{lex}$ $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\vee\mu\}\rangle$. But $d(s,\lambda\vee\mu) \leq d(s,\mu)$ and thus $\langle d(s'',\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle \leq_{lex}$ $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$. So it cannot be the case that for every $s' \in Mod(\top)$ it holds that $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle \leq_{lex} \langle d(s',\alpha) \mid \alpha \in \{\kappa,\lambda\}\rangle$ or $\langle d(s,\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle \leq_{lex}$ $\langle d(s',\alpha) \mid \alpha \in \{\kappa,\mu\}\rangle$ or both. Contradiction. But then $s \in Mod(\kappa \triangle_{GMax} (\lambda\vee\mu))$. Since s was chosen arbitrarily, the result holds in the other direction as well. But then $\triangle_{GMax}$ satisfies postulate (**LS** $\triangle$ **6**).

(**LS** $\triangle$ **7**) It will be shown that it not always the case that $Mod(\kappa \triangle_{GMax} \lambda) \subseteq Mod(\kappa\vee\lambda)$. The proof is by counterexample, using the same example as for proposition

6.9. So $Mod(\kappa) = \{00\}$ and $Mod(\lambda) = \{11\}$ with $d(\kappa) = \{(11, 2), (10, 1), (01, 1), (00, 0)\}$ and $d(\lambda) = \{(11, 0), (10, 1), (01, 1), (00, 2)\}$. But then $d_{GMax}(\{\kappa, \lambda\}) = \{(11, \langle 2, 0\rangle), (10, \langle 1, 1\rangle), (01, \langle 1, 1\rangle), (00, \langle 2, 0\rangle)\}$ and thus $Mod(\kappa \triangle_{GMax} \lambda) = \{10, 01\}$. But $Mod(\kappa \vee \lambda) = \{00, 11\}$. So $Mod(\kappa \triangle_{GMax} \lambda) \not\subseteq Mod(\kappa \vee \lambda)$ and thus postulate (**LS** $\triangle$ **7**) is not satisfied.

(**LS** $\triangle$ **8**) It will be shown that it not always the case that $Mod(\kappa) \cap Mod(\kappa \triangle_{GMax} \lambda) \neq \varnothing$. Using the same example as for postulate (**LS** $\triangle$ **7**), it follows that $Mod(\kappa) \cap Mod(\kappa \triangle_{GMax} \lambda) = \{00\} \cap \{10, 01\} = \varnothing$. So postulate (**LS** $\triangle$ **8**) is not satisfied. ■

The three families of merging operations differ primarily in their choice of the function $\odot$, which, as mentioned earlier, determines an overall distance between a possible world and the set of knowledge bases to be merged. These functions are also referred to as *aggregation functions* (Konieczny, Lang, and Marquis, 2002, 2004) and abstractly defined as a function $\odot : \mathbb{N} \times \mathbb{N} \times \ldots \times \mathbb{N} \rightarrow \mathbb{N}$ that satisfies a number of constraints. The abstract distance function d and the abstract aggregation function $\odot$ provide two parameters for defining specific merging operations. By adding an additional aggregation function $\oplus$, DA^2 merging operators are obtained (Konieczny, Lang, and Marquis, 2004). The D is short for 'Distance-based' and the A^2 short for '2 Aggregation steps'. The second aggregation function $\oplus$ is only applicable when the set of knowledge bases to be merged are belief bases and syntactic methods for constructing merging operations are called for. The basic idea is that the arbitrary set of sentences that comprises a belief base represents pieces of information that can be aggregated. In the context of diagrammable systems, the aggregation function $\oplus$ is irrelevant.

The three families of merging operations that were considered all operate at the level of belief sets (or knowledge bases). In the next section, the focus turns to merging at the level of epistemic states.

6.3 Merging epistemic states

In the approach of Meyer (2001a) epistemic states are described as structures in the style of Spohn (1988) based on a finitely generated propositional language under a traditional truth-value semantics. The epistemic state of an agent is taken to be a function e from the set S of possible worlds to $\mathbb{N}$ while the knowledge base $\kappa(e) \in L$ associated with e is defined by $Mod(\kappa(e)) = \{s \in S \mid e(s) = 0\}$. From this definition it follows that a knowledge base $\kappa(e)$ may be unsatisfiable. For every epistemic state e, $min(e) = \min\{e(s) \mid s \in S\}$ and $max(e) = \max\{e(s) \mid s \in S\}$. An *epistemic list* $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ is a finite list of epistemic states of length m. A merging operation $\triangle$ on epistemic states is a function from the set of all non-empty epistemic lists to the set of all epistemic states.

Meyer (2001a) proposes the following properties for the merging of epistemic states:

($\mathbf{M} \triangle \mathbf{0}$) $\triangle([e])(s) = e(s) - min(e)$

($\mathbf{M} \triangle \mathbf{1}$) $\exists s$ s.t. $\triangle(\mathbf{E})(s) = 0$

($\mathbf{M} \triangle \mathbf{2}$) If $e_i(s) = e_j(s)$ $\forall e_i, e_j \in \mathbf{E}$, $e_i(s) \leq e_i(s')$ $\forall e_i \in \mathbf{E}$, and $e_j(s) < e_j(s')$ for some $e_j \in \mathbf{E}$ then $\triangle(\mathbf{E})(s) < \triangle(\mathbf{E})(s')$[1]

($\mathbf{M} \triangle \mathbf{3}$) If $e_i(s) \leq e_i(s')$ $\forall e_i \in \mathbf{E}$ then $\triangle(\mathbf{E})(s) \leq \triangle(\mathbf{E})(s')$

($\mathbf{M} \triangle \mathbf{4}$) If $\triangle(\mathbf{E})(s) \leq \triangle(\mathbf{E})(s')$ then $e_i(s) \leq e_i(s')$ for some $e_i \in \mathbf{E}$

Property ($\mathbf{M} \triangle \mathbf{0}$) ensures that trivial merging (with a singleton epistemic list) results in a satisfiable knowledge base while property ($\mathbf{M} \triangle \mathbf{1}$) ensures the same for non-trivial merging. Property ($\mathbf{M} \triangle \mathbf{2}$) says that whenever all of the epistemic states in $\mathbf{E}$ have the same level of plausibility for a specific state s then, in the epistemic state obtained from merging, s would be strictly more plausible than any state s' which is at most as

[1] This formulation is the simplifiied version used in Chopra, Ghose, and Meyer (2006).

plausible as s in every epistemic state in $\mathbf{E}$, but less plausible than s in at least one epistemic state in $\mathbf{E}$. Property ($\mathbf{M} \triangle \mathbf{3}$) states that if s is at least as plausible as s' in every epistemic state in $\mathbf{E}$, then it should remain so in the epistemic state obtained from merging. Property ($\mathbf{M} \triangle \mathbf{4}$) states that if s is at least as plausible as s' in the epistemic state obtained from merging the epistemic states in $\mathbf{E}$, then this should have been the case for at least one epistemic state in $\mathbf{E}$.

When the **KP** postulates are phrased in terms of a multiset of knowledge bases, then property ($\mathbf{M} \triangle \mathbf{1}$) is the counterpart of postulate ($\mathbf{KP} \triangle \mathbf{1}$) while property ($\mathbf{M} \triangle \mathbf{2}$) can be seen as a generalisation of postulate ($\mathbf{KP} \triangle \mathbf{2}$). To see this, recall that in the semantic representation of a knowledge base $\kappa \in L$ the models of κ are strictly below (and hence more plausible than) the nonmodels of κ. The counterpart for postulate ($\mathbf{KP} \triangle \mathbf{3}$) is called property (**Comm**) because it is a commitment to commutativity (as is postulate ($\mathbf{KP} \triangle \mathbf{3}$) when phrased in terms of multisets). It relies on a notion of equivalence between epistemic lists. Two epistemic lists $\mathbf{E}$ and $\mathbf{E}'$ are said to be *element-equivalent*, written as $\mathbf{E} \approx \mathbf{E}'$, iff for every element e of $\mathbf{E}$ there is a unique element e' (position-wise) of $\mathbf{E}'$ such that $e = e'$ and vice versa.

(**Comm**) If $\mathbf{E} \approx \mathbf{E}'$ then $\triangle(\mathbf{E}) = \triangle(\mathbf{E}')$

Postulate ($\mathbf{KP} \triangle \mathbf{7}$) for majority merging and the (original) postulate ($\mathbf{KP} \triangle \mathbf{8}'$) for arbitration of Konieczny and Pino-Pérez are also generalised for merging operations on epistemic states.

(**Maj**) $\exists m$ s.t. $\forall s, s' \in S$, $e(s) \leq e(s')$ if $\triangle(\mathbf{E} \sqcup e^m)(s) \leq \triangle(\mathbf{E} \sqcup e^m)(s')$

(**Arb**) $\forall m$ $\triangle(\mathbf{E} \sqcup [e])(s) \leq \triangle(\mathbf{E} \sqcup [e])(s')$ iff $\triangle(\mathbf{E} \sqcup e^m)(s) \leq \triangle(\mathbf{E} \sqcup e^m)(s')$

The effect of property (**Maj**) for majority merging is to produce a refined version of epistemic state e when the epistemic state e has been included enough times in an epistemic list. This is similar to postulate (**KP** $\triangle$ **7**) for knowledge base majority merging where the effect is to produce a refined version of knowledge base κ (a subset of $Mod(\kappa)$) when κ has been included enough times in the multiset. Property (**Arb**) for arbitration says (similarly to postulate (**KP** $\triangle$ **8**$'$) for knowledge base arbitration) that the result of arbitration is independent of the frequency with which an epistemic state (a knowledge base) appears in an epistemic list (a multiset).

The feasibility of these properties is supported by the construction of several merging operations on epistemic states, of which generalisations of the three families of merging operations on knowledge bases are examples. As pointed out by Meyer (2001a), the distance $d(s, \kappa)$ between a possible world s and a knowledge base κ can be used to induce an epistemic state e by taking $e(s) = d(s, \kappa)$ for every $s \in S$. It then follows that $e(s) = 0$ iff $s \in Mod(\kappa)$ so that $\kappa(e) \equiv \kappa$, i.e. so that κ is the knowledge base associated with epistemic state e. However, an epistemic state need not be induced in this manner.

The first collection of merging operations on epistemic states includes generalisations of the $\triangle_\Sigma$ and $\triangle_{Max}$ operations for merging knowledge bases and introduces a new merging operation on epistemic states, denoted by $\triangle_{Min}$, which is inspired by an arbitration operation proposed by Liberatore and Schaerf (1998).

Definition 6.8 *Let* $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ *be an epistemic list and let* $s \in S$*. Then the merging operations* $\triangle_\Sigma$*,* $\triangle_{Max}$*, and* $\triangle_{Min}$ *are defined as follows:*

- $\triangle_\Sigma(\mathbf{E})(s) = Sum_\mathbf{E}(s) - \min\{Sum_\mathbf{E}(s) \mid s \in S\}$
  where $Sum_\mathbf{E}(s) = \sum e_i(s) \mid e_i \in \mathbf{E}$

- $\triangle_{Max}(\mathbf{E})(s) = Max_\mathbf{E}(s) - \min\{Max_\mathbf{E}(s) \mid s \in S\}$
  where $Max_\mathbf{E}(s) = \max\{e_i(s) \mid e_i \in \mathbf{E}\}$

- $\triangle_{Min}(\mathbf{E})(s) = Min_\mathbf{E}(s) - \min\{Min_\mathbf{E}(s) \mid s \in S\}$

$$\textit{where } Min_{\mathbf{E}}(s) = \begin{cases} e_1(s) & \textit{if } m = 1 \\ 2\min\{e_i(s) \mid e_i \in \mathbf{E}\} & \textit{if } e_i(s) = e_j(s) \forall e_i, e_j \in \mathbf{E} \\ 2\min\{e_i(s) \mid e_i \in \mathbf{E}\} + 1 & \textit{otherwise} \end{cases}$$

In all of these merging operations the uniform subtraction is a form of normalisation to ensure that the knowledge base associated with the resulting epistemic state is satisfiable, as required by property $(\mathbf{M} \triangle \mathbf{1})$.

The next collection of merging operations on epistemic states is a refinement of the first collection and includes a generalisation of the $\triangle_{GMax}$ operation for merging knowledge bases. As with the family of $GMax$ operations, some of these merging operations use a lexicographical order $\leq_{lex}$ on (finite) sequences of $\mathbb{N}$; others a total preorder. Given any set Seq of finite sequences of $\mathbb{N}$ and any total preorder $\leq$ on Seq, Meyer defines a function $\Omega_{Seq,\leq} : Seq \to \{0, 1, \ldots, card(Seq) - 1\}$ which assigns consecutive natural numbers to the elements of Seq in the order imposed by $\leq$, starting by assigning 0 to the elements lowest down in $\leq$. Associated with every epistemic list $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ is the set $Seq_{\mathbf{E}}$ of all sequences of length m of $\mathbb{N}$, ranging from 0 to $\max\{max(e_i) \mid e_i \in \mathbf{E}\}$, the subset $Seq_{\mathbf{EA}}$ of all sequences of $Seq_{\mathbf{E}}$ in ascending order, and the subset $Seq_{\mathbf{ED}}$ of all sequences of $Seq_{\mathbf{E}}$ in descending order.

For every $s \in S$, $q_{\mathbf{E}}(s)$ denotes the sequence $\langle e_1(s), e_2(s), \ldots, e_m(s) \rangle$ while $q_{\mathbf{EA}}(s)$ and $q_{\mathbf{ED}}(s)$ respectively denotes $q_{\mathbf{E}}(s)$ in ascending and descending order. For $q \in Seq_{\mathbf{E}}$, $sum_{\mathbf{E}}(q) = \sum_{i=1}^{m} q_i$ and $d_{\mathbf{E}}(q) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} |q_i - q_j|$ where q_i denotes the i-th element of q. A total preorder $\leq_{\mathbf{E}}$ on $Seq_{\mathbf{E}}$ is defined by taking $q \leq_{\mathbf{E}} q'$ iff $sum_{\mathbf{E}}(q) < sum_{\mathbf{E}}(q')$ or $(sum_{\mathbf{E}}(q) = sum_{\mathbf{E}}(q')$ and $d_{\mathbf{E}}(q) \leq d_{\mathbf{E}}(q'))$.

Meyer (2001a) defines as refinements of $\triangle_{\Sigma}$, $\triangle_{Max}$, and $\triangle_{Min}$ respectively the merging operations $\triangle_{R\Sigma}$, $\triangle_{RMax}$, and $\triangle_{RMin}$ on epistemic states (with $\triangle_{RMax}$ the generalisation of $\triangle_{GMax}$.)

Definition 6.9 *Let* $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ *be an epistemic list and let* $s, s' \in S$. *Then the* *merging operation* $\triangle_{R\Sigma}$, $\triangle_{RMax}$, *and* $\triangle_{RMin}$ *are defined as follows:*

- $\triangle_{R\Sigma}(\mathbf{E})(s) = RSum_{\mathbf{E}}(s) - \min\{RSum_{\mathbf{E}}(s) \mid s \in S\}$
  where $RSum_{\mathbf{E}}(s) = \Omega_{Seq_{\mathbf{E}}, \leq_{\mathbf{E}}}(q_{\mathbf{E}}(s))$

- $\triangle_{RMax}(\mathbf{E})(s) = RMax(s) - \min\{RMax_{\mathbf{E}}(s) \mid s \in S\}$
  where $RMax(s) = \Omega_{Seq_{\mathbf{ED}}, \leq_{lex}}(q_{\mathbf{ED}}(s))$

- $\triangle_{RMin}(\mathbf{E})(s) = RMin(s) - \min\{RMin_{\mathbf{E}}(s) \mid s \in S\}$
  where $RMin_{\mathbf{E}}(s) = \Omega_{Seq_{\mathbf{EA}}, \leq_{lex}}(q_{\mathbf{EA}}(s))$

Both collections of merging operations satisfy properties ($\mathbf{M} \triangle \mathbf{0}$) to ($\mathbf{M} \triangle \mathbf{4}$) and (**Comm**) supporting the view that these properties may be regarded as basic properties of merging operations on epistemic states.

Proposition 6.12 (Meyer, 2001) *$\triangle_{\Sigma}$, $\triangle_{Max}$, $\triangle_{Min}$,$\triangle_{R\Sigma}$, $\triangle_{RMax}$, and $\triangle_{RMin}$ all satisfy properties* ($\mathbf{M} \triangle \mathbf{0}$) *to* ($\mathbf{M} \triangle \mathbf{4}$) *and* (**Comm**).

Proposition 6.13 (Meyer, 2001) *$\triangle_{\Sigma}$ and $\triangle_{R\Sigma}$ both satisfy property* (**Maj**) *while $\triangle_{Min}$ and $\triangle_{Max}$ both satisfy property* (**Arb**).

An interesting result is that none of the merging operations of Meyer satisfies postulate ($\mathbf{KP} \triangle \mathbf{4}$) while all but $\triangle_{R\Sigma}$ satisfy postulate ($\mathbf{KP} \triangle \mathbf{5}$) (when the **KP** postulates are rephrased for multisets). Postulate ($\mathbf{KP} \triangle \mathbf{6}$) is only satisfied by $\triangle_{\Sigma}$, $\triangle_{RMax}$, and $\triangle_{RMin}$. Although some of the merging operations of Meyer are based on the constructions of Konieczny and Pino-Pérez, the notion of distance between possible worlds differs between these approaches. Proposition 6.13 supports the use of property (**Maj**) as a suitable postulate for the subclass of majority merging operations on epistemic states and is compatible with results from knowledge base merging. For the subclass of arbitration operations on epistemic states, the use of property (**Arb**) conflicts, to some extent, with the results from knowledge base merging (Meyer, 2001a).

The last collection of merging operations on epistemic states comprises non-commutative merging operations, of which the operation $\triangle_{Lex}$ is presented as an example. This operation is a lexicographic refinement of the epistemic states in an epistemic list $\mathbf{E}$ under the assumption that the epistemic states in $\mathbf{E}$ are (strictly) ranked according to reliability.

Definition 6.10 *Let* $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ *be a ranked epistemic list and let* $s \in S$*. Then the merging operation* $\triangle_{Lex}$ *is defined as* $\triangle_{Lex}(\mathbf{E})(s) = Lex_{\mathbf{E}}(s) - \min\{Lex_{\mathbf{E}}(s) \mid s \in S\}$ *where* $Lex_{\mathbf{E}}(s) = \Omega_{Seq_{\mathbf{E}}, \leq_{\mathbf{lex}}}(q_{\mathbf{E}}(s))$*.*

Proposition 6.14 (Meyer, 2001) $\triangle_{Lex}$ *satisfies properties* $(\mathbf{M} \triangle \mathbf{0})$ *to* $(\mathbf{M} \triangle \mathbf{4})$ *but not property* (**Comm**)*.*

Maynard-Zhang and Lehmann (2003) make explicit the ranking on epistemic states when they define a total preorder $\sqsupseteq$ over a finite set of sources by taking $a \sqsupseteq a'$ iff $rank(a) \geq rank(a')$ where *rank* is function from a finite set $\mathbf{A}$ of sources to any totally ordered finite set *Rank* of ranks[2]. The intuition is that the epistemic state of an agent can be constructed when the agent is informed by a set of sources (agents) of varying degrees of reliability by *aggregating* (or merging) the epistemic states associated with each source. The epistemic state of an agent, called a *generalised belief state*, is taken to be a modular, transitive relation $\prec$ on S. The relation $\prec$ is one of strict likelihood with $s \prec s'$ interpreted as saying that 'there is reason to consider s as strictly more likely than s''. Associated with every relation $\prec$ is an *equal likelihood relation* $\sim$ representing 'agnosticism', defined such that $s \sim s'$ iff $s \not\prec s'$ and $s' \not\prec s$, and a *conflict relation* $\bowtie$, defined such that $s \bowtie s'$ iff $s \prec s'$ and $s' \prec s$. The epistemic state associated with a source a is denoted by $\prec_a$.

[2]The relations $\sqsupseteq$ and $\geq$ should be interpreted as saying that "greater" is "better".

Aggregating (the epistemic states associated with) strictly ranked sources can be achieved by a lexicographic refinement operation where lower ranked sources refine the epistemic states of higher ranked sources.

Definition 6.11 *Let* $\mathbf{A}$ *be a finite set of sources and let* $\sqsupseteq$ *be a total partial order on* $\mathbf{A}$*. Then the aggregation operation* ARG_{Rf} *is defined as* $ARG_{Rf}(\mathbf{A}) = \{(s, s') \mid (\exists a \in \mathbf{A})\ s \prec_a s' \wedge (\forall a' \in \mathbf{A},\ a' \sqsupset a \text{ implies } s \sim_{a'} s')\}$*.*

For the special case in which all of the sources are equally ranked, a different aggregation operation is proposed.

Definition 6.12 *Let* $\mathbf{A}$ *be a finite set of sources and let* $\sqsupseteq$ *be a total preorder on* $\mathbf{A}$ *that is fully connected. Then the aggregation operation* ARG_{Un} *is defined as* $ARG_{Un}(\mathbf{A}) = Un(\mathbf{A})^+$ *where* $Un(\mathbf{A}) = \bigcup_{a \in \mathbf{A}} \prec_a$*.*

A more general form of aggregation, in which several ranks are possible and more than one source may have the same rank, is defined by first applying the lexicographic refinement operation and then applying closure.

Definition 6.13 *Let* $\mathbf{A}$ *be a finite set of sources and let* $\sqsupseteq$ *be a total preorder on* $\mathbf{A}$*. Then the aggregation operation* ARG *is defined as* $ARG(\mathbf{A}) = ARG_{Rf}(\mathbf{A})^+$*.*

From the generalised aggregation operation ARG, the aggregation operations ARG_{Rf} and ARG_{Un} can be obtained as special cases:

- if $\sqsupseteq$ is a total partial order, then $ARG(\mathbf{A}) = ARG_{Rf}(\mathbf{A})$ and
- if $\sqsupseteq$ is fully connected, then $ARG(\mathbf{A}) = ARG_{Un}(\mathbf{A})$.

The approach of Maynard-Zhang and Lehmann was inspired by earlier work of Maynard-Reid II and Shoham (2001) in which the epistemic state of an agent is taken to be a total preorder on S and the reliability of sources is assumed to be strictly ranked. The decision not to use total preorders was motivated by results from social choice theory. The classical approach in social choice theory is to use total preorders to represent (collective) preferences but Arrow's (1951, 1963) celebrated Impossibility Theorem showed that no aggregation operation based on total preorders exists that satisfies a particular small set of conditions deemed desirable. Maynard-Zhang and Lehmann show that the aggregation operation ARG, which is based on the lexicographic refinement of modular transitive relations representing the epistemic states of agents, satisfies suitably modified versions of Arrow's conditions. A similar result is obtained by Andréka, Ryan and Schobbens (2002) who show that the lexicographic rule is the only way of combining binary relations, called preference relations, which satisfies natural conditions that are similar to those of Arrow. As stated by Maynard-Zhang and Lehmann (2003):

> "The social choice community has dealt extensively with the problem of representing collective preferences (cf. Sen, 1986). However, the problem is formally equivalent to that of representing collective beliefs, so the results are applicable."

Stated differently, the problem of merging epistemic states may be viewed as technically similar to the problem of aggregating preferences in social choice theory. The connection between belief merging (in the broader sense) and social choice theory is generally seen as a fruitful area for future research (Delgrande et al., 2005; Chopra, Ghose, and Meyer, 2006; Pigozzi, 2006). It seems reasonable then to take note of the results in social choice theory before proposing the notion of templated merging.

6.4 Social choice theory

Social choice theory is concerned about social (or collective) decision-making and the logical foundations of welfare economics (Arrow, Sen, and Suzumura, 2002). The focus of this section will be on social decision-making and, in particular, the conditions imposed by Arrow. The origins of social choice theory can be traced back to the works of Borda (1781) and Condorcet (1785) on voting in elections. Condorcet discovered that pairwise majority voting may lead to indeterminacy in the social choice, a paradox generally referred to as the *Condorcet paradox*. Borda proposed a method of *rank-order voting*, which supported his view that the entire ordering of individual voters over alternative candidates is needed for social decision. An extensive history of the theory of social choice, starting with the works of Borda and Condorcet, is provided by Black (1958) who also uncovered the work of Dodgson (Lewis Carroll) from unpublished manuscripts.

One of the most significant contributions to the theory of social choice is due to Arrow (1951, 1963). In contrast to earlier work, Arrow did not focus on specific voting schemes but on the process or rule by which individual preferences determine social choice. This process or rule for aggregating individual preferences into a social preference is captured by the notion of a *social welfare function*. Arrow's approach is axiomatic in nature in the sense that a set of conditions is imposed for social welfare functions to be deemed reasonable. The startling result from Arrow's General Possibility Theorem (or Impossibility Theorem) is that there exists no social welfare function satisfying this set of conditions.

In Arrow's framework, the society of individuals is assumed to be finite and fixed. The preference of an individual i over a finite set A of alternative social states is taken to be a total preorder R_i[3]. If x and y are alternatives in A, then $(x, y) \in R$ states that

[3] Arrow used the term "ordering" instead of "total preorder".

'x is preferred or indifferent to y'. Associated with every total preorder R is a (strict) *preference relation* P defined such that $(x, y) \in P$ iff $(x, y) \in R$ but $(y, x) \notin R$ and an *indifference relation* I defined such that $(x, y) \in I$ iff $(x, y) \in R$ and $(y, x) \in R$. The concept of choice is defined in terms of R as the *choice set* $C_R(A) = \{x \in A \mid (x, y) \in R$ for every $y \in A\}$. For every non-empty subset $A' \subseteq A$ the choice set $C_R(A')$ is non-empty (Sen, 1970). It is interesting to note that the preference relation P is nothing other than the strict modular partial order Q corresponding to R (as shown in proposition 3.8) and that the choice set $C_R(A)$ is the set of minimal elements in A with respect to R, i.e. $C_R(A) = Min_R(A)$.

Definition 6.14 *A* ***social welfare function*** *is a function swf that maps a m-tuple of total preorders $R_1, R_2, \ldots, R_m$ on A (representing individual preference orderings) to a corresponding total preorder R_{SC} on A (representing the social preference ordering).*

The definition of the social welfare function reflects additional characteristics of Arrow's framework: firstly, social choice depends only on individual preferences; secondly, the outcome of the process for determining social choice is a social preference ordering; and thirdly, the social preference ordering is a total preorder.

As is to be expected, frameworks outside this classic framework abound (Campbell and Kelly, 2002). One such framework, characterised by the assumption that society does not require a social preference ordering in order to have a social choice, deserves special mention by virtue of its link with nonmonotonic reasoning. It relies on the notion of a *choice function* to determine the choice set, without assuming the pre-existence of some order relation, and a set of rationality conditions to impose *coherence constraints* on choices. Both Arrow (1959) and Sen (1969, 1970) contributed to the development of this framework, sometimes referred to as the theory of rational choice, of which a survey is provided by Moulin (1985). The link between this framework of social choice theory and nonmonotonic logic is investigated, amongst others, by Lehmann (2001) and Rott (1998, 2001)[4].

Arrow imposed four conditions[5] on social welfare functions (Arrow, 1963 ch.VIII).

- Condition **U** (*Unrestricted domain*) All logically possible total preorders on the set of alternative social states are admissible.

The rationale behind condition **U** is that each individual in society should be free to form and express any preference of social states and that the social welfare function should be robust enough to aggregate the individual preference orderings into a social preference ordering.

- Condition **P** (*Pareto principle*) If every individual strictly prefers some social state to another, then so should society.

The Pareto principle ensures that the social welfare function faithfully reflects unanimous (strict) preference expressed by all individuals over a pair of social states. Formally, it is defined as follows: If $(x, y) \in P_i$ for every individual i, then $(x, y) \in P_{SC}$.

- Condition **IIA** (*Independence of irrelevant alternatives*) The social choice between two alternatives should be independent of other alternatives.

Condition **IIA** ensures that if the choice is between alternatives x and y, and some change occurs that leaves the individual choices between x and y the same (for all individuals), then, irrespective of other changes amongst alternatives, the social choice

[4] An earlier study linking nonmonotonic and default reasoning with social choice theory in the classic framework may be found in Doyle and Wellman (1991).

[5] We follow Sen (1970) in labelling the conditions except for condition **IIA**, which Sen labelled as condition **I**. (Condition **U** corresponds to Arrow's condition **1′**, condition **IIA** to his condition **3**, and condition **D** to his condition **5**.)

between between x and y will remain the same. Formally condition **IIA** can be defined as follows: Let $R_1, R_2, \ldots, R_m$ and $R'_1, R'_2, \ldots, R'_m$ be two m-tuples of total preorders on A. For every $x, y \in A$, it holds that if for every individual i, $(x, y) \in R_i$ iff $(x, y) \in R'_i$, then $(x, y) \in R_{SC}$ iff $(x, y) \in R'_{SC}$.[6]

- Condition **D** (*Nondictatorship*) The social welfare function should not be dictatorial.

A social welfare function is said to be *dictatorial* if there exists an individual i such that for every m-tuple of total preorders $R_1, R_2, \ldots, R_m$ on A, $(x, y) \in P_i$ implies $(x, y) \in P_{SC}$ for every $x, y \in A$.[7] Condition **D** ensures that social choice is not based on the preference of only one individual in society.

From Arrow's *Impossibility Theorem* it follows that there exists no social welfare function that can simultaneously satisfies all four conditions **U**, **P**, **IIA**, and **D**.

Theorem 6.1 (Arrow, 1963 ch.VIII) *Conditions* **U***,* **P***,* **IIA***, and* **D** *are inconsistent.*

The method of majority voting (or decision-making) in which society prefers alternative x to y iff the number of individuals in society who prefer x to y is at least as great as the number of individuals who prefer y to x has been shown to satisfy conditions **P**, **IIA**, and **D**. However, it fails to satisfy condition **U** as shown by the Condorcet paradox, which can be illustrated by considering three voters 1, 2, and 3 and three candidates x, y, and z and assuming that voter 1 prefers x to y and y to z, voter 2 prefers y to z and z to x, and voter 3 prefers z to x and x to y. By majority voting, each candidate defeats the

[6] In the original formulation (Arrows,1963), condition **IIA** is defined in terms of the choice set associated with a social preference.

[7] This definition of condition **D** makes explicit the inclusion of every element in the domain of the social welfare function as suggested by Sen(1970).

others by two votes so that society prefers x to y and y to z and z to x yielding a social preference ordering cycle. This leads to inconsistencies because transitivity requires that society prefers x to z.

The rank-order method of voting assigns a specific score to each candidate for being ranked first in any voter's preference, a smaller score for being ranked second in anyone's preference, and so on. The individual scores received by each candidate over all voters are then added up, sometimes referred to as the *Borda count*, and the winner is the candidate with the highest score. The rank-order method of voting satisfies conditions **U**, **P**, and **D** but fails to satisfy condition **IIA**. To use an illustration from Sen (1970, pp.39), consider three voters 1, 2, and 3 and three candidates x, y, z and assume that the scores for first, second, and third ranked choices are 3, 2, and 1 respectively. Suppose that voter 1 ranks candidates in the order x, y, z while voters 2 and 3 rank them in the order z, x, y. With a highest score of 7, candidates x and z are tied for first place. Now consider another scenario in which voter 1 changes his mind about (the now irrelevant) candidate y so that his ranking becomes x, z, y. Applying the method to this scenario yields a highest score of 8 for candidate z making him the winner. Everyone's ordering of x and z are still the same, but the social choice between x and z is not the same, hence violating condition **IIA**. The central concepts and results underlying the Borda method (and to a lesser extent the Condorcet method) are reviewed by Pattanaik (2002) in the context of positional rules for collective decision-making.

Much of the work that has been done in the field of social choice theory since the publication of Arrow's Impossibility Theorem either attempts to show the robustness of the theorem or attempts to find escape routes around the theorem by relaxing the conditions or by modifying the framework (or both). A classic example of relaxing a condition is the *single-peaked preference* approach of Black (1948) which relaxes condition **U** so that the method of majority decision-making (voting) becomes a social welfare function.

The basic idea is that if individuals classify or categorise alternatives in terms of some dimension, so that there are patterns of preferences, and vote for the alternative closest to their own position, then the individual's preference pattern is single-peaked. In the context of political voting, one such classification or categorisation may be extreme left, moderate left, moderate right, and extreme right. The assumption of single-peakedness requires that every triple be arranged in a manner characteristic of partial agreement, in other words, everyone agrees that a particular alternative in the triple is not worst (or not best, or not medium). This restriction is called *value restriction* and if it holds for every triple, then the method of majority voting is a social welfare function, provided the number of voters is odd. Sen and Pattanaik (1969) generalised the assumption of single-peakedness to an arbitrary number of voters by imposing a restriction called *extremal restriction*, which requires that if someone prefers x to y and y to z, then z is uniquely best in an individual's preference ordering iff x is uniquely worse in its preference ordering. Extremal restriction permits a wide variety of relations (for example, 'partly similar' or 'sharply opposite') and if it holds for every triple of alternatives, then the method of majority voting is a social welfare function. For a discussion of other types of domain restrictions, the reader is referred to Gaertner (2002).

The approach of Maynard-Zhang and Lehmann discussed earlier can be seen as an example where both the conditions are relaxed and the framework is modified. By adopting modular transitive relations as the notion of preference instead of total preorders, the Arrow framework is modified significantly. Recall that each modular transitive relation has an associated equal likelihood relation and an associated conflict relation. In the presence of conflicts, condition **IIA** is weakened to say that the social preference between two alternatives should be independent of other alternatives unless these alternatives put them in conflict. In the approach of Maynard-Zhang and Lehmann, information sources (individuals) are ranked according to credibility (or reliability) with the most credible information sources viewed as 'dictators' in the sense of Arrow. Condition **D** is therefore restricted to information sources of equal rank only. The relaxation of conditions **IIA** and **D** raise the question to what extent the Arrow conditions are applicable in a given context.

In the context of diagrammable systems, there would be a large degree of compatibility or similarity of preferences among agents. For example, consider two agents 1 and 2, each with a two-level ordering such that the states in agent's 1 upper level are in agent's 2 lower level and vice versa. These agents do not share a single untautological belief and therefore their beliefs cannot be expected to be sensibly reconcilable. The condition of unrestricted domain is therefore inappropriate in the context of diagrammable systems.

The setting for decision-making by control room agents is a complex and dynamic environment in which decisions are made, typically by a group of agents, on the basis of experience. Decision-making is these environments is characterised by high uncertainty, incomplete information, time pressures, organisational influences, and high stakes because of the potential consequences for the decision-makers[8]. In decision-making events such as these, interaction and cooperation between agents are essential and often lead to changes in the preferences of individual agents. The nature of decision-making by control room agents rules out the condition of nondictatorship as simply inappropriate. The Pareto principle, on the other hand, is a natural condition for decision-making by control room agents. If every agent participating in a decision-making event agrees that some alternative is preferred to another, then the Pareto principle will ensure that they adopt the most preferred alternative.

Experiments conducted by Tversky, Kahneman, and others showing that human behaviour rarely adheres to the condition of independence of irrelevant alternatives are

[8]The term *NDM* or *Naturalistic Decision Making* (Zsambok and Klein, 1997) refers to the study of decision-making in environments such as these.

applicable for decision-making by control room agents. These experiments demonstrate, for example, that the way in which alternatives are framed may affect decision makers' choices (Tverksy and Kahneman, 1981, 1986); that decision makers look for reasons to prefer one alternative to another (Tversky and Shafir, 1992); that in choice under uncertainty, decision makers value losses more heavily than corresponding gains (Kahneman and Tversky, 1979); that systematic errors or biases occur when humans make decisions (Kahneman, Slovic, and Tversky, 1982); and that humans don't choose objects per se, they tend to choose collections of features (or aspects) of which the objects are composed (Tversky, 1972). More recent material is covered in Kahneman and Tversky (2000). The condition of independence of irrelevant alternatives has also received some criticism on technical grounds (Hansson, 1973).

To summarise, the large degree of homogeneity among control room agents and the complex and dynamic setting for decision-making by control room agents mean that the conditions of unrestricted domain, nondictatorship, and independence of irrelevant alternatives are not considered to be applicable in the context of diagrammable systems. The Pareto principle, however, is applicable and will play a key role in templated merging.

6.5 Templated merging

In a recent survey on logic-based approaches to information fusion (Grégoire and Konieczny, 2006) the notion of belief merging is broadened to include other types of information such as knowledge, goals, regulations, and observations. The logic-based approaches in the context of *possibilistic logic* (Dubois, Lang, and Prade, 1994) are gaining in popularity from the earlier application to belief change (Dubois and Prade, 1991, 1992) to more recent applications to information fusion (Dubois and Prade, 2004; Benferhat, Dubois, Kaci, and Prade, 2002, 2006). Grounded in fuzzy methods (Zadeh, 1978), these possibilistic approaches may be seen as numerical and hence not applicable in the context of our qualitative treatment of control room agents.

It is necessary to distinguish between information that is symbolic, which can be communicated by agents in a linear fashion using sentences of the knowledge representation language, and information that is nonsymbolic, which can be 'communicated' by an agent's sensors as 'sub-agents'. Symbolic information communicated from another agent can be merged with an agent's own information requiring only a binary operation. But sensory input communicated by an agent's sub-agents typically happens in parallel and requires an m-ary operation to be merged. Note that both symbolic and nonsymbolic information can be represented, to the extent that they bear upon the exclusion of states of the system, by t-orderings in normal form, as discussed in section 3.4.

Templated merging refers to the merging of possibly conflicting information represented by regular t-orderings. An *information tuple* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ is an m-tuple of regular t-orderings. An information tuple $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ is said to be contradictory iff $\bigcup_{t_i \in \mathbf{T}} top(t_i) = S$ and satisfiable iff $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Two information tuples $\mathbf{T}$ and $\mathbf{T}'$ are *element-equivalent*, written as $\mathbf{T} \approx \mathbf{T}'$, iff for every t-ordering t of $\mathbf{T}$ there is a unique t-ordering t' (position-wise) of $\mathbf{T}'$ such that $t = t'$ and vice versa.

Formally, a templated merging operation $\triangle$ is an m-ary operation on regular t-orderings. Adopted as properties for templated merging are the uncontroversial postulates $(\mathbf{KP} \triangle \mathbf{1})$ to $(\mathbf{KP} \triangle \mathbf{3})$ of Konieczny and Pino-Pérez, the property of commutativity, and the Pareto principle from social choice theory. Suitably modified for regular t-orderings, the properties proposed for templated merging are the following:

$(\mathbf{TM} \triangle \mathbf{1})$ $bottom(\triangle(\mathbf{T})) \neq \varnothing$

$(\mathbf{TM} \triangle \mathbf{2})$ If $\mathbf{T}$ is satisfiable, then $bottom(\triangle(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$

$(\mathbf{TM} \triangle \mathbf{3})$ If $\mathbf{T} \approx \mathbf{T}'$, then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$

($\mathbf{TM} \bigtriangleup \mathbf{4}$) If $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$, then $\bigtriangleup(\mathbf{T})(s) \leq \bigtriangleup(\mathbf{T})(s')$

($\mathbf{TM} \bigtriangleup \mathbf{5}$) If $\bigtriangleup(\mathbf{T})(s) \leq \bigtriangleup(\mathbf{T})(s')$, then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$

Properties ($\mathbf{TM} \bigtriangleup \mathbf{1}$) to ($\mathbf{TM} \bigtriangleup \mathbf{3}$) are the respective counterparts to postulates ($\mathbf{KP} \bigtriangleup \mathbf{1}$) to ($\mathbf{KP} \bigtriangleup \mathbf{3}$) when phrased in terms of multisets. A separate property for commutativity is not required because property ($\mathbf{TM} \bigtriangleup \mathbf{3}$) implies commutativity. Property ($\mathbf{TM} \bigtriangleup \mathbf{4}$) is a formal expression of the Pareto principle while property ($\mathbf{TM} \bigtriangleup \mathbf{5}$) can be seen as guaranteeing that the merging operation doesn't introduce fictitious preferences that no individual agent entertains. Properties ($\mathbf{TM} \bigtriangleup \mathbf{4}$) and ($\mathbf{TM} \bigtriangleup \mathbf{5}$) correspond directly to Meyer's properties ($\mathbf{M} \bigtriangleup \mathbf{3}$) and ($\mathbf{M} \bigtriangleup \mathbf{4}$) respectively.

The feasibility of these properties will be illustrated by the construction of several templated merging operations on regular t-orderings. The proposed principle of Qualitativeness, which applies generally to control room agents, will be used to evaluate the concrete templated merging operations. The method for constructing templated merging operations relies on the notion of an *indexing function* and an associated (well-ordered) *index set*. For the family of *basic* templated merging operations, an indexing function $f_\odot : S \to \{0, 1, \ldots, k\}$ (with respect to information tuple $\mathbf{T}$) is defined abstractly as $f_\odot(s) = \odot\{t_i(s) \mid t_i \in \mathbf{T}\}$. Different indexing functions may have different codomains depending on the value of k which represents the maximum index. Every indexing function $f_\odot$ induces an index set $A_\odot$, which is a subset of $\{0, 1, \ldots, k\}$, defined as $A_\odot = ran(f_\odot)$. Since $A_\odot$ is finite and $\leq$ is a linear order on $A_\odot$ it follows that $\leq$ is a well-ordering on $A_\odot$. A basic templated merging operation $\bigtriangleup_\odot$, based on $f_\odot$, can now be defined abstractly.

Definition 6.15 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Let* $f_\odot : S \to \{0, 1, \ldots, k\}$ *be an indexing function (with respect to* $\mathbf{T}$*) and* $A_\odot$ *its associated index set, well-ordered by the linear order* $\leq$*. A* ***basic templated merging operation*** $\bigtriangleup_\odot$ *is defined as*

- $\triangle_{\odot}(\mathbf{T})(s) = \begin{cases} card(seg(f_{\odot}(s))) & \text{if } f_{\odot}(s) < k \text{ or } card(A_{\odot}) = 1 \\ n & \text{otherwise} \end{cases}$
  where $seg(f_{\odot}(s))$ is the initial segment of $f_{\odot}(s) \in A_{\odot}$.

Proposition 6.15 *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Every t-ordering produced by a basic templated merging operation $\triangle_{\odot}$ is regular.*

Proof. See proof in appendix D, section D.2. ∎

Basic templated merging operations differ only in their choice of indexing function. It is not hard to see that for every $s \in S$ the construction of $\triangle_{\odot}(\mathbf{T})(s)$ depends only on the the linear order $\leq$ on the index set $A_{\odot} = ran(f_{\odot})$. Provided the indexing function $f_{\odot}$ does not make use of arithmetic operations on template B, any basic merging operation $\triangle_{\odot}$ will satisfy the principle of Qualitativeness.

Three basic templated merging operations are proposed, namely, *minimisation*, *maximisation*, and *sum*, denoted respectively by $\triangle_{Min}$, $\triangle_{Max}$, and $\triangle_{\Sigma}$. They are based respectively on the indexing functions $f_{Min} : S \to \{0, 1, \ldots, n\}$, $f_{Max} : S \to \{0, 1, \ldots, n\}$, and $f_{\Sigma} : S \to \{0, 1, \ldots, m \times n\}$.

Definition 6.16 *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. The **indexing functions** $f_{Min} : S \to \{0, 1, \ldots, n\}$, $f_{Max} : S \to \{0, 1, \ldots, n\}$, and $f_{\Sigma} : S \to \{0, 1, \ldots, m \times n\}$ (with respect to $\mathbf{T}$) are defined as follows:*

- $f_{Min}(s) = \begin{cases} t_1(s) & \text{if } t_i(s) = t_j(s) \text{ for every } t_i, t_j \in \mathbf{T} \\ succ(\min\{t_i(s) \mid t_i \in \mathbf{T}\}) & \text{otherwise} \end{cases}$
  where $succ(x)$ is the successor of x in template B

- $f_{Max}(s) = \max\{t_i(s) \mid t_i \in \mathbf{T}\}$

- $f_{\Sigma}(s) = \sum\{t_i(s) \mid t_i \in \mathbf{T}\}$

The basic templated merging operations $\triangle_{Min}$, $\triangle_{Max}$, and $\triangle_{\Sigma}$ are defined using definition 6.15. From the definitions of indexing functions f_{Min}, f_{Max}, and f_{Σ}, it follows that templated merging operations $\triangle_{Min}$ and $\triangle_{Max}$ satisfy the principle of Qualitativeness while templated merging operation f_{Σ} , which makes use of addition, fails to do so.

The minimisation operation $\triangle_{Min}$ is inspired by the *(P,P)-capped* merging operation $\triangle_{Min_2}$ of Chopra, Ghose, and Meyer (2006). Its effect is to create a merged regular t-ordering in which the states are as low as possible by considering the minimum level assigned to a state by any of the t-orderings in the information tuple, but distinguishing the special case in which all of the t-orderings assign the same level to a state. The minimisation operation ensures that the level assigned to a state $s \in S$ in the merged result when the t-orderings in the information tuple are not all in agreement in terms of their assignment of s, is minimally higher than when all of the t-orderings in the information template are in agreement in terms of their assignment of s. This requirement ensures that property (**TM** $\triangle$ **2**) is satisfied but is not the most extreme form of minimisation. The minimise operation $\bigtriangledown$, which was defined in section 5.5, is the most extreme form of minimisation. It can be reformulated in terms of an indexing function $f_{\bigtriangledown}$ by taking $f_{\bigtriangledown}(s) = \min\{t_i(s) \mid t_i \in \mathbf{T}\}$, thus making implicit the process of normalisation, which is explicit in the original formulation of operation $\bigtriangledown$. The minimise operation $\bigtriangledown$ is however not a templated merging operation because it violates property (**TM** $\triangle$ **2**). To see this, take $\mathbf{T} = (t_1, t_2)$. By property 3 of proposition 5.11, it follows that $bottom(t_1 \bigtriangledown t_2) = bottom(t_1) \cup bottom(t_2)$. Property (**TM** $\triangle$ **2**), on the other hand, requires that if $\mathbf{T}$ is satisfiable, then $bottom(t_1 \bigtriangledown t_2) = bottom(t_1) \cap bottom(t_2)$.

In contrast to minimisation, the maximisation operation $\triangle_{Max}$ creates a merged regular t-ordering in which the states are as high as possible by considering the maximum level assigned to a state by any of the t-orderings in the information tuple. It is a generalisation of the family of knowledge base merging operations called Max operations by

Konieczny and Pino-Pérez and similar to Meyer's $\triangle_{Max}$ operation on epistemic states. The essential difference between the templated maximisation operation $\triangle_{Max}$ and that of Meyer lies in the notion of normalisation; Meyer's notion of normalisation can be described as an (explicit) downward shift of all states that leaves the distance between states unchanged while the notion of templated normalisation can be described as an (implicit) downward shift of non-definitely excluded states that leave the ordering between states unchanged.

The sum operation $\triangle_{\Sigma}$ is a generalisation of the family of knowledge base merging operations called Σ operations by Konieczny and Pino-Pérez. It is similar to Meyer's $\triangle_{\Sigma}$ operation on epistemic states but, again, differs in the way normalisation is performed.

The following example illustrates:

Example 6.2 *Let $\mathbf{T} = (t_1, t_2, t_3, t_4)$ be an information tuple where*
$t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$, $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$,
$t_3 = \{(11, 0), (10, 2), (01, 4), (00, 1)\}$, *and* $t_4 = \{(11, 1), (10, 2), (01, 2), (00, 0)\}$. *But then*
$f_{Min} = \{(11, 1), (10, 1), (01, 3), (00, 1)\}$ *with* $A_{Min} = (1, 3)$,
$f_{Max} = \{(11, 1), (10, 2), (01, 4), (00, 3)\}$ *with* $A_{Max} = (1, 2, 3, 4)$, *and*
$f_{\Sigma} = \{(11, 1), (10, 5), (01, 12), (00, 5)\}$ *with* $A_{\Sigma} = (1, 5, 12)$.
The merged regular t-orderings $\triangle_{Min}(\mathbf{T})$, $\triangle_{Max}(\mathbf{T})$, and $\triangle_{\Sigma}(\mathbf{T})$ are depicted in figure 6-1 below.

Lemma 6.1 *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Let $\triangle_{\odot}$ be a basic templated merging operation based on an indexing function $f_{\odot} : S \rightarrow \{0, 1, \ldots, k\}$ (with respect to $\mathbf{T}$). For every $s, s' \in S$,*

- *if $f_{\odot}(s) = f_{\odot}(s')$ then $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$ and*
- *if $f_{\odot}(s) < f_{\odot}(s')$ then $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$.*

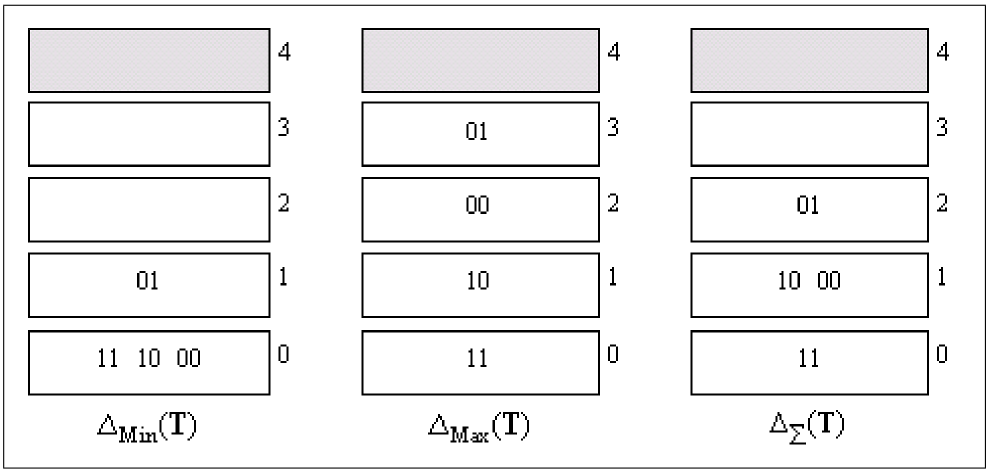

Figure 6-1: Basic templated merging

Proof. See proof in appendix D, section D.2. ■

Proposition 6.16 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Min}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

Proposition 6.17 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Max}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

Proposition 6.18 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{\Sigma}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

Using a similar method of construction, *refined* versions of the basic templated merging operations can be defined by using a refined indexing function. Following Meyer

(2001a), the idea behind the refined versions is that a merged regular t-ordering can be refined by creating a distinction between states at the same level without disturbing the relative ordering of states at different levels. However, since states at the top level of t-orderings are definitely excluded, no distinction will be created between states at the top level of the merged regular t-ordering. A *refined indexing function* $f_{R\odot} : S \rightarrow \{0, 1, \ldots, k\}^l$ (with respect to information tuple $\mathbf{T}$) assigns to each state $s \in S$ an l-tuple over $\{0, 1, \ldots, k\}$, defined abstractly as $f_{R\odot}(s) = (x_1, x_2, \ldots, x_l)$ ordered in accordance with $\odot$. Different refined indexing functions may have different codomains depending on the value of k, which represents the maximum index, and the value of l, which represents the indexing range. Every refined indexing function $f_{R\odot}$ induces a *refined index set* $A_{R\odot}$ defined exactly as before as $A_{R\odot} = ran(f_{R\odot})$. The difference is that $A_{R\odot}$ is a set of l-tuples over $\{0, 1, \ldots, k\}$, i.e. a subset of the Cartesian product $\{0, 1, \ldots, k\}^l$. Since $\{0, 1, \ldots, k\}$ is finite with $\leq$ a linear order on $\{0, 1, \ldots, k\}$, it follows from lemma 4.1 that the lexicographic ordering $\preceq$ on $A_{R\odot}$ is a well-ordering on $A_{R\odot}$.

Definition 6.17 *Let* $f_{R\odot} : S \rightarrow \{0, 1, \ldots, k\}^l$ *be a refined indexing function and* $A_{R\odot}$ *its associated index set, well-ordered by the lexicographic ordering* $\preceq$*. Let* $f_{R\odot}(s) = (x_1, x_2, \ldots, x_l)$ *and* $f_{R\odot}(s') = (y_1, y_2, \ldots, y_l)$*. Then* $f_{R\odot}(s) \prec f_{R\odot}(s')$ *iff* $x_1 < y_1$ *or* $[(\exists j > 1)(\forall i < j)\ x_i = y_i$ *and* $x_j < y_j]$*.*

A refined templated merging operation $\triangle_{R\odot}$, based on $f_{R\odot}$, can now be defined abstractly.

Definition 6.18 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Let* $f_{R\odot} : S \rightarrow \{0, 1, \ldots, k\}^l$ *be a refined indexing function (with respect to* $\mathbf{T}$*) and* $A_{R\odot}$ *its associated index set, well-ordered by the lexicographic ordering* $\preceq$*. A* ***refined templated merging operation*** $\triangle_{R\odot}$ *is defined as*

- $\triangle_{R\odot}(\mathbf{T})(s) = \begin{cases} card(seg(f_{R\odot}(s))) & \textit{if } first(f_{R\odot}(s)) < k \textit{ or } first(\min(A_{R\odot})) = k \\ n & \textit{otherwise} \end{cases}$
  where $seg(f_{R\odot}(s))$ *is the initial segment of* $f_{R\odot}(s) \in A_{R\odot}$
  and $first(x_1, x_2, \ldots, x_l)$ *returns the first element* x_1 *of the* l*-tuple.*

Proposition 6.19 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Every t-ordering produced by a refined templated merging operation* $\triangle_{R\odot}$ *is regular.*

Proof. See proof in appendix D, section D.2 ■

As was the case with basic templated merging operations, refined templated merging operations differ only in their choice of refined indexing function. To ensure that the refined merging operations do not create a merged regular t-ordering in which the states at the top level are refined, the refined indexing functions have to ensure that the first element of the l-tuple assigned to a state is the maximum index iff it is the maximum index assigned to that state by the corresponding indexing function. It is not hard to see that for every $s \in S$ the construction of $\triangle_{R\odot}(\mathbf{T})(s)$ depends only on the the lexicographic ordering $\preceq$ on the refined index set $A_{R\odot} = ran(f_{R\odot})$. Provided the refined indexing function $f_{R\odot}$ does not make use of arithmetic operations on template B, any refined merging operation $\triangle_{R\odot}$ will satisfy the principle of Qualitativeness.

Three refined templated merging operations are proposed, namely, *refined minimisation*, *refined maximisation*, and *refined sum*, denoted respectively by $\triangle_{RMin}$, $\triangle_{RMax}$, and $\triangle_{R\Sigma}$. They are based respectively on the refined indexing functions $f_{RMin} : S \rightarrow \{0, 1, \ldots, n\}^m$, $f_{RMax} : S \rightarrow \{0, 1, \ldots, n\}^m$, and $f_{R\Sigma} : S \rightarrow \{0, 1, \ldots, m \times n\}^n$.

Definition 6.19 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. The* ***refined*** *indexing functions* $f_{RMin} : S \rightarrow \{0, 1, \ldots, n\}^m$*,* $f_{RMax} : S \rightarrow \{0, 1, \ldots, n\}^m$*, and* $f_{R\Sigma} : S \rightarrow \{0, 1, \ldots, m \times n\}^n$ *(with respect to* $\mathbf{T}$*) are defined as follows:*

- $f_{RMin}(s) = (t_i(s) \mid t_i \in \mathbf{T})$ *ordered increasingly*

- $f_{RMax}(s) = (t_i(s) \mid t_i \in \mathbf{T})$ *ordered decreasingly*
- $f_{R\Sigma}(s) = (x_1, x_2, \ldots, x_n)$ *such that* $x_j = \sum\{t_i(s) \mid t_i \in \mathbf{T}$ *and* $t_i(s) > j-1\}$

The refined templated merging operations $\triangle_{RMin}$, $\triangle_{RMax}$, and $\triangle_{R\Sigma}$ are defined using definition 6.18. The effect of the refinement operations are to refine the merged regular t-ordering created by the corresponding basic templated merging operations by considering all of the levels assigned to a state by the t-orderings in the information tuple. From the definitions of refined indexing functions f_{RMin}, f_{RMax}, and $f_{R\Sigma}$, it follows that refined templated merging operations $\triangle_{RMin}$ and $\triangle_{RMax}$ satisfy the principle of Qualitativeness while, as would be expected, refined templated merging operation $f_{R\Sigma}$ fails to satisfy the principle of Qualitativeness.

The refined minimisation operation $\triangle_{RMin}$ and maximisation operation $\triangle_{RMax}$ are similar to Meyer's refinement operations $\triangle_{RMin}$ and $\triangle_{RMax}$ on epistemic states. Apart from different notions of normalisation, the templated refined minimisation and maximisation operations differ from those of Meyer in their treatment of states at the top level of the t-ordering obtained from the basic merging operations; these states, being definitely excluded, are not refined. Note that the refined maximisation operation $\triangle_{RMax}$ is a generalisation of the family of knowledge base merging operations called $GMax$ operations by Konieczny and Pino-Pérez.

The refined sum operation $\triangle_{R\Sigma}$ is similar in spirit to Meyer's refinement operation $\triangle_{R\Sigma}$ on epistemic states but, apart from differences in normalisation and in the treatment of definitely excluded states, it differs in the construction of the underlying ordering. Our construction creates n-tuples over $\{0, 1, \ldots, m \times n\}$ through the refined indexing function $f_{R\Sigma}$ which is then well-ordering by the lexicographic ordering $\preceq$ on $A_{R\Sigma}$. Meyer's construction creates a total preorder $\leq_{\mathbf{E}}$ on the set $Seq_{\mathbf{E}}$ of all sequences of length m of $\mathbb{N}$, ranging from 0 to $\max\{max(e_i) \mid e_i \in \mathbf{E}\}$, where $\mathbf{E} = [e_1, e_2, \ldots, e_m]$ is an epistemic list of epistemic states.

The following example illustrates the effect of each of the three refined templated merging operations:

Example 6.3 *Let* $\mathbf{T} = (t_1, t_2, t_3, t_4)$ *be the information tuple of example 6.2 (on page 257). But then,*
$f_{RMin} = \{(11,(0,0,0,1)),(10,(0,1,2,2)),(01,(2,2,4,4)),(00,(0,1,1,3))\}$
with $A_{RMin} = ((0,0,0,1),(0,1,1,3),(0,1,2,2),(2,2,4,4))$,
$f_{RMax} = \{(11,(1,0,0,0)),(10,(2,2,1,0)),(01,(4,4,2,2)),(00,(3,1,1,0))\}$
with $A_{RMax} = ((1,0,0,0),(2,2,1,0),(3,1,1,0),(4,4,2,2))$, *and*
$f_{R\Sigma} = \{(11,(1,0,0,0)),(10,(5,4,0,0)),(01,(12,12,8,8)),(00,(5,3,3,0))\}$
with $A_{R\Sigma} = ((1,0,0,0),(5,3,3,0),(5,4,0,0),(12,12,8,8))$.
The merged regular t-orderings $\triangle_{RMin}(\mathbf{T})$, $\triangle_{RMax}(\mathbf{T})$, *and* $\triangle_{R\Sigma}(\mathbf{T})$ *are depicted in figure 6-2 below.*

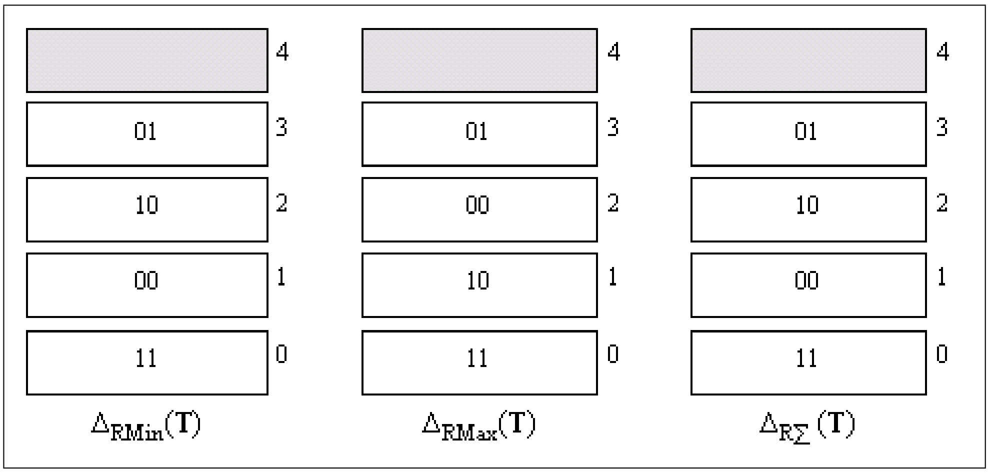

Figure 6-2: Refined templated merging

Lemma 6.2 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Let* $\triangle_{R\odot}$ *be a refined templated merging operation based on a refined indexing function* $f_{R\odot} : S \to \{0, 1, \ldots, k\}^l$

(with respect to $\mathbf{T}$*). For every* $s, s' \in S$*,*

- *if* $f_{R\odot}(s) = f_{R\odot}(s')$ *then* $\triangle_{R\odot}(\mathbf{T})(s) = \triangle_{R\odot}(\mathbf{T})(s')$
- *if* $f_{R\odot}(s) \prec f_{R\odot}(s')$ *and* $(first(f_{\odot}(s)) < k$ *or* $first(f_{\odot}(s')) < k$ *or* $first(\min(A_{R\odot})) = k)$ *then* $\triangle_{R\odot}(\mathbf{T})(s) < \triangle_{R\odot}(\mathbf{T})(s')$
- *if* $f_{R\odot}(s) \prec f_{R\odot}(s')$ *and* $(first(f_{\odot}(s)) = k$ *and* $first(f_{\odot}(s')) = k$ *and* $first(\min(A_{R\odot})) \neq k)$ *then* $\triangle_{R\odot}(\mathbf{T})(s) = \triangle_{R\odot}(\mathbf{T})(s')$

Proof. See proof in appendix D, section D.2. ■

Proposition 6.20 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RMin}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

Proposition 6.21 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RMax}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

Proposition 6.22 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{R\Sigma}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{5})$*.*

Proof. See proof in appendix D, section D.2. ■

These results, together with those for basic templated merging operations, support the view that properties $(\mathbf{TM} \triangle \mathbf{1})$ to $(\mathbf{TM} \triangle \mathbf{5})$ may be regarded as basic properties of templated merging operations. However, properties $(\mathbf{TM} \triangle \mathbf{1})$ to $(\mathbf{TM} \triangle \mathbf{5})$ cannot be viewed as a canonical list of properties for templated merging operations.

In addition to properties ($\mathbf{TM} \triangle \mathbf{1}$) to ($\mathbf{TM} \triangle \mathbf{5}$) for templated merging operations, properties for the subclass of templated majority merging operations and the subclass of templated arbitration operations are also defined. Let $\mathbf{T}_{t_j,k} = (t_1, t_2, \ldots, t_m, t_j, t_j, \ldots, t_j)$ be the information tuple obtained from the information tuple $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ by adding the regular t-ordering $t_j \in T_E$ to $\mathbf{T}$ exactly k times. Property ($\mathbf{TM} \triangle \mathbf{6}$) captures the idea behind majority merging while property ($\mathbf{TM} \triangle \mathbf{7}$) captures the idea behind arbitration.

($\mathbf{TM} \triangle \mathbf{6}$) $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle(\mathbf{T}_{t_j,k})(s) \leq \triangle(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$

($\mathbf{TM} \triangle \mathbf{7}$) $\forall t_j \in T_E \; \forall k \; \triangle(\mathbf{T}_{t_j,k}) = \triangle(\mathbf{T}_{t_j,1})$

Property ($\mathbf{TM} \triangle \mathbf{6}$) is the counterpart of property (**Maj**) for epistemic states and says that if a regular t-ordering $t_j \in T_E$ is added enough times to an information tuple, then the result of templated majority merging will reflect the ordering of t-ordering t_j. Property ($\mathbf{TM} \triangle \mathbf{7}$) is the counterpart for postulate ($\mathbf{KP} \triangle \mathbf{8'}$) and says that the result of templated arbitration is fully independent of the frequency with which a regular t-ordering appears in an information tuple. It is somewhat stricter than property (**Arb**) for epistemic states.

Proposition 6.23 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Property* ($\mathbf{TM} \triangle \mathbf{6}$) *is satisfied by templated merging operation* $\triangle_{\Sigma}$ *and* $\triangle_{R\Sigma}$ *but not satisfied by* $\triangle_{Min}$*,* $\triangle_{Max}$*,* $\triangle_{RMin}$*, and* $\triangle_{RMax}$*.*

Proof. See proof in appendix D, section D.2. ∎

Proposition 6.24 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Property* ($\mathbf{TM} \triangle \mathbf{7}$) *is satisfied by templated merging operations* $\triangle_{Min}$ *and* $\triangle_{Max}$*, but not satisfied by* $\triangle_{\Sigma}$*,* $\triangle_{RMin}$*,* $\triangle_{RMax}$*, and* $\triangle_{R\Sigma}$*.*

Proof. See proof in appendix D, section D.2. ■

From the results, it follows that templated merging operations $\triangle_{Min}$ and $\triangle_{Max}$ belong to the subclass of templated arbitration operations. This is consistent with the results of Meyer (2001a) whose operations $\triangle_{Min}$ and $\triangle_{Max}$ on epistemic states, albeit slightly different from ours, also belong to a subclass of arbitration operations. In contrast, Konieczny and Pino-Pérez's (1998) operation $\triangle_{Max}$ on knowledge bases does not belong to a subclass of arbitration operations. The difference is not unexpected given that property (**TM** $\triangle$ **7**) for templated arbitration and Meyer's property (**Arb**) for arbitration are both based on Konieczny and Pino-Pérez's original postulate (**KP** $\triangle$ **8**$'$) for arbitration as opposed to postulate (**KP** $\triangle$ **8**).

Templated merging operations $\triangle_{\Sigma}$ and $\triangle_{R\Sigma}$ belong to the subclass of templated majority merging operations, a result that is consistent with that of Meyer. It supports the view that property (**TM** $\triangle$ **6**) may be regarded as an appropriate property for the subclass of templated majority merging operations.

Lastly, the results show that the refined templated merging operations $\triangle_{RMin}$ and $\triangle_{RMax}$ belong to neither of the subclasses of templated merging operations. Again, this is consistent with Meyer's results for the refinement operations $\triangle_{RMin}$ and $\triangle_{RMax}$ on epistemic states.

The templated merging operations that were defined do not rely on a notion of distance between possible worlds as is the case with most semantic methods for constructing (knowledge base) merging operations. Rather, they rely on the notion of an indexing function (refined indexing function) and an associated well-ordered index set (refined index set) to construct a basic templated merging operation (refined templated merging operation). Because these templated merging operations are defined abstractly, it is possible to construct different kinds of templated merging operations. However, the abstract definitions of basic and refined templated merging operations do not guarantee that the principle of Qualitativeness will be satisfied. Indeed, both templated majority merging operations that were defined fail to satisfy the principle of Qualitativeness. Because of the importance we attach to the principle of Qualitativeness, two new purely qualitative templated majority merging operations will be constructed in the next section, both of which are inspired by the notion of an infobase.

6.6 Content-based merging

The notion of information communicated by independent sources is closely linked to the concept of an *infobase* (Meyer, Labuschagne, and Heidema, 2000b; Meyer, 2001b). Formally, an infobase IB is a finite list of sentence $[\alpha_1, \alpha_2, \ldots, \alpha_m]$ representing information obtained from independent sources by a specific point in time. Each sentence in an infobase is associated with one and only one source. This means that repeated communications from a single source would be represented cumulatively by a single sentence.

Every infobase defines a total preorder on the set of states, which is used in the construction of infobase change operations (Meyer, 2001b). The total preorder is defined according to the number of sentences in the infobase that the states satisfy; the more sentences a state satisfies, the lower down in the ordering the state will be. For $s \in S$, the *IB-number* s_{IB} of s is defined as the number of sentences α_i in the infobase IB such that $s \in Mod(\alpha_i)$. The total preorder R on S is then obtained by taking $(s, s') \in R$ iff $s'_{IB} \leq s_{IB}$.

Since every sentence $\alpha \in L$ induces a definite t-ordering t_α such that $bottom(t_\alpha) = Mod_M(\alpha)$ and $top(t_\alpha) = NMod_M(\alpha)$, where $M = \langle S, l \rangle$ is an extensional interpretation of L, and since every definite t-ordering $t_\alpha \in T_D$ has a syntactic expression in the form of a sentence $\alpha \in L$ that is a finite axiomatisation of $bottom(t_\alpha)$ under M, it follows that every infobase $IB = [\alpha_1, \alpha_2, \ldots, \alpha_m]$ can be represented as an information tuple

$\mathbf{T} = (t_{\alpha_1}, t_{\alpha_2}, \ldots, t_{\alpha_m})$ and vice versa, provided none of the sentences in the infobase is a contradiction. (A contradiction $\perp$ induces a definite t-ordering $t_\perp$ that is not regular.)

To generate a t-ordering from an infobase, as opposed to a total preorder on the set of states, provision must be made for the exclusion of states. This will be accomplished by generating the t-ordering from the nonmodels (or semantic content) of the sentences in the infobase, instead of from the models of the sentences. Recall from semantic information theory that the (semantic) content of a sentence α is defined as $Cont_M(\alpha) = NMod_M(\alpha)$, under an extensional interpretation $M = \langle S, l \rangle$. The fewer sentences a state is a nonmodel of, the lower down in the t-ordering it will be. If a state is an element of the (semantic) content of every sentence in the infobase, then the state will be definitely excluded in the t-ordering generated from the infobase.

Definition 6.20 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and let $IB = [\alpha_1, \alpha_2, \ldots, \alpha_m]$ be an infobase. For $s \in S$, the IB-**content** of s is defined as $cont_{IB} : S \rightarrow \{0, 1, \ldots, m\}$ by $cont_{IB}(s) = card(\{\alpha_i \in IB \mid s \in Cont_M(\alpha_i)\})$.*

An IB-content function can be seen as a form of indexing function (with respect to IB) from which an index set $A_{IB} \subseteq \{0, 1, \ldots, m\}$ can be defined by taking $A_{IB} = ran(cont_{IB})$. (Since A_{IB} is finite and $\leq$ is a linear order on A_{IB} it follows that $\leq$ is a well-ordering on A_{IB}.) The t-ordering generated from an infobase IB will be denoted by $\mathbf{T}_{IB}$.

Definition 6.21 *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$ and let $IB = [\alpha_1, \alpha_2, \ldots, \alpha_m]$ be an infobase. Then the t-ordering $\mathbf{T}_{IB}$ generated from infobase IB is defined as*

- $\mathbf{T}_{IB}(s) = \begin{cases} card(seg(cont_{IB}(s))) & \textit{if } cont_{IB}(s) < m \\ n & \textit{otherwise} \end{cases}$

  where $seg(cont_{IB}(s))$ is the initial segment of $cont_{IB}(s) \in A_{IB}$.

The idea of generating a t-ordering from the (semantic) content of the sentences in an infobase will serve as the inspiration for constructing a new type of basic templated merging operation, called a *content-based* templated merging operation, by using the same method as for basic templated merging operations. A *content-based* indexing function $f_{Cont} : S \to \{0, 1, \ldots, m\}$ (with respect to information tuple $\mathbf{T}$) assigns to each state the number of t-orderings in $\mathbf{T}$ in which the state is excluded. Note that exclusion here means both definitely excluded and tentatively excluded. This is because content-based templated merging operations should apply to every information tuple, and not only to information tuples that correspond to infobases. If a state is excluded by every t-ordering in the information tuple, definitely or tentatively, then the state will be definitely excluded in the merged regular t-ordering.

Definition 6.22 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. The* ***content-based*** *indexing function* $f_{Cont} : S \to \{0, 1, \ldots, m\}$ *(with respect to* $\mathbf{T}$*) is defined as* $f_{Cont}(s) = card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\})$.

The *content-based templated merging operation* $\triangle_{Cont}$ based on indexing function $f_{Cont} : S \to \{0, 1, \ldots, m\}$ is defined by using definition 6.15. From the definition of content-based indexing function f_{Cont} and the fact that for every $s \in S$ the construction of $\triangle_{\odot}(\mathbf{T})(s)$ depends only on the the linear order $\leq$ on the index set $A_{\odot} = ran(f_{\odot})$, it follows that content-based templated merging operation $\triangle_{Cont}$ satisfies the principle of Qualitativeness. The effect of the content-based operation $\triangle_{Cont}$ is to create a merged regular t-ordering in which the states reflect the combined indefinite content of all t-orderings in the information tuple by considering the number of t-orderings in the information tuple for which a state $s \in S$ is tentatively (or definitely) excluded by a t-ordering $t_i \in \mathbf{T}$, i.e. for which $s \in Cont_0(t_i)$. The following example illustrates:

Example 6.4 *Let* $\mathbf{T} = (t_1, t_2, t_3, t_4)$ *be the information tuple of example 6.2 (on page*

257) where $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$, $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$, $t_3 = \{(11,0),(10,2),(01,4),(00,1)\}$, *and* $t_4 = \{(11,1),(10,2),(01,2),(00,0)\}$. *But then* $f_{Cont} = \{(11,1),(10,3),(01,4),(00,3)\}$ *with* $A_{Cont} = (1,3,4)$. *The merged regular t-ordering* $\triangle_{Cont}(\mathbf{T})$ *is depicted in figure 6-3.*

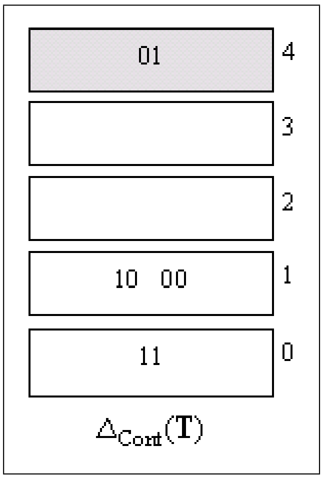

Figure 6-3: Content-based merging

The main difference between $\triangle_{Cont}$ and the other basic templated merging operations is that a state will be definitely excluded in the merged t-ordering if it is excluded, definitely or tentatively, by every t-ordering in the information tuple. In the case of $\triangle_{Min}$ and $\triangle_{\Sigma}$, a state will only be definitely excluded in the merged t-ordering if it is definitely excluded by every t-ordering in the information tuple whereas in the case of $\triangle_{Max}$, a state will be definitely excluded in the merged t-ordering if it is definitely excluded by some t-ordering in the information tuple. (Note, however, that the definition of basic templated merging operations ensures that in the event of every state being targeted for definite exclusion in the merged t-ordering, the resulting merged t-ordering would be tautological rather than strongly contradictory and thus regular as shown by proposition

6.15.) Despite these differences, content-based templated merging operations also satisfy properties ($\mathbf{TM} \triangle \mathbf{1}$) to ($\mathbf{TM} \triangle \mathbf{5}$).

Proposition 6.25 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Cont}$ *satisfies properties* ($\mathbf{TM} \triangle \mathbf{1}$) *to* ($\mathbf{TM} \triangle \mathbf{6}$) *but fails to satisfy property* ($\mathbf{TM} \triangle \mathbf{7}$).

Proof. See proof in appendix D, section D.3. ■

This result strengthens the view that properties ($\mathbf{TM} \triangle \mathbf{1}$) to ($\mathbf{TM} \triangle \mathbf{5}$) may be regarded as basic properties of templated merging operations. From the proposition it follows that basic content-based merging operation $\triangle_{Cont}$ is a new type of templated majority merging operation.

In the case where the information tuple comprises only definite (regular) t-orderings, the result of a content-based templated merging operation is precisely the t-ordering generated from the infobase corresponding to the information tuple.

Proposition 6.26 *Let* $M = \langle S, l \rangle$ *be an extensional interpretation of* L *with* $card(S) = n$ *and let* $IB = [\alpha_1, \alpha_2, \ldots, \alpha_m]$ *be an infobase such that no* $\alpha_i \equiv \bot$. *Let* $\mathbf{T} = (t_{\alpha_1}, t_{\alpha_2}, \ldots, t_{\alpha_m})$ *be the information tuple corresponding to* IB. *Then* $\mathbf{T}_{IB} = \triangle_{Cont}(\mathbf{T})$.

Proof. Let $\alpha \in L$. If $s \in Cont_M(\alpha)$ then by definition 3.3, $s \in NMod_M(\alpha)$. But then $s \in top(t_\alpha)$, i.e. $t_\alpha(s) > 0$. So $card(\{\alpha_i \in IB \mid s \in Cont_M(\alpha_i)\}) = card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\})$. But then $card(seg(f_{Cont}(s))) = card(seg(cont_{IB}(s)))$, and hence by definitions 6.21 and 6.15, it follows that $\mathbf{T}_{IB} = \triangle_{Cont}(\mathbf{T})$. ■

As was the case with the other basic templated merging operations, it is possible to define a refined version of content-based merging operations based on the definition of a refined content-based indexing function.

Definition 6.23 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. The* ***refined content-based*** *indexing function* $f_{RCont} : S \to \{0, 1, \ldots, m\}^n$ *(with respect to* $\mathbf{T}$*) is defined as* $f_{RCont}(s) = (x_1, x_2, \ldots, x_n)$ *such that* $x_j = card(\{t_i \in \mathbf{T} \mid t_i(s) > j - 1\})$*.*

The *refined content-based templated merging operation* $\triangle_{RCont}$ based on indexing function $f_{RCont} : S \to \{0, 1, \ldots, m\}^n$ is defined by using definition 6.18. From the definition of refined content-based indexing function f_{RCont} and the fact that for every $s \in S$ the construction of $\triangle_{R\odot}(\mathbf{T})(s)$ depends only on the the lexicographic ordering $\preceq$ on the refined index set $A_{R\odot} = ran(f_{R\odot})$, it follows that refined content-based templated merging operation $\triangle_{RCont}$ satisfies the principle of Qualitativeness. The effect of the refined content-based operation $\triangle_{RCont}$ is to refine the merged regular t-ordering obtained from the basic content-based merging operation $\triangle_{Cont}$ by creating a distinction between states at the same level without disturbing the relative ordering of states at different levels or refining the set of definitely excluded states. The following example illustrates:

Example 6.5 *Let* $\mathbf{T} = (t_1, t_2, t_3, t_4)$ *be the information tuple of example 6.2 (on page 257) where* $t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$*,* $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$*,* $t_3 = \{(11, 0), (10, 2), (01, 4), (00, 1)\}$*, and* $t_4 = \{(11, 1), (10, 2), (01, 2), (00, 0)\}$*. But then* $f_{RCont} = \{(11, (1, 0, 0, 0)), (10, (3, 2, 0, 0)), (01, (4, 4, 2, 2)), (00, (3, 1, 1, 0))\}$
with $A_{RCont} = ((1, 0, 0, 0), (3, 1, 1, 0), (3, 2, 0, 0), (4, 4, 2, 2))$*.*
The merged regular t-ordering $\triangle_{RCont}(\mathbf{T})$ *is depicted in figure 6-4 below.*

Proposition 6.27 *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RCont}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{6})$ *but fails to satisfy property* $(\mathbf{TM} \triangle \mathbf{7})$*.*

Proof. See proof in appendix D, section D.3. ■

From the results it follows that the refined content-based merging operation $\triangle_{RCont}$ is another templated merging operation that belongs to the subclass of templated majority

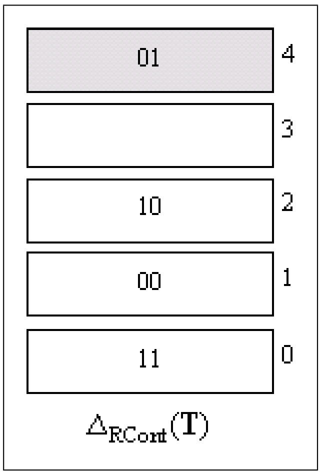

Figure 6-4: Refined content-based merging

merging operations. Both content-based merging operations depend on the principle of Trustworthiness, which says that the information received by agents is trustworthy. Therefore, a state may be definitely excluded in the merged result if it is excluded, even tentatively, by every agent (or sub-agent) communicating information. However, a state which is tentatively excluded in the merged result of a content-based merging operation must have been excluded, at least tentatively, by at least one agent (or sub-agent) communicating information.

Proposition 6.28 *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Then, the following holds for each $s \in S$:*

1. $s \in Cont_D(\triangle_{Cont}(\mathbf{T}))$ *iff* $s \in Cont_0(t_i)$ *for every* $t_i \in \mathbf{T}$
2. *if* $s \in Cont_0(\triangle_{Cont}(\mathbf{T}))$ *then* $s \in Cont_0(t_i)$ *for some* $t_i \in \mathbf{T}$
3. $s \in Cont_D(\triangle_{RCont}(\mathbf{T}))$ *iff* $s \in Cont_0(t_i)$ *for every* $t_i \in \mathbf{T}$
4. *if* $s \in Cont_0(\triangle_{RCont}(\mathbf{T}))$ *then* $s \in Cont_0(t_i)$ *for some* $t_i \in \mathbf{T}$

Proof. See proof in appendix D, section D.3. ■

It may be argued that content-based merging operations $\triangle_{Cont}$ and $\triangle_{RCont}$ are too credulous. For agents operating in a competitive setting where information communicated by other agents is not guaranteed to be trustworthy, such as in games, this may be true. But for control room agents operating in a cooperative setting where information communicated by other agents (or sub-agents) is trustworthy, it need not be true. However, should a more sceptical content-based merging operation be required, the content-based indexing function may be modified to increase the threshold of exclusion. In other words, a content-based indexing function $f_{Cont,j} : S \rightarrow \{0, 1, \ldots, m\}$ may be defined by taking $f_{Cont,j}(s) = card(\{t_i \in \mathbf{T} \mid t_i(s) > j \text{ for } 0 < j < n\})$ where j indicates the threshold of exclusion. The most credulous of content-based indexing functions would be $f_{Cont,0}$ whereas the most sceptical of content-based indexing functions would be $f_{Cont,n-1}$. Albeit interesting, the idea of sceptical content-based merging operations will not be pursued any further.

Chapter 7

Conclusion

Anyway I'll make an end, for I want to.... ... Is it really the end?

Raskolnikov in *Crime and Punishment* (Fyodor Dostoyevsky)

7.1 Summary

One of the goals of the thesis has been to emphasise the benefits of taking an agent-oriented view in the broader area of knowledge representation and reasoning. The agent-oriented view led to the demarcation of a specific class of systems, the class of diagrammable systems, and a compatible class of agents, the class of control room agents. These determined the choice of knowledge representation languages, namely, finitely generated transparent propositional languages under a possible worlds semantics (as opposed to the classical truth-value semantics). By taking an agent-oriented view, subtleties were discovered in the general area of belief change that may otherwise have been left undiscovered.

Within the broader agent-oriented perspective, an information-theoretic approach was adopted, which gave rise to the concept of a templated ordering (t-ordering), a novel semantic construct that allows for the combination of both definite and indefinite information in the same structure. One of the attractive features of this data structure is that (different classes of) t-orderings are expressible (by different normal forms) in the knowledge representation language and recoverable from such. This was accomplished by extending the obvious state description normal form or SDNF (semantic content normal form or SCNF) into state description cumulative form or SDCF (semantic content cumulative form or SCCF) and state description templated form or SDTF (semantic content templated form or SCTF).

Finding appropriate representations of the epistemic states of agents is an important issue in the area of knowledge representation and reasoning. The information-theoretic representation of an epistemic state as a regular t-ordering is different from most other approaches, even the ordinal conditional functions of Spohn, in the sense that it allows for the representation of both knowledge and (degrees of) belief. In accordance with the agent-oriented view, knowledge is not viewed in the Platonic tradition as 'true justified belief', which would make sense only from the perspective of an omniscient godlike superagent, but is instead related to a psychological notion of 'entrenchment of belief'. No attempt is made to specify a particular degree of entrenchment that constitutes the threshold of conviction at which mere belief becomes knowledge. Not only is this likely to be context-dependent, but a full account must take into account emotions (or emotion-analogs), since it would seem that one of the mechanisms whereby belief is replaced by knowledge is the recruitment of emotions. The process of converting a tentative preference into a deeply entrenched conviction that the agent is unwilling to give up in the face of argument has, in contexts such as vegetarianism and beliefs about the desirability of cigarette smoking, been called 'moralization' and associated with the emotion of disgust by Rozin (1997). Evidence that such convictions do not yield to mere argument is cited in Haidt (2001). For present purposes, therefore, we are content to rely on the evidence that there is indeed a point at which tentative belief is sufficiently deeply entrenched to acquire the status of a conviction.

Similarly to beliefs, convictions may be surrendered, although not lightly. For example, an agent may somehow be brought to realise that circumstances have changed. In everyday human social intercourse this realisation may come about in a variety of ways, for example as a result of cognitive dissonance (Festinger, 1957). For control room agents the relevant mechanism is paradox-resolution - the agent resolves the paradox created by the reception of trustworthy new information that is inconsistent with what the agent strongly believes to be fact, by realising that the state of the system has changed. This mechanism is built into the proposed epistemic change algorithm. The algorithm provides a control room agent with a novel mechanism to choose between different epistemic change operations, most notably, between revision and update.

An area of general neglect in the representation of epistemic states is the construction of an initial epistemic state. This has been addressed by defining a refinement operation on t-orderings in normal form and using it to refine the agent's fixed information by its default rules about the system.

The area of belief change has been an active research area for at least two decades with the major focus on belief revision and, to a lesser extent, on belief (or knowledge base) update. One of the features of templated revision, apart from extending the notion of belief revision to include knowledge, is that a notion of informational gain (or loss) is available through the definition of the (semantic) information content of t-orderings. As has been shown, templated revision results in monotonic growth of the control room agent's knowledge and nonmonotonic growth of its beliefs. The epistemic change algorithm proved instrumental in providing an alternative justification for rejecting the controversial postulate (**C2**) for iterated revision and in casting some shadow over recent claims that the DP-postulates are overly permissive.

The main contribution of the thesis to the area of belief change lies, however, in templated update. Recognising that the distinctive postulate for update, called the Disjunction Rule, requires an operation to be defined on epistemic states so that the rule holds on the associated knowledge bases, a modification of the KM postulates for update was proposed which allows for the transition from update as a function of belief sets to update as a function of epistemic states. The other main insight about update stems from adopting an agent-oriented view in which diagrammable systems are defined as discrete systems. This led to the subtle discovery that belief update has essentially been cast in the realm of continuous systems. Recasting epistemic update into the realm of discrete systems resulted in effectively rejecting one of the KM postulates for update and gave rise to the notion of an epistemic transition function. Templated revision and templated update support iterated epistemic change in a unified framework and, for each, a representation result has been proved and a concrete operation constructed.

In recent times, research in the area of belief change has tended to shift more towards belief merging and the field of information fusion. In accordance with our focus on control room agents, only the Pareto principle from Arrow's set of conditions in social choice theory was adopted as a rationality property for templated merging. Through the definition of an abstract basic templated merging operation and an associated abstract refined templated merging operation, a flexible foundation was provided for constructing merging operations that do not rely on any notion of distance. Several concrete instantiations of these abstract templated merging operations were constructed, thereby illustrating the flexibility of the foundation. Based on ideas from infobase change, the proposal for content-based merging gave rise to a new family of purely qualitative templated majority merging operations. Using the foundation provided by the abstract templated merging operations, two concrete content-based merging operations were constructed, which could potentially be transformed into more 'sceptical' forms of content-based merging.

7.2 Future research

In this thesis, a control room agent is viewed as a first-order intentional system having both informational attitudes and motivational attitudes, but the focus has been exclusively on the informational attitudes of knowledge and belief. Of particular interest would be to (attempt to) model the role that emotion plays in the reasoning (Simon, 1967; Oatley and Johnson-Laird, 1987; Picard, 1997; Moore and Oaksford, 2002) of control room agents. One possible avenue of exploration is to augment the epistemic framework of control room agents with an ordering on sentences that reflects a preference ordering of 'emotion', which may serve to direct and limit (by conserving resources) the reasoning of the agent. The ordering may be induced from a t-ordering by a suitable power construction, such as those used in verisimilitude (Brink and Heidema, 1987; Burger and Heidema, 1995). Whether these ideas will prove useful remains to be seen. Another line of research would be to extend the notion of a control room agent as a first-order intentional system by formalising some of the motivational attitudes of a control room agent.

Moving away from a single agent, an obvious next step would be to generalise templated interpretations under the information-theoretic semantics for epistemic logic so as to represent the epistemic states of multiple agents. In this connection, the use of t-orderings has an advantage to offer, specially in respect of the treatment of agents' fixed information. Given a single agent, the usual treatment of fixed information is to permanently exclude from all consideration the states that are (known to be) unrealisable. However, given multiple agents, each with limited information, it becomes necessary to record the differences in these agents' fixed information when modelling their epistemic states. Such differences cannot be recorded if the usual treatment is used because it does not provide a means for choosing which states to permanently exclude amongst those known by different agents (or even an omniscient godlike superagent) to be unrealisable. In the context of multiple agents, t-orderings make it possible to explicitly represent difference in the agents' fixed information.

The information-theoretic approach to belief change is purely qualitative and, as such, avoids the problems associated with quantitative approaches relating to numbers, fundamental among which is the problem 'Where do the numbers come from?'. However, as pointed out by Fermé and Rott (2004), quantitative approaches are more expressive than qualitative approaches in allowing degrees of acceptance of the input sentence, for example, Spohn's conditionalisation $k * (\alpha, m)$. Whilst in agreement with the arguments against the use of numbers, Fermé and Rott also agree that there should be ways of specifying the strength of new beliefs. Their proposal, called *revision by comparison*, is to provide a qualitative analogue in the form of a reference sentence β which calls for an input sentence α to be accepted with a degree of plausibility that is at least equal to that of β, by using an epistemic entrenchment ordering $\sqsubseteq$. The revision by comparison operation $K * (\alpha \sqsubseteq \beta)$ instructs the agent to ensure that the entrenchment of β is at least as firm as the entrenchment of α in the revised epistemic state. From an agent-oriented view, a key question is how the reference sentence (degree of acceptance) is determined and by whom, given that agents obtain information in the form of observations and communications from other agents. Incorporating qualitative analogues for such quantitative constructs into templated revision and templated update seems a useful area for future research provided, from an agent-oriented view, that it is meaningful for the class of agents under consideration.

Belief merging and belief revision, whether at the level of belief sets or epistemic states, are generally viewed as alternative forms of belief change. However, at the level of epistemic states, iterated revision can also be viewed as a form of prioritised merging (Maynard-Reid II and Shoham, 2001; Delgrande, Dubois, and Lang, 2006). In the context of information fusion, where the notion of belief merging relates to the combination of (possibly conflicting) information received from multiple sources, yet another view emerges. In this view, which can be traced back to the Decision Ladder of Rasmussen (1986), the result of belief merging (combining information) serves as the input for belief revision (as well as for belief expansion and belief update). The intuitive meaning is that the control room agent is presented with both certain and uncertain information. Since the result of templated merging is a regular t-ordering, it means that the epistemic change algorithm, and hence all the templated operations, will have to be redefined to accept as input not a definite t-ordering but a regular t-ordering. Technically, templated update may be redefined as a templated maximisation operation whereas templated revision may be redefined by a suitable modification of the partial refinement operation. The potential implications are not yet clear, especially for templated update. From an agent-oriented perspective, the notion of scepticism (credulousness) should form part of any proper account.

The last area for future research to be considered is the computational complexity (Papadimitriou, 1994) of the proposed templated operations for epistemic change and for combining (merging) information. The complexity of deciding whether a sentence is a semantic consequence of a revised (updated) knowledge base are studied by Eiter and Gottlob (1992) while Liberatore and Schaerf (2001) study whether a possible world is a model of a revised (updated) knowledge base. The analysis of Cadoli, Donini, Liberatore, and Schaerf (1999) focusses on a specific computational aspect, namely, the size of the revised (updated) knowledge base. Liberatore (1997) considers the complexity of iterated revision operations. Some complexity results for distance-based merging operations is studied by Konieczny, Lang, and Marquis (2002) and it would be constructive to see whether the complexity results for templated merging operations differ significantly from those of distance-based merging operations.

Appendix A

Proofs for Chapter 3

A.1 Proofs for Section 3.2

Proposition A.1 (3.7) *Let X be a set and let R be a total preorder on X. Then R produces a strict linearly ordered partition of X.*

Proof. To construct the partition of X, start by constructing an equivalence relation $\approx$ on X from R. Let $\approx$ be the relation on X given by $x \approx y$ iff $(x, y) \in R$ and $(y, x) \in R$. Then $\approx$ is reflexive on X by reflexivity of R, symmetric by definition of $\approx$, and transitive by transitivity of R, i.e. $\approx$ is an equivalence relation on X. The equivalence classes form the relevant partition P_R of X. Let $<$ be the relation on P_R defined as $[x] < [y]$ iff $(x, y) \in R$ and $(y, x) \notin R$. (Note that $<$ is independent of choice of representatives. To see this suppose $[x] < [y]$. So $(x, y) \in R$ and $(y, x) \notin R$. Choose any $x' \in [x]$ and $y' \in [y]$. Since $x \approx x'$, $(x, x') \in R$ and $(x', x) \in R$. Similarly, $(y, y') \in R$ and $(y', y) \in R$. By transitivity of R, $(x', y') \in R$. Suppose $(y', x') \in R$. By transitivity of R, $(y, x) \in R$. Contradiction. So $(y', x') \notin R$.) It is easy to see that the relation $<$ is irreflexive and antisymmetric. To see that $<$ is transitive, suppose that $[x] < [y]$ and $[y] < [z]$. So $(x, y) \in R$ and $(y, x) \notin R$ and $(y, z) \in R$ and $(z, y) \notin R$. By transitivity of R, $(x, z) \in R$.

But $(z, x) \notin R$ otherwise, by transitivity of R, $(z, y) \in R$ resulting in a contradiction. Hence, $[x] < [z]$. To see that $<$ is connected, pick any $[x], [y] \in P_R$. Suppose $[x] \not< [y]$. But then, by definition of $<$, $(x, y) \notin R$ or $(y, x) \in R$ or both. If both, then $[y] < [x]$. If $(x, y) \notin R$ but $(y, x) \notin R$ we have a contradiction since R is total. If $(y, x) \in R$ but $(x, y) \in R$, then $x \approx y$ and hence $[x] = [y]$. Thus, for every $[x], [y] \in P_R$, either $[x] < [y]$ or $[y] < [x]$ or $[x] = [y]$. So $<$ is a strict linear order on P_R. ■

Proposition A.2 (3.9) *Let X be a set and let Q be a strict modular partial order on X. Then Q produces a strict linearly ordered partition of X.*

Proof. To construct the partition of X, begin by constructing an equivalence relation $\simeq$ on X from Q. Let $\simeq$ be the relation on X given by $x \simeq y$ iff $(x, y) \notin Q$ and $(y, x) \notin Q$. By irreflexivity of Q, for every $x \in X$, $(x, x) \notin Q$. So for every $x \in X$, $x \simeq x$, i.e. $\simeq$ is reflexive on X. Suppose $x \simeq y$. So $(x, y) \notin Q$ and $(y, x) \notin Q$. But then $y \simeq x$, i.e. $\simeq$ is symmetric. Suppose $x \simeq y$ and $y \simeq z$. So $(x, y) \notin Q$ and $(y, x) \notin Q$, and $(y, z) \notin Q$ and $(z, y) \notin Q$. Suppose $(z, x) \in Q$. But $(x, y) \notin Q$ and $(y, x) \notin Q$. So, by the first condition of modularity of Q, $(z, y) \in R_S$. Contradiction. So $(z, x) \notin Q$. Similarly, $(x, z) \notin Q$. So $x \simeq z$, i.e. $\simeq$ is transitive. But then $\simeq$ is an equivalence relation on X. The equivalence classes form the relevant partition P_Q of X. Let $\prec$ be the relation on P_Q defined as $[x] \prec [y]$ iff $(x, y) \in Q$. (Note that $\prec$ is independent of choice of representatives. To see this suppose $[x] \prec [y]$. So $(x, y) \in Q$. Choose any $x' \in [x]$ and $y' \in [y]$. Since $x \simeq x'$, $(x, x') \notin Q$ and $(x', x) \notin Q$ by definition. Similarly, since $y \simeq y'$, $(y, y') \notin Q$ and $(y', y) \notin Q$. By the first condition of modularity of Q, $(x, y') \in Q$. By the second condition of modularity of Q, either $(x', y') \in Q$ or $(x, x') \in Q$. But $(x, x') \notin Q$. So $(x', y') \in Q$.) To show that $\prec$ is linear and strict, it must be shown to be irreflexive, antisymmetric, transitive, and connected. But $\prec$ is irreflexive on P_Q by irreflexivity of Q, antisymmetric by antisymmetry of Q, and transitive by transitivity of Q. To see that

$\prec$ is connected, pick any $[x], [y] \in P_Q$. Suppose $[x] \not\prec [x]$, $[y] \not\prec [x]$, and $[x] \neq [y]$. So $(x, y) \notin Q$ and $(y, x) \notin Q$ and $x \neq y$. Then, by definition of $\simeq$, $x \simeq y$. So $[x] = [y]$. Contradiction. So $\prec$ is connected as well. But then $\simeq$ is a strict linear order on P_Q. ■

Proposition A.3 (3.10) *Let X be a set, R a total preorder on X, and Q the corresponding strict modular partial order on X. Then R and Q produce the same strict linearly ordered partition of X.*

Proof. By proposition 3.7, R produces a a strict linearly ordered partition $<$ on P_R and by proposition 3.9, Q produces a strict linearly ordered partition $\prec$ on P_Q where $Q = R - \{(x, y) \mid (y, x) \in R\}$. It must be shown that $P_R = P_Q$ and that the relations $<$ on P_R and $\prec$ on P_Q are equal. Suppose $x \approx y$. Thus, by definition, $(x, y) \in R$ and $(y, x) \in R$. But then $(x, y) \notin Q$ and $(y, x) \notin Q$ since $Q = R - \{(x, y) \mid (y, x) \in R\}$. So, by definition of $\simeq$, $x \simeq y$. Conversely, suppose $x \simeq y$. Thus, by definition, $(x, y) \notin Q$ and $(y, x) \notin Q$. But R is total and thus $(x, y) \in R$ and $(y, x) \in R$. So, by definition of $\approx$, $x \approx y$. To show that the relations $<$ on P_R and $\prec$ on P_Q are equal it must be shown that if $[x] < [y]$, then $[x] \prec [y]$ and, conversely, if $[x] \prec [y]$, then $[x] < [y]$. Choose any $[x] < [y]$. By definition of $<$, $(x, y) \in R$ and $(y, x) \notin R$. But $Q = R - \{(x, y) | (y, x) \in R\}$. So $(x, y) \in Q$. Thus, by definition of $\prec$, $[x] \prec [y]$. Conversely, choose any $[x] \prec [y]$. So $(x, y) \in Q$. But $Q \subseteq R$ and so $(x, y) \in R$. Suppose $(y, x) \in R$. By definition of Q, $(x, y) \notin Q$. Contradiction. So $(y, x) \notin R$. Thus $[x] < [y]$. It may therefore be concluded that R and Q produce the same strict linearly ordered partition of X. ■

A.2 Proofs for Section 3.3

Proposition A.4 (3.13) *Let $t \in T_S$, $i, j \in B$, and $s \in S$. Then the following constraints hold:*

1. $bottom(t) \subseteq get_{\uparrow}(t, j)$ *and* $top(t) \subseteq get_{\downarrow}(t, j)$

2. $get_{\rightarrow}(t, i) \subseteq get_{\uparrow}(t, j)$ *if* $i \leq j$ *and* $get_{\rightarrow}(t, i) \subseteq get_{\downarrow}(t, j)$ *if* $i \geq j$

3. $get_{\uparrow}(t, j) \cup get_{\downarrow}(t, j) = S$ *and* $get_{\uparrow}(t, j) \cap get_{\downarrow}(t, j) = get_{\rightarrow}(t, j)$

4. $pull(push(t, s, j), s, t(s)) = t$

5. $push(pull(t, s, j), s, t(s)) = t$

Proof. The first three results follow directly from the definitions.

4. Suppose $j > t(s)$. Then $push(t, s, j) = t \circledast \{(s, j)\}$. So $push(t, s, j)(s) = j$. Thus $push(t, s, j)(s) > t(s)$. But then $pull(push(t, s, j), s, t(s)) = push(t, s, j) \circledast \{(s, t(s))\}$. So $pull(push(t, s, j), s, t(s))(s) = t(s)$. But only s was involved in the overriding and thus $pull(push(t, s, j), s, t(s)) = t$. Now suppose that $j \leq t(s)$. So $push(t, s, j) = t$, i.e. $push(t, s, j)(s) = t(s)$. But then $pull(push(t, s, j), s, t(s)) = t$.

5. Suppose $j < t(s)$. Then $pull(t, s, j) = t \circledast \{(s, j)\}$. So $pull(t, s, j)(s) = j$. Thus $pull(t, s, j)(s) < t(s)$. But then $push(pull(t, s, j), s, t(s)) = pull(t, s, j) \circledast \{(s, t(s))\}$. So $push(pull(t, s, j), s, t(s))(s) = t(s)$. But only s was involved in the overriding and thus $push(pull(t, s, j), s, t(s)) = t$. Now suppose that $j \geq t(s)$. So $pull(t, s, j) = t$, i.e. $pull(t, s, j)(s) = t(s)$. But then $push(pull(t, s, j), s, t(s)) = t$. ∎

Lemma A.1 (3.1) *Let A be a non-empty proper subset of template B and let g be the normalise function for A. Then $g(x) < n$ for every $x \in A$ such that $x \neq n$.*

Proof. Pick any element $x \in A$ such that $x \neq n$. So $g(x) = card(seg(x))$. But A is well-ordered and so $x = succ(seg(x))$ and $n = succ(seg(n))$. Since $x < n$ it follows that $seg(x) \subset seg(n)$ and thus $card(seg(x)) < card(seg(n))$. But $seg(n) \cup \{n\} = B$ and so $card(seg(n)) = n$. Hence $card(seg(x)) < n$, i.e. $g(x) < n$. ∎

Lemma A.2 (3.2) *Let A be a non-empty proper subset of template B and let g be the normalise function for A. Then g is injective.*

Proof. If A is a singleton then g is trivially injective. Suppose A is not a singleton. Pick any elements $x, x' \in A$. To prove that g is injective it must be shown that if $x \neq x'$ then $g(x) \neq g(x')$. Assume that $x \neq x'$. Suppose further that $x \neq n$ and $x' \neq n$. Since A is well-ordered it follows that $x = succ(seg(x))$ and $x' = succ(seg(x'))$. But then $seg(x) \neq seg(x')$ otherwise $x = x'$. So $card(seg(x)) \neq card(seg(x'))$ and thus $g(x) \neq g(x')$. Now suppose that $x = n$. Then $g(x) = n$ and $g(x') < n$ (by lemma 3.1) and so $g(x) \neq g(x')$. Similarly when it is supposed that $x' = n$. Hence g is injective. ■

Lemma A.3 (3.3) *Let A be a non-empty proper subset of template B and let g be the normalise function for A. If $j < n$ and $j \in ran(g)$, then $k \in ran(g)$ for every $0 \leq k < j$.*

Proof. Suppose $j < n$ and $j \in ran(g)$. But g is injective (by lemma 3.2). So there is exactly one $x \in A$ such that $g(x) = card(seg(x)) = j$. Let $g|seg(x)$ be the restriction of g to $seg(x)$. We shall show that $ran(g|seg(x)) = \{0, 1, \ldots, j-1\}$. To see this pick any $k \in ran(g|seg(x))$. But g is injective (by lemma 3.2). So there is exactly one $x' \in seg(x)$ such that $g(x') = card(seg(x')) = k$. But $x' < x$ and since A is well-ordered it follows that $\varnothing \subseteq seg(x') \subset seg(x)$. But then $0 \leq card(seg(x')) < card(seg(x))$, i.e. $0 \leq g(x') < g(x)$, i.e. $0 \leq k < j$. So $k \in \{0, 1, \ldots, j-1\}$. Conversely, pick any $k \in \{0, 1, \ldots, j-1\}$. Let x' be the predecessor of x. Since A is well-ordered it follows that $x' \in seg(x)$ and $seg(x') = seg(x) - \{x'\}$. So $card(seg(x')) = j-1$. If $j-1 \neq k$ then the process repeats until $card(seg(x'')) = k$ for some $x'' \in seg(x))$. So $k \in ran(g|seg(x))$. Thus $ran(g|seg(x)) = \{0, 1, \ldots, j-1\}$. So if it is the case that $j < n$ and $g(x) = j$ then $ran(g|seg(x)) = \{0, 1, \ldots, j-1\}$. In other words, if $j < n$ and $j \in ran(g)$, then $k \in ran(g)$ for every $0 \leq k < j$. ■

A.3 Proofs for Section 3.4

Lemma A.4 (3.4) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = \bigwedge_{i=0}^{n} \beta_i$ of L in SDCF, $Mod_M(\gamma) = Mod_M(\beta_0)$.*

Proof. Pick any sentence $\gamma = \bigwedge_{i=0}^{n} \beta_i$ of L in SDCF. By definition of SDCF, each β_i is a sentence in extensional SDNF. Pick any state description, say σ_j, occurring in β_i. Then, since γ is in SDCF, σ_j occurs in every β_k where $k > i$. But for every state description σ_j in β_i, there exists a corresponding state $s_j \in S$ such that $Mod_M(\sigma_j) = \{s_j\}$ (because M is extensional). So $Mod_M(\beta_i) \subseteq Mod_M(\beta_k)$ for every $k > i$. But then $\bigcap_{i=0}^{n} Mod_M(\beta_i) = Mod_M(\beta_0)$. Since $Mod_M(\phi \wedge \psi) = Mod_M(\phi) \cap Mod_M(\psi)$ for every $\phi, \psi \in L$, it follows that $Mod_M(\gamma) = \bigcap_{i=0}^{n} Mod_M(\beta_i)$. But then $Mod_M(\gamma) = Mod_M(\beta_0)$. ■

Lemma A.5 (3.5) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = \bigwedge_{i=0}^{m} \beta_i$ of L in SCCF, $Mod_M(\gamma) = Mod_M(\beta_0)$.*

Proof. Pick any sentence $\gamma = \bigwedge_{i=0}^{m} \beta_i$ of L in SCCF. By definition of SCCF, each β_i is a sentence in extensional SCNF. Pick any content element, say ε_j, occurring in β_i. Then, since γ is in SCCF, ε_j occurs in every β_k where $k < i$. But for every content element ε_j in β_i, there exists a corresponding state $s_j \in S$ such that $NMod_M(\varepsilon_j) = \{s_j\}$ (because M is extensional). So $NMod_M(\beta_i) \subseteq NMod_M(\beta_k)$ for every $k < i$. Hence $Mod_M(\beta_k) \subseteq Mod_M(\beta_i)$ for every $k < i$. But then $\bigcap_{i=0}^{m} Mod_M(\beta_i) = Mod_M(\beta_0)$. Since $Mod_M(\phi \wedge \psi) = Mod_M(\phi) \cap Mod_M(\psi)$ for every $\phi, \psi \in L$, it follows that $Mod_M(\gamma) = \bigcap_{i=0}^{n} Mod_M(\beta_i)$. But then $Mod_M(\gamma) = Mod_M(\beta_0)$. ■

Lemma A.6 (3.6) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ of L in SDTF, $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$.*

Proof. Pick any sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n} \beta_i)$ of L in SDTF. By definition of SDTF, each β_i is a sentence in extensional SDNF. Pick any state description, say σ_j, occurring in β_i. Then, since γ is in SDTF, σ_j occurs in every β_k where $k > i$. But for every state description σ_j in β_i, there exists a corresponding state $s_j \in S$ such that $Mod_M(\sigma_j) = \{s_j\}$ (because M is extensional). So $Mod_M(\beta_i) \subseteq Mod_M(\beta_k)$ for every $k > i$. Let $\delta = \bigvee_{i=0}^{n-1} \beta_i$. Since $Mod_M(\phi \vee \psi) = Mod_M(\phi) \cup Mod_M(\psi)$ for every $\phi, \psi \in L$, it follows that $Mod_M(\delta) = \bigcup_{i=0}^{n-1} Mod_M(\beta_i)$. But $Mod_M(\beta_i) \subseteq Mod_M(\beta_k)$ for every $k > i$ and thus $Mod_M(\delta) = Mod_M(\beta_{n-1})$. Hence $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$. ■

Lemma A.7 (3.7) *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Then for every sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ of L in SCTF, $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$.*

Proof. Pick any sentence $\gamma = (\bigvee_{i=0}^{n-1} \beta_i) \wedge (\bigwedge_{i=0}^{n-1} \beta_i)$ of L in SCTF. By definition of SCTF, each β_i is a sentence in extensional SCNF. Pick any content element, say ε_j, occurring in β_i. Then, since γ is in SCCF, ε_j occurs in every β_k where $k < i$. But for every content element ε_j in β_i, there exists a corresponding state $s_j \in S$ such that $NMod_M(\varepsilon_j) = \{s_j\}$ (because M is extensional). So $NMod_M(\beta_i) \subseteq NMod_M(\beta_k)$ for every $k < i$. Hence $Mod_M(\beta_k) \subseteq Mod_M(\beta_i)$ for every $k < i$. Let $\delta = \bigvee_{i=0}^{n-1} \beta_i$. Since $Mod_M(\phi \vee \psi) = Mod_M(\phi) \cup Mod_M(\psi)$ for every $\phi, \psi \in L$, it follows that $Mod_M(\delta) = \bigcup_{i=0}^{n-1} Mod_M(\beta_i)$. But $Mod_M(\beta_k) \subseteq Mod_M(\beta_i)$ for every $k < i$ and thus $Mod_M(\delta) = Mod_M(\beta_{n-1})$. Hence $Mod_M(\bigvee_{i=0}^{n-1} \beta_i) = Mod_M(\beta_{n-1})$. ■

A.4 Proofs for Section 3.6

Lemma A.8 (3.8) *Let $T = \langle S, t, l\rangle$ be a templated interpretation of L and let $\alpha \in L$. Then $Min_t(\alpha) = \{s \in Mod_T(\alpha) \mid t(s) = dt(t, \alpha)\}$.*

Proof. If $Min_t(\alpha) = \varnothing$ then the result holds trivially. Suppose $Min_t(\alpha) \neq \varnothing$. Choose any $s \in Min_t(\alpha)$. So $s \in Mod_T(\alpha)$ and there is no $s' \in Mod_T(\alpha)$ such that $t(s') \leq t(s)$ unless $t(s) \leq t(s')$. But $s \in get_{\uparrow}(t, t(s))$. So $s \in get_{\uparrow}(t, t(s)) \cap Mod_T(\alpha)$. But $t(s)$ is the least j such that $get_{\uparrow}(t, j) \cap Mod_T(\alpha) \neq \varnothing$ otherwise $s \notin Min_t(\alpha)$. But then $t(s) = dt(t, \alpha)$. Since s was chosen arbitrarily it follows that $Min_t(\alpha) \subseteq \{s \in Mod_T(\alpha) \mid t(s) = dt(t, \alpha)\}$.

Conversely, if $Mod_T(\alpha) = \varnothing$ then the result holds trivially. Suppose $Mod_T(\alpha) \neq \varnothing$. Choose any $s \in Mod_T(\alpha)$ such that $t(s) = dt(t, \alpha)$. Suppose $dt(t, \alpha) = j$ for $j \leq n$. But then $get_{\uparrow}(t, j) \cap Mod_T(\alpha) \neq \varnothing$ and $get_{\uparrow}(t, i) \cap Mod_T(\alpha) = \varnothing$ for every $i < j$. It is claimed there is no $s' \in Mod_T(\alpha)$ such that $t(s') \leq j$ unless $j \leq t(s')$ otherwise it cannot be the case that $get_{\uparrow}(t, i) \cap Mod_T(\alpha) = \varnothing$ for every $i < j$. But then $s \in Min_t(\alpha)$. Since s was chosen arbitrarily, it follows that $\{s \in Mod_T(\alpha) \mid t(s) = dt(t, \alpha)\} \subseteq Min_t(\alpha)$. But then $Min_t(\alpha) = \{s \in Mod_T(\alpha) \mid t(s) = dt(t, \alpha)\}$. ■

Lemma A.9 (3.9) *Let $T = \langle S, t, l \rangle$ be a templated interpretation of L and let $\alpha \in L$ and $s \in S$. If $s \in Mod_T(\alpha \wedge \beta)$ and $s \in Min_t(\alpha)$, then $s \in Min_t(\alpha \wedge \beta)$.*

Proof. Let $s \in Mod_T(\alpha \wedge \beta)$ and $s \in Min_t(\alpha)$. By lemma 3.8, $s \in Mod_T(\alpha)$ and $t(s) = dt(t, \alpha)$. Suppose $s \notin Min_t(\alpha \wedge \beta)$. So there must be some $s' \in Mod_T(\alpha \wedge \beta)$ such that $t(s') < t(s)$. So $t(s') < dt(t, \alpha)$. But then $s' \notin Mod_T(\alpha)$. Contradiction. So $s \in Min_t(\alpha \wedge \beta)$. ■

Appendix B

Proofs for Chapter 4

B.1 Proofs for Section 4.3

Proposition B.1 (4.6) *Let $\langle S, F\rangle$ be a frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_{n-1}T_{n-1}4_{n-1}5_{n-1}$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_{n-1}, T_{n-1}, 4_{n-1}, or 5_{n-1} iff for every $s, s' \in S$,*

1. *s is accessible from itself and*

2. *if s' is accessible from s, then $Cont_D(F(s')) = Cont_D(F(s))$.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of $K_{n-1}T_{n-1}4_{n-1}5_{n-1}$. Then by proposition 4.5 it follows that for every $s \in S$, s is accessible from itself. To show that condition 2 holds, suppose that there are $s, s' \in S$ such that $s' \in get_{\uparrow}(F(s), n-1)$ but $Cont_D(F(s')) \neq Cont_D(F(s))$, i.e. $get_{\uparrow}(F(s), n-1) \neq get_{\uparrow}(F(s'), n-1)$. Then a templated interpretation based on $\langle S, F\rangle$ can be constructed that is not a model of all the instances of schema 5_{n-1}, in particular the sentence $\neg[n{-}1]P(a) \rightarrow [n{-}1]\neg[n{-}1]P(a)$. Take $T = \langle S, F, l\rangle$ such that $v_{l(s')}(P(a)) = 0$ for state $s' \in get_{\uparrow}(F(s), n{-}1)$ and $v_{l(s'')}(P(a)) = 1$ for every $s'' \in get_{\uparrow}(F(s'), n-1)$. So $P(a)$ is not satisfied at state $s' \in get_{\uparrow}(F(s), n-1)$

and hence $[n-1]P(a)$ is not satisfied at state s from which it follows that $\neg[n-1]P(a)$ is satisfied at state s. However, $P(a)$ is satisfied at every state $s'' \in get_{\uparrow}(F(s'), n-1)$ and thus $[n-1]P(a)$ is satisfied at state s' from which it follows that $\neg[n-1]P(a)$ is not satisfied at state s'. But $s' \in get_{\uparrow}(F(s), n-1)$ and thus $[n-1]\neg[n-1]P(a)$ is not satisfied at state s. But then s fails to satisfy $\neg[n-1]P(a) \rightarrow [n-1]\neg[n-1]P(a)$ at state s. So $\langle S, F\rangle$ cannot be a frame-model of $\neg[n-1]P(a) \rightarrow [n-1]\neg[n-1]P(a)$. Contradiction. Hence, if s' is accessible from s, then $Cont_D(F(s')) = Cont_D(F(s))$.

Conversely, suppose that for every $s, s' \in S$, s is accessible from itself and if s' is accessible from s, then $Cont_D(F(s')) = Cont_D(F(s))$. It must be shown that $\langle S, F\rangle$ is a frame-model of $K_{n-1}T_{n-1}4_{n-1}5_{n-1}$. By propositions 4.4 and 4.5 $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schemas K_{n-1} and T_{n-1} respectively.
To show that $\langle S, F\rangle$ is a a frame-model of every sentence that is an instance of schema 4_{n-1} choose any sentence $\gamma \in L'$ of the form $[n-1]\alpha \rightarrow [n-1][n-1]\alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $T = \langle S, F, l\rangle$ of L' and some state $s \in S$ such that s fails to satisfy γ in T. So it must be the case that $[n-1]\alpha$ is satisfied at s but $[n-1][n-1]\alpha$ is not. If $[n-1][n-1]\alpha$ is not satisfied at s then there must be some state $s' \in get_{\uparrow}(F(s), n-1)$ at which $[n-1]\alpha$ is not satisfied. Hence there must be some state $s'' \in get_{\uparrow}(F(s'), n-1)$ that fails to satisfy α. Recall that $[n-1]\alpha$ is satisfied at s. So it must be the case that α is satisfied at every state $s''' \in get_{\uparrow}(F(s), n-1)$, including s'. But $get_{\uparrow}(F(s'), n-1) = get_{\uparrow}(F(s), n-1)$ and thus $s'' \in get_{\uparrow}(F(s), n-1)$ from which it follows that α is satisfied at s''. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema 4_{n-1}.
To show that $\langle S, F\rangle$ is a a frame-model of every sentence that is an instance of schema 5_{n-1} choose any sentence $\gamma \in \mathcal{L}'$ of the form $\neg[n-1]\alpha \rightarrow [n-1]\neg[n-1]\alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $\mathcal{T} = \langle S, F, l\rangle$

of $\mathcal{L}'$ and some state $s \in S$ such that s fails to satisfy γ in $\mathcal{T}$. So it must be the case that $\neg[n-1]\alpha$ is satisfied at s but $[n-1]\neg[n-1]\alpha$ is not. If $[n-1]\neg[n-1]\alpha$ is not satisfied at s then there must be some state $s' \in get_{\uparrow}(F(s), n-1)$ at which $\neg[n-1]\alpha$ is not satisfied, i.e. at which $[n-1]\alpha$ is satisfied. So it must be the case that α is satisfied at every $s'' \in get_{\uparrow}(F(s'), n-1)$. Recall that $\neg[n-1]\alpha$ is satisfied at s and thus $[n-1]\alpha$ is not. So there must be some $s''' \in get_{\uparrow}(F(s), n-1)$ that fails to satisfy α. But since $s' \in get_{\uparrow}(F(s), n-1)$ it follows that $get_{\uparrow}(F(s), n-1) = get_{\uparrow}(F(s'), n-1)$ and hence that $s''' \in get_{\uparrow}(F(s'), n-1)$. But then α must be satisfied at s'''. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema 5_{n-1}.

■

Proposition B.2 (4.7) *Let $\langle S, F\rangle$ be a frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_0D_04_05_0$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_0, D_0, 4_0, or 5_0 iff*

1. *for every $s \in S$ there exists some $s' \in S$ such that s is accessible to degree 0 from s' and*
2. *for every $s, s' \in S$, if s' is accessible to degree 0 from s, then $Cont_0(F(s')) = Cont_0(F(s))$.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of $K_0D_04_05_0$. To show that condition 1 holds, choose any $s \in S$. It must be shown that there is some state $s' \in S$ such that $s' \in get_{\uparrow}(F(s), 0)$. Suppose it is not the case. So $get_{\uparrow}(F(s), 0) = \varnothing$. Then a templated interpretation based on $\langle S, F\rangle$ can be constructed that is not a model of all the instances of schema D_0, in particular the sentence $[0]P(a) \rightarrow \neg[0]\neg P(a)$. Since $get_{\uparrow}(F(s), 0) = \varnothing$, it hold vacuously that $P(a)$ is satisfied at every $s' \in get_{\uparrow}(F(s), 0)$ and that $\neg P(a)$ is satisfied at every $s' \in get_{\uparrow}(F(s), 0)$. So $[0]P(a)$ is satisfied at s and $[0]\neg P(a)$ is satisfied at s. But if $[0]\neg P(a)$ is satisfied at s then $\neg[0]\neg P(a)$ fails to be satisfied at s. But then $\langle S, F\rangle$ cannot be a frame-model of $[0]P(a) \rightarrow \neg[0]\neg P(a)$. Contradiction. Hence, there must be some state $s' \in S$ such that $s' \in get_{\uparrow}(F(s), 0)$. To show that condition 2 holds, suppose that there are $s, s' \in S$ such that $s' \in get_{\uparrow}(F(s), 0)$ but $Cont_0(F(s')) \neq Cont_0(F(s))$. The proof can now be constructed in exactly the same way as the proof for proposition 4.6 by constructing a templated interpretation based on $\langle S, F\rangle$ that is not a model of all the instances of schema 5_0, using sentence $\neg[0]P(a) \rightarrow [0]\neg[0]P(a)$ as an instances of schema 5_0.

Conversely, suppose that conditions 1 and 2 hold. It must be shown that $\langle S, F\rangle$ is a frame-model of $K_0D_04_05_0$. By proposition 4.4, $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schema K_0. To show that $\langle S, F\rangle$ is a a frame-model of every sentence that is an instance of schema D_0 choose any sentence $\gamma \in \mathcal{L}'$ of the form $[0]\alpha \rightarrow \neg[0]\neg\alpha$. Suppose $\langle S, F\rangle$ is not a frame-model of γ. So there must be some templated interpretation $\mathcal{T} = \langle S, F, l\rangle$ of $\mathcal{L}'$ and some state $s \in S$ such that s fails to satisfy γ in $\mathcal{T}$. So it must be the case that $[0]\alpha$ is satisfied at s but $\neg[0]\neg\alpha$ is not. If $\neg[0]\neg\alpha$ is not satisfied at s, then $[0]\neg\alpha$ must be satisfied at s and hence, $\neg\alpha$ must be satisfied at every state $s' \in get_{\uparrow}(F(s), 0)$. Since $get_{\uparrow}(F(s), 0) \neq \varnothing$ it follows that α fails to be satisfied at every $s' \in get_{\uparrow}(F(s), 0)$. But then $[0]\alpha$ cannot be satisfied at s. Contradiction. Hence, $\langle S, F\rangle$ is a frame-model of the arbitrarily chosen sentence γ and therefore of every instance of schema D_0. In addition, $\langle S, F\rangle$ can be shown to be a frame-model of every sentence that is an instance of schema 4_0 by choosing any sentence $\gamma \in \mathcal{L}'$ of the form $[0]\alpha \rightarrow [0][0]\alpha$ and following the same argument as for the proof of schema 4_{n-1} in proposition 4.6. Similarly, by choosing any sentence $\gamma \in \mathcal{L}'$ of the form $\neg[0]\alpha \rightarrow [0]\neg[0]\alpha$ and using the same argument as for the proof of schema 5_{n-1} in proposition 4.6, it can be shown that $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schema 5_0. ■

Proposition B.3 (4.10) *Let $\langle S, F\rangle$ be a frame of $\mathcal{L}'$. Then $\langle S, F\rangle$ is a frame-model of the set $K_iT_{n-1}4_{Ki}5_{Ki}$ of sentences of $\mathcal{L}'$ that are instances of one of the schemas K_i, T_{n-1}, 4_{Ki}, or 5_{Ki} iff for every $s, s' \in S$,*

1. *s is accessible from itself and*

2. *if s' is accessible to degree i from s, then $Cont_i(F(s)) = Cont_i(F(s'))$ and $F(s')(x) = F(s)(x)$ for every $x \in get_\uparrow(F(s'), i)$.*

Proof. Suppose that $\langle S, F\rangle$ is a frame-model of $K_iT_{n-1}4_{Ki}5_{Ki}$. Then by proposition 4.5 it follows that for every $s \in S$, s is accessible from itself. So condition 1 holds. For every $s, s' \in S$, if s' is accessible to degree i from s, then by proposition 4.8 it holds that $Cont_i(F(s)) \subseteq Cont_i(F(s'))$ and $F(s')(x) \geq F(s)(x)$ for every $x \in get_\uparrow(F(s'), i)$ and by proposition 4.9 it holds that $Cont_i(F(s')) \subseteq Cont_i(F(s))$ and $F(s)(x) \geq F(s')(x)$ for every $x \in get_\uparrow(F(s), i)$. But then condition 2 holds.

Conversely, suppose that for every $s, s' \in S$, s is accessible from itself and if s' is accessible to degree i from s, then $Cont_i(F(s)) = Cont_i(F(s'))$ and $F(s')(x) = F(s)(x)$ for every $x \in get_\uparrow(F(s'), i)$. It must be shown that $\langle S, F\rangle$ is a frame-model of $K_iT_{n-1}4_{Ki}5_{Ki}$. By propositions 4.4 and 4.5 $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schemas K_i and T_{n-1} respectively. Proposition 4.8 shows that $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schema 4_{Ki} while 4.9 shows that $\langle S, F\rangle$ is a frame-model of every sentence that is an instance of schema 5_{Ki}. ■

B.2 Proofs for Section 4.6

Lemma B.1 (4.1) *Suppose X and Y are two finite sets on which there are linear orders $\leq_X$ and $\leq_Y$ respectively. Then the lexicographic ordering $\preceq$ on any subset A of the*

Cartesian product $X \times Y$ is a well-ordering on A.

Proof. It will be shown that $\preceq$ is linear on A and use the finiteness of A to conclude that $\preceq$ is a well-ordering on A. To show that $\preceq$ is reflexive on A pick any element $(x, y) \in A$. Then (by reflexivity of $\leq_X$ and $\leq_Y$) $x = x$ and $y = y$. So $(x, y) = (x, y)$ thus establishing the reflexivity of $\preceq$ on R. To show that $\preceq$ is antisymmetric suppose that $(x, y) \preceq (x', y')$ and $(x, y) \neq (x', y')$. So $x \leq_X x'$ and if $x = x'$ then $y <_Y y'$ for otherwise it would be the case that $(x, y) = (x', y')$. But then $((x', y'), (x, y)) \notin \preceq$. Thus $\preceq$ is antisymmetric. To show that $\preceq$ is transitive suppose that $(x, y) \preceq (x', y')$ and $(x', y') \preceq (x'', y'')$. So $x \leq_X x'$ and if $x = x'$ then $y \leq_Y y'$. Furthermore, $x' \leq_X x''$ and if $x' = x''$ then $y' \leq_Y y''$. But then (by transitivity of $\leq_X$), $x \leq_X x''$. Suppose $x = x''$. So $x = x'$ and $x' = x''$. But then, by transitivity of $\leq_Y$, $y \leq_Y y''$. So $(x, y) \preceq (x'', y'')$, i.e. $\preceq$ is transitive. To show that $\preceq$ is linear pick any elements $(x, y), (x', y') \in R$. But $\leq_X$ and $\leq_Y$ are both linear and so either $x \preceq_X x'$ or $x' \leq_X x$ and either $y \leq_X y'$ or $y' \leq_X y$. Suppose that $x \preceq_X x'$. If $x = x'$ then either $(x, y) \preceq (x', y')$ or $(x, y') \preceq (x', y)$. Now suppose that $x' \leq_X x$. If $x = x'$ then either $(x', y) \preceq (x, y')$ or $(x', y') \preceq (x, y)$. Thus, $\preceq$ is a linear order on A. But A is finite and so $\preceq$ is a well-ordering on A. ∎

Proposition B.4 (4.15) *Let S be a non-empty set of states with $card(S) = n$. Let $t_1, t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2 \rangle$ and well-ordered by the lexicographic ordering $\preceq$. Then the following holds for every $s, s' \in S$:*

1. $t_1(s) \leq (t_1 \boxminus t_2)(s)$
2. *if* $t_1(s) < t_1(s')$ *then* $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$
3. *if* $t_1(s) = t_1(s') \neq n$ *and* $t_2(s) < t_2(s')$ *then* $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$
4. $bottom(t_1 \boxminus t_2) \subseteq bottom(t_1)$

5. $Cont_D(t_1) = Cont_D(t_1 \boxminus t_2)$

6. $Cont_0(t_1) \subseteq Cont_0(t_1 \boxminus t_2)$

Proof. 1. Suppose that $t_1(s) = n$. But then $(t_1 \boxminus t_2)(s) = n$ and thus $t_1(s) = (t_1 \boxminus t_2)(s)$. Suppose that $t_1(s) = j$ where $j < n$. But t_1 is in normal form and so there must be at least j elements (i, k) in A such that $i = 0, 1, \ldots, j-1$ and $0 \leq k \leq n$. But then $card(seg(t_1(s), t_2(s))) \geq j$, i.e. $(t_1 \boxminus t_2)(s) \geq j$. Hence $t_1(s) \leq (t_1 \boxminus t_2)(s)$.

2. Let $t_1(s) < t_1(s')$. So $t_1(s) \neq n$. Suppose that $(t_1 \boxminus t_2)(s) \geq (t_1 \boxminus t_2)(s')$. So $card(seg(t_1(s), t_2(s))) \geq card(seg(t_1(s'), t_2(s')))$. But then $(t_1(s'), t_2(s')) \preceq (t_1(s), t_2(s))$. So $t_1(s') \leq t_1(s)$. Contradiction. Hence $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$.

3. Let $t_1(s) = t_1(s') \neq n$ and $t_2(s) < t_2(s')$. Suppose $(t_1 \boxminus t_2)(s) \geq (t_1 \boxminus t_2)(s')$. So $card(seg(t_1(s), t_2(s))) \geq card(seg(t_1(s'), t_2(s')))$. But then $(t_1(s'), t_2(s')) \preceq (t_1(s), t_2(s))$. So since $t_1(s') = t_1(s)$, it follows that $t_2(s') \leq t_2(s)$. Contradiction. Hence $(t_1 \boxminus t_2)(s) < (t_1 \boxminus t_2)(s')$.

4. It must be shown that $bottom(t_1 \boxminus t_2) \subseteq bottom(t_1)$. Choose any $s \in bottom(t_1 \boxminus t_2)$. So $(t_1 \boxminus t_2)(s) = 0$. But then $card(seg(t_1(s), t_2(s))) = 0$. Since t_1 is in normal form it follows that $t_1(s) = 0$. So $s \in bottom(t_1)$. Since s was chosen arbitrarily it follows that $bottom(t_1 \boxminus t_2) \subseteq bottom(t_1)$.

5. It must be shown that $top(t_1) = top(t_1 \boxminus t_2)$. Choose any $s \in top(t_1)$. So $t_1(s) = n$. But then $(t_1 \boxminus t_2)(s) = n$, i.e. $s \in top(t_1 \boxminus t_2)$. Since s was chosen arbitrarily it follows that $top(t_1) \subseteq top(t_1 \boxminus t_2)$. Conversely, choose any $s \in top(t_1 \boxminus t_2)$. Then $top(t_1 \boxminus t_2) \subseteq top(t_1)$ follows similarly. Hence $top(t_1) = top(t_1 \boxminus t_2)$.

6. It must be shown that $get(t_1, 1, n) \subseteq get(t_1 \boxminus t_2, 1, n)$. Choose any $s \in get(t_1, 1, n)$. So $t_1(s) \geq 1$. But $t_1(s) \leq (t_1 \boxminus t_2)(s)$ by the first condition. So $(t_1 \boxminus t_2)(s) \geq 1$. But then $s \in get(t_1 \boxminus t_2, 1, n)$. So $get(t_1, 1, n) \subseteq get(t_1 \boxminus t_2, 1, n)$. ■

Appendix C

Proofs for Chapter 5

C.1 Proofs for Section 5.4

Proposition C.1 (5.1) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E, t_{E'} \in T_E$ and let $\alpha, \alpha' \in L$ be consistent with $bel(t_E)$ and $bel(t_{E'})$ respectively. Then it holds that if $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$ then $t_E + \alpha = t_{E'} + \alpha'$.*

Proof. Let $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$. Let $X = Mod_M(\alpha)$ and $S - X = NMod_M(\alpha)$. So for every $s, s' \in X$, $t_E(s) \leq t_E(s')$ iff $(t_E + \alpha)(s) \leq (t_E + \alpha)(s')$. Since $t_E(s) = t_{E'}(s)$ for every $s \in S$, it follows that for every $s, s' \in X$, $(t_E + \alpha)(s) \leq (t_E + \alpha)(s')$ iff $(t_{E'} + \alpha)(s) \leq (t_{E'} + \alpha)(s')$. Hence $t_E + \alpha$ and $t_{E'} + \alpha$ are order-equivalent on X. But $Mod_M(know(t_E + \alpha)) = Mod_M(know(t_E) \wedge \alpha) = Mod_M(know(t_E)) \cap Mod_T(\alpha)$ and thus $S - X \not\subseteq Mod_M(know(t_E + \alpha))$, i.e. $S - X \not\subseteq get_{\uparrow}(t_E + \alpha, n - 1)$. So for every $s, s' \in S - X$, $(t_E + \alpha)(s) = n$. Since $t_E(s) = t_{E'}(s)$ for every $s \in S$, it follows that for every $s, s' \in S - X$, $(t_{E'} + \alpha)(s) = n$. Hence $t_E + \alpha$ and $t_{E'} + \alpha$ are order-equivalent on $S - X$. But $\alpha \equiv_M \alpha'$ and thus $X = Mod_M(\alpha')$ and $S - X = NMod_M(\alpha')$. But then $t_E + \alpha$ and $t_{E'} + \alpha'$ are order-equivalent. Since both $t_E + \alpha$ and $t_{E'} + \alpha'$ are regular, it follows by proposition 3.17 that $t_E + \alpha = t_{E'} + \alpha'$. ■

Proposition C.2 (5.2) *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be consistent with $bel(t_E)$. Then the following holds:*

1. $Cont_D(t_E) \subseteq Cont_D(t_E + \alpha)$
2. $Cont_0(t_E) \subseteq Cont_0(t_E + \alpha)$

Proof. 1. If $Cont_D(t_E) = \varnothing$ then the result holds vacuously. Assume that $Cont_D(t_E) \neq \varnothing$. Choose any $s \in Cont_D(t_E)$. So $s \in top(t_E)$, i.e. $s \notin get_\uparrow(t_E, n-1)$, i.e. $s \notin Mod_M(know(t_E))$, i.e. $s \notin Mod_M(know(t_E)) \cap Mod_M(\alpha)$, i.e. $s \notin Mod_M(know(t_E + \alpha))$, i.e. $s \notin get_\uparrow(t_E + \alpha, n-1)$. So $s \in top(t_E + \alpha)$. Since s was chosen arbitrarily, it follows that $Cont_D(t_E) \subseteq Cont_D(t_E + \alpha)$.

2. If $Cont_0(t_E) = \varnothing$ then the result holds vacuously. Assume that $Cont_0(t_E) \neq \varnothing$. Choose any $s \in Cont_0(t_E)$. So $s \in get(t_E, 1, n)$, i.e. $s \notin bottom(t_E)$, i.e. $s \notin Mod_M(bel(t_E))$, i.e. $s \notin Mod_M(bel(t_E)) \cap Mod_M(\alpha)$, i.e. $s \notin Mod_M(bel(t_E + \alpha))$, i.e. $s \notin bottom(t_E + \alpha)$. So $s \in get(t_E + \alpha, 1, n)$. Since s was chosen arbitrarily, it follows that $Cont_0(t_E) \subseteq Cont_0(t_E + \alpha)$. ■

Proposition C.3 (5.3) *Let $T = \langle S, t_E, l\rangle$ and $T' = \langle S, t_E + \alpha, l\rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and consistent with $bel(t_E)$. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E + \alpha, \alpha)$
2. $dt(t_E, \alpha) = dt(t_E + \alpha, \alpha)$

Proof. 1. If α is T-valid, then $pl(t_E, \alpha) = pl(t_E + \alpha, \alpha) = n$. Assume that α is not T-valid. Now $Mod_T(know(t_E + \alpha)) = Mod_T(know(t_E)) \cap Mod_T(\alpha)$, i.e. $Mod_T(know(t_E + \alpha)) \subseteq Mod_T(\alpha)$, i.e. $get_\uparrow(t_E + \alpha, n-1) \subseteq Mod_T(\alpha)$. So $pl(t_E + \alpha, \alpha) = n - 1$. But $pl(t_E, \alpha) \leq n - 1$ and hence $pl(t_E, \alpha) \leq pl(t_E + \alpha, \alpha)$.

2. Since α is consistent with $bel(t_E)$ it follows that $Mod_T(bel(t_E)) \cap Mod_T(\alpha) \neq \varnothing$, i.e. $bottom(bel(t_E)) \cap Mod_T(\alpha) \neq \varnothing$. So $dt(t_E, \alpha) = 0$. But $Mod_T(bel(t_E + \alpha)) = Mod_T(bel(t_E)) \cap Mod_T(\alpha)$. So $Mod_T(bel(t_E + \alpha)) \subseteq Mod_T(\alpha)$, i.e. $bottom(bel(t_E + \alpha)) \subseteq Mod_T(\alpha)$, i.e. $bottom(bel(t_E + \alpha)) \cap Mod_T(\alpha) \neq \varnothing$. So $dt(t_E + \alpha, \alpha) = 0$. Hence $dt(t_E, \alpha) = dt(t_E + \alpha, \alpha)$. ∎

Theorem C.1 (5.14) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S)$ $= n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. A templated revision operation $* : T_E \times L \to T_E$ satisfies postulates $(\mathbf{TR}*\mathbf{1})$ to $(\mathbf{TR}*\mathbf{8})$ iff for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$ the following holds:*

1. $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$
2. $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha)$
3. *For every* $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ *iff* $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$

Proof. ($\Longrightarrow$) Suppose that $*$ satisfies postulates $(\mathbf{TR}*\mathbf{1})$ to $(\mathbf{TR}*\mathbf{8})$. It must be shown that conditions 1, 2, and 3 hold.

1. It must be shown that $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$. Choose any $s \in Mod_M(bel(t_E * \alpha))$. By postulate $(\mathbf{TR}*\mathbf{1})$, $s \in Mod_M(\alpha)$. Suppose $s \notin Min_{t_E}(Mod_M(\alpha))$. So there must be some $s' \in Mod_M(\alpha)$ such that $t_E(s') < t_E(s)$. Now either $s' \in Mod_M(bel(t_E))$ or $s' \notin Mod_M(bel(t_E))$. Suppose $s' \in Mod_M(bel(t_E))$. But then $bel(t_E) \wedge \alpha$ is M-satisfiable, and thus by postulate $(\mathbf{TR}*\mathbf{2})$, $Mod_M(bel(t_E * \alpha)) = Mod_M(bel(t_E) \wedge \alpha)$. Since $s \in Mod_M(bel(t_E * \alpha))$, it follows that $s \in Mod_M(bel(t_E))$, i.e. $s \in bottom(t_E)$. But then $t_E(s') < t_E(s)$. Contradiction. So $s \in Min_{t_E}(Mod_M(\alpha))$. Suppose $s' \notin Mod_M(bel(t_E))$. Let β be a finite axiomatisation of $\{s, s'\}$, i.e. $Mod_M(\beta) = \{s, s'\}$. Now by postulate $(\mathbf{TR}*\mathbf{1})$, $Mod_M(bel(t_E * \beta)) \subseteq Mod_M(\beta)$. So let $s' \in Mod_M(bel(t_E * \beta))$

and $s \notin Mod_M(bel(t_E * \beta))$. Since $Mod_M(\beta) \cap Mod_M(\alpha) = Mod_M(\beta)$, it follows that $\alpha \wedge \beta \equiv_M \beta$. Thus, by postulate ($\mathbf{TR} * \mathbf{4}$), $Mod_M(bel(t_E * (\alpha \wedge \beta))) = Mod_M(bel(t_E * \beta))$ and hence, by postulate ($\mathbf{TR} * \mathbf{5}$), $Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E * \beta))$. Since $s \in Mod_M(bel(t_E * \alpha))$ and $s \in Mod_M(\beta)$ it follows that $s \in Mod_M(bel(t_E * \beta))$. Contradiction. So $s \in Min_{t_E}(Mod_M(\alpha))$. Since s was chosen arbitrarily it follows that $Mod_M(bel(t_E * \alpha)) \subseteq Min_{t_E}(Mod_M(\alpha))$.

Conversely, choose any $s \in Min_{t_E}(Mod_M(\alpha))$. So $s \in Mod_M(\alpha)$. Suppose that $s \notin Mod_M(bel(t_E*\alpha))$. By postulate ($\mathbf{TR} * \mathbf{3}$), $Mod_M(bel(t_E*\alpha)) \neq \varnothing$. So there must be some $s' \in Mod_M(bel(t_E*\alpha))$. But then, by postulate ($\mathbf{TR} * \mathbf{1}$), $s' \in Mod_M(\alpha)$. Let β be a finite axiomatisation of $\{s, s'\}$, i.e. $Mod_M(\beta) = \{s, s'\}$. Since $Mod_M(\beta) \cap Mod_M(\alpha) = Mod_M(\beta)$, it follows that $\alpha \wedge \beta \equiv_M \beta$. Thus, by postulate ($\mathbf{TR} * \mathbf{4}$), $Mod_M(bel(t_E * (\alpha \wedge \beta))) = Mod_M(bel(t_E * \beta))$ and hence, by postulate ($\mathbf{TR} * \mathbf{5}$), $Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E * \beta))$. Since $s' \in Mod_M(bel(t_E * \alpha))$ and $s' \in Mod_M(\beta)$ it follows that $bel(t_E * \alpha) \wedge \beta$ is M-satisfiable, and hence, from postulate ($\mathbf{TR} * \mathbf{6}$), that $Mod_M(bel(t_E * \beta)) \subseteq Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta)$. But then $Mod_M(bel(t_E * \beta)) = Mod_M(bel(t_E*\alpha)) \cap Mod_M(\beta)$. Since $s' \in Mod_M(bel(t_E*\alpha))$ and $s \notin Mod_M(bel(t_E*\alpha))$, it follows that $Mod_M(bel(t_E * \beta)) = \{s'\}$. So $t_E * \beta(s') < t_E * \beta(s)$. If $t_E * \beta(s) = n$ then $s \notin Mod_M(know(t_E * \beta))$, and hence from postulate ($\mathbf{TR} * \mathbf{7}$) it follows that either $s \notin Mod_M(know(t_E))$ or $s \notin Mod_M(\beta)$. But $s \in Mod_M(\beta)$ and thus $s \notin Mod_M(know(t_E))$, i.e. $t_E(s) = n$. So $t_E(s') \leq t_E(s)$. On the other hand, if $s \in Mod_M(know(t_E * \beta))$, then by postulate ($\mathbf{TR} * \mathbf{8}$), $t_E(s') \leq t_E(s)$. But $s \in Min_{t_E}(Mod_M(\alpha))$ and thus $t_E(s) \leq t_E(s')$. So $t_E(s) = t_E(s')$. But $s' \in Mod_M(bel(t_E * \alpha))$, i.e. $s' \in bottom(t_E * \alpha)$, i.e. $(t_E * \alpha)(s') = 0$. But then $(t_E * \alpha)(s) = 0$, i.e. $s \in bottom(t_E * \alpha)$, i.e. $s \in Mod_M(bel(t_E * \alpha))$. Contradiction. Since s was chosen arbitrarily it follows that $Min_{t_E}(Mod_M(\alpha)) \subseteq Mod_M(bel(t_E * \alpha))$. Hence $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$.

2. It must be shown that $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E)) \cap Mod_M(\alpha)$,

which follows directly by postulate ($\mathbf{TR} * \mathbf{7}$).

3. It must be shown that for every $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ iff $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$. Suppose there is some $s, s' \in Mod_M(\alpha)$ such that $t_E(s) \leq t_E(s')$ but $(t_E * \alpha)(s) \not\leq (t_E * \alpha)(s')$. Suppose that $t_E(s') = n$, i.e. $s' \notin Mod_M(know(t_E))$. But then, by postulate ($\mathbf{TR} * \mathbf{7}$), $s' \notin Mod_M(know(t_E * \alpha))$, i.e. $(t_E * \alpha)(s') = n$. So $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$. Contradiction. Suppose that $t_E(s') < n$. Let β be a finite axiomatisation of $\{s, s'\}$, i.e. $Mod_M(\beta) = \{s, s'\}$. Since $Mod_M(\beta) \cap Mod_M(\alpha) = Mod_M(\beta)$, it follows that $\alpha \wedge \beta \equiv_M \beta$. Thus, by postulate ($\mathbf{TR} * \mathbf{4}$), $Mod_M(bel(t_E * (\alpha \wedge \beta))) = Mod_M(bel(t_E * \beta))$ and hence, by postulate ($\mathbf{TR} * \mathbf{5}$), $Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E * \beta))$. But $Mod_M(bel(t_E * \beta)) \subseteq Mod_M(\beta)$ by postulate ($\mathbf{TR} * \mathbf{1}$). Suppose that $Mod_M(bel(t_E * \beta)) = \{s\}$. So $s \in Mod_M(bel(t_E * \alpha))$ and $s' \notin Mod_M(bel(t_E * \alpha))$. But then $(t_E * \alpha)(s) < (t_E * \alpha)(s')$. Contradiction. So for every $s, s' \in Mod_M(\alpha)$ such that $t_E(s) \leq t_E(s')$ it holds that $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$.

Conversely, suppose there is some $s, s' \in Mod_M(\alpha)$ such that $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$ but $t_E(s) \not\leq t_E(s')$. Suppose that $(t_E * \alpha)(s') = n$, i.e. $s' \notin Mod_M(know(t_E * \alpha))$. But $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E)) \cap Mod_M(\alpha)$ by postulate ($\mathbf{TR} * \mathbf{7}$). So either $s' \notin Mod_M(know(t_E))$ or $s' \notin Mod_M(\alpha)$. But $s' \in Mod_M(\alpha)$ and thus it must be the case that $s' \notin Mod_M(know(t_E))$, i.e. $t_E(s') = n$. But then $t_E(s) \leq t_E(s')$. Contradiction. Suppose that $(t_E * \alpha)(s') < n$. But then, by postulate ($\mathbf{TR} * \mathbf{8}$), $t_E(s) \leq t_E(s')$. Contradiction. So for every $s, s' \in Mod_M(\alpha)$ such that $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$ it holds that $t_E(s) \leq t_E(s')$.

($\Longleftarrow$) Suppose that conditions 1, 2, and 3 hold. It must be shown that $*$ satisfies postulates ($\mathbf{TR} * \mathbf{1}$) to ($\mathbf{TR} * \mathbf{8}$).

($\mathbf{TR} * \mathbf{1}$) It must be shown that $Mod_M(bel(t_E * \alpha)) \subseteq Mod_M(\alpha)$. But $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$ and since $Min_{t_E}(Mod_M(\alpha)) \subseteq Mod_M(\alpha)$ it follows that

$Mod_M(bel(t_E * \alpha)) \subseteq Mod_M(\alpha)$.

(**TR** $*$ **2**) Suppose that $bel(t_E) \wedge \alpha$ is M-satisfiable, i.e. $bottom(t_E) \cap Mod_M(\alpha) \neq \varnothing$. It must be shown that $Mod_M(bel(t_E * \alpha)) = Mod_M(bel(t_E) \wedge \alpha)$. Choose any $s \in Mod_M(bel(t_E * \alpha))$, i.e. $s \in Min_{t_E}(Mod_M(\alpha))$. So there can be no $s' \in Mod_M(\alpha)$ such that $t(s') \leq t(s)$, unless $t(s) \leq t(s')$. But then $s \in bottom(t_E)$ otherwise $bel(t_E) \wedge \alpha$ is not M-satisfiable. So $s \in Mod_M(bel(t_E))$. But then $s \in Mod_M(bel(t_E) \wedge \alpha)$. Since s was chosen arbitrarily, it follows that $Mod_M(bel(t_E * \alpha)) \subseteq Mod_M(bel(t_E) \wedge \alpha)$. Conversely, choose any $s \in Mod_M(bel(t_E) \wedge \alpha)$. So $s \in bottom(t_E)$ and $s \in Mod_M(\alpha)$. But then there can be no $s' \in Mod_M(\alpha)$ such that $t(s') \leq t(s)$, unless $t(s) \leq t(s')$. Hence $s \in Min_{t_E}(Mod_M(\alpha))$. Since s was chosen arbitrarily, it follows that $Mod_M(bel(t_E) \wedge \alpha) \subseteq Min_{t_E}(Mod_M(\alpha))$. Hence $Mod_M(bel(t_E * \alpha)) = Mod_M(bel(t_E) \wedge \alpha)$.

(**TR** $*$ **3**) It must be shown that $bel(t_E * \alpha)$ and $know(t_E * \alpha)$ are M-satisfiable. Since α is consistent with $know(t_E)$ it follows that $Mod_M(know(t_E) \wedge \alpha) \neq \varnothing$. But $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha)$ and thus $know(t_E * \alpha)$ is M-satisfiable. But if $Mod_M(know(t_E) \wedge \alpha) \neq \varnothing$ then $Mod_M(\alpha) \neq \varnothing$ and thus $Min_{t_E}(Mod_M(\alpha)) \neq \varnothing$. Since $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$ it follows that $bel(t_E * \alpha)$ is M-satisfiable.

(**TR** $*$ **4**) Suppose that $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$. It must be shown that $t_E * \alpha = t_{E'} * \alpha'$. Let $X = Mod_M(\alpha)$ and $S - X = NMod_M(\alpha)$. So for every $s, s' \in X$, $t_E(s) \leq t_E(s')$ iff $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$. Since $t_E(s) = t_{E'}(s)$ for every $s \in S$, it follows that for every $s, s' \in X$, $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$ iff $(t_{E'} * \alpha)(s) \leq (t_{E'} * \alpha)(s')$. Hence $t_E * \alpha$ and $t_{E'} * \alpha$ are order-equivalent on X. Now $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha) = Mod_M(know(t_E)) \cap Mod_M(\alpha)$ and thus $S - X \not\subseteq Mod_M(know(t_E * \alpha))$, i.e. $S - X \not\subseteq get_{\uparrow}(t_E * \alpha, n-1)$. So for every $s, s' \in S - X$, $(t_E * \alpha)(s) = n$. Since $Mod_M(know(t_{E'})) = Mod_M(know(t_E))$, it follows that for every $s, s' \in S - X$, $(t_{E'} * \alpha)(s) = n$. Hence $t_E * \alpha$ and $t_{E'} * \alpha$ are order-equivalent on $S - X$. But $\alpha \equiv_M \alpha'$ and thus $X = Mod_M(\alpha')$ and $S - X = NMod_M(\alpha')$. But then $t_E * \alpha$ and $t_{E'} * \alpha'$ are order-equivalent. Since both

$t_E * \alpha$ and $t_{E'} * \alpha'$ are regular, it follows by proposition 3.17 that $t_E * \alpha = t_{E'} * \alpha'$.

(**TR** $*$ **5**) It must be shown that $Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E * (\alpha \wedge \beta)))$. Choose any $s \in Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta)$. So $s \in Mod_M(bel(t_E * \alpha))$ and $s \in Mod_M(\beta)$, i.e. $s \in Min_{t_E}(Mod_M(\alpha))$ and $s \in Mod_M(\beta)$, i.e. $s \in Mod_M(\alpha)$ and $s \in Mod_M(\beta)$, i.e. $s \in Mod_M(\alpha) \cap Mod_M(\beta)$, i.e. $s \in Mod_M(\alpha \wedge \beta)$. Suppose that $s \notin Mod_M(bel(t_E * (\alpha \wedge \beta)))$, i.e. $s \notin Min_{t_E}(Mod_M(\alpha \wedge \beta))$. But $s \in Mod_M(\alpha \wedge \beta)$ and hence there must be some $s' \in Min_{t_E}(Mod_M(\alpha \wedge \beta))$ and $t_E(s') < t_E(s)$. But if $s' \in Min_{t_E}(Mod_M(\alpha \wedge \beta))$ then $s' \in Mod_M(\alpha \wedge \beta)$, i.e. $s' \in Mod_M(\alpha)$. But then $s \notin Min_{t_E}(Mod_M(\alpha))$. Contradiction. So $s \in Mod_M(bel(t_E * (\alpha \wedge \beta)))$. Since s was chosen arbitrarily it follows that $Mod_M(bel(t_E*\alpha))\cap Mod_M(\beta) \subseteq Mod_M(bel(t_E*(\alpha\wedge\beta)))$.

(**TR** $*$ **6**) Suppose that $bel(t_E * \alpha) \wedge \beta$ is M-satisfiable, i.e. $Min_{t_E}(Mod_M(\alpha)) \cap Mod_M(\beta) \neq \varnothing$. It must be shown that $Mod_M(bel(t_E * (\alpha \wedge \beta))) \subseteq Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta)$. Choose any $s \in Mod_M(bel(t_E*(\alpha\wedge\beta)))$. So $s \in Min_{t_E}(Mod_M(\alpha\wedge\beta))$, i.e. $s \in Mod_M(\alpha)$ and $s \in Mod_M(\beta)$ and there is no $s' \in Mod_M(\alpha)$ and $s' \in Mod_M(\beta)$ such that $t(s') \leq t(s)$, unless $t(s) \leq t(s')$. Suppose that $s \notin Mod_M(bel(t_E*\alpha)) \cap Mod_M(\beta)$, i.e. $s \notin Min_{t_E}(Mod_M(\alpha)) \cap Mod_M(\beta)$, i.e. $s \notin Min_{t_E}(Mod_M(\alpha))$ or $s \notin Mod_M(\beta)$. Since $s \in Mod_M(\beta)$ it follows that $s \notin Min_{t_E}(Mod_M(\alpha))$. However, $Min_{t_E}(Mod_M(\alpha)) \cap Mod_M(\beta) \neq \varnothing$ and hence there must be some $s'' \in Min_{t_E}(Mod_M(\alpha))$ such that $s'' \in Mod_M(\beta)$, i.e. $s'' \in Mod_M(\alpha)$ and $s'' \in Mod_M(\beta)$. Since $s \notin Min_{t_E}(Mod_M(\alpha))$ but $s \in Mod_M(\alpha)$ it follows that $t_E(s'') < t_E(s)$. Contradiction. So $s \in Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta)$. Since s was chosen arbitrarily it follows that $Mod_M(bel(t_E * (\alpha \wedge \beta))) \subseteq Mod_M(bel(t_E * \alpha)) \cap Mod_M(\beta)$.

(**TR** $*$ **7**) It must be shown that $Mod_M(know(t_E*\alpha)) = Mod_M(know(t_E) \wedge \alpha)$, which follows directly from condition 2.

(**TR** $*$ **8**) Suppose that $(t_E * \alpha)(s) \leq (t_E * \alpha)(s') < n$. It must be shown that $t_E(s) \leq$

$t_E(s')$. Since $(t_E * \alpha)(s) \leq (t_E * \alpha)(s') < n$, it follows that $s, s' \in Mod_M(know(t_E * \alpha))$. But $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha)$ and thus $s, s' \in Mod_M(\alpha))$. But for every $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ iff $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$. Hence $t_E(s) \leq t_E(s')$.
■

Proposition C.4 (5.5) *Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. Then the revision operation $*_{\boxminus}$ satisfies postulates $(\mathbf{DP} * \mathbf{1})$ to $(\mathbf{DP} * \mathbf{6})$ under the classical interpretation $M_0 = \langle S, l \rangle$ of L.*

Proof. It follows from proposition 4.12 that t_E induces a total preorder $\preceq_E$ on the set of states that is faithful with respect to $Th(bottom(t))$ (or $Cn(bel(t_E))$) and from proposition 5.4 that if $t_E = t_{E'}$ then $\preceq_E = \preceq_{E'}$. Hence the requirement by theorem 5.10 for a faithful assignment is satisfied. If it can be shown that $Mod(bel(t_\alpha \boxminus t_E)) = Min_{\preceq_E}(\alpha)$ then, by theorem 5.10, it follows that the revision operation $*_{\boxminus}$ satisfies postulates $(\mathbf{DP} * \mathbf{1})$ to $(\mathbf{DP} * \mathbf{6})$. For every sentence $\alpha \in L$, $Mod(\alpha) = bottom(t_\alpha)$ and $NMod(\alpha) = top(t_\alpha)$ for $t_\alpha \in T_D$ the definite t-ordering induced by α. Choose any $s \in Mod(bel(t_\alpha \boxminus t_E))$. So $s \in bottom(t_\alpha \boxminus t_E) \neq \varnothing$, i.e. $(t_\alpha \boxminus t_E)(s) = 0$. But if $(t_\alpha \boxminus t_E)(s) = 0$ then $card(seg(t_\alpha(s), t_E(s))) = 0$ and by proposition 4.15(1), $t_\alpha(s) = 0$. So $s \in bottom(t_\alpha)$, i.e. $s \in Mod(\alpha)$. Suppose there is some $s' \in Mod(\alpha)$ such that $s' \preceq_E s$ but $s \not\preceq_E s'$. But then $card(seg(t_\alpha(s), t_E(s))) \neq 0$. Contradiction. Hence, there can be no $s' \in Mod(\alpha)$ such that $s' \preceq_E s$, unless $s \preceq_E s'$. So $s \in Min_{\preceq_E}(\alpha)$. Since s was chosen arbitrarily is follows that $Mod(bel(t_\alpha \boxminus t_E)) \subseteq Min_{\preceq_E}(\alpha)$.

Conversely, choose any $s \in Min_{\preceq_E}(\alpha)$. So $s \in Mod(\alpha)$ and there is no $s' \in Mod(\alpha)$ such that $s' \preceq_E s$, unless $s \preceq_E s'$. Hence $t_\alpha(s) = 0$ and $t_E(s) \leq t_E(s')$. But then $card(seg(t_\alpha(s), t_E(s))) = 0$ and since $t_\alpha(s) < n$ it follows that $(t_\alpha \boxminus t_E)(s) = 0$. So $s \in bottom(t_\alpha \boxminus t_E)$, i.e. $s \in Mod(bel(t_\alpha \boxminus t_E))$. Since s was chosen arbitrarily is follows that $Min_{\preceq_E}(\alpha) \subseteq Mod(bel(t_\alpha \boxminus t_E))$. But $Mod(bel(t_\alpha \boxminus t_E)) \subseteq Min_{\preceq_E}(\alpha)$ and thus $Mod(bel(t_\alpha \boxminus t_E)) = Min_{\preceq_E}(\alpha)$. ■

Proposition C.5 (5.6) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S)$ $= n$. Let $t_E \in T_E$ and let $\alpha \in L$ be consistent with $know(t_E)$. Then $Cont_D(t_E) \subseteq Cont_D(t_E * \alpha)$.*

Proof. If $Cont_D(t_E) = \varnothing$ then the result holds vacuously. Assume that $Cont_D(t_E) \neq \varnothing$. Choose any $s \in Cont_D(t_E)$. So $s \in top(t_E)$, i.e. $s \notin get_\uparrow(t_E, n-1)$, i.e. $s \notin Mod_M(know(t_E))$, i.e. $s \notin Mod_M(know(t_E)) \cap Mod_M(\alpha)$, i.e. $s \notin Mod_M(know(t_E * \alpha))$, i.e. $s \notin get_\uparrow(t_E * \alpha, n-1)$. So $s \in top(t_E * \alpha)$. Since s was chosen arbitrarily, it follows that $Cont_D(t_E) \subseteq Cont_D(t_E * \alpha)$. ■

Proposition C.6 (5.7) *Let S be a non-empty set of states with $card(S) = n$. Let $t_1 \in T_D$ and $t_2 \in T_N$ and let A be the index set induced by $\langle t_1, t_2 \rangle$ and well-ordered by the lexicographic ordering $\preceq$. Then the following holds for every $s, s' \in S$:*

1. $t_1(s) \leq (t_1 \boxplus t_2)(s)$
2. *if* $t_1(s) < t_1(s')$ *then* $(t_1 \boxplus t_2)(s) \leq (t_1 \boxplus t_2)(s')$
3. *if* $t_1(s) = t_1(s') \neq n$ *and* $t_2(s) < t_2(s')$ *then* $(t_1 \boxplus t_2)(s) < (t_1 \boxplus t_2)(s')$
4. $bottom(t_1 \boxplus t_2) \subseteq bottom(t_1)$
5. $Cont_D(t_1) \subseteq Cont_D(t_1 \boxplus t_2)$
6. $Cont_0(t_1) \subseteq Cont_0(t_1 \boxplus t_2)$
7. $Cont_D(t_2) \subseteq Cont_D(t_1 \boxplus t_2)$

Proof. 1. Suppose that $t_1(s) = n$. But then $(t_1 \boxplus t_2)(s) = n$ and thus $t_1(s) = (t_1 \boxplus t_2)(s)$. Suppose that $t_1(s) \neq n$. Since t_1 is definite it follows that $t_1(s) = 0$. But then $card(seg(t_1(s), t_2(s))) \geq 0$, i.e. $(t_1 \boxplus t_2)(s) \geq 0$. Hence $t_1(s) \leq (t_1 \boxplus t_2)(s)$.

2. Let $t_1(s) < t_1(s')$. Since t_1 is definite it follows that $t_1(s) = 0$. Suppose that $(t_1 \boxplus t_2)(s) > (t_1 \boxplus t_2)(s')$. So $card(seg(t_1(s), t_2(s))) > card(seg(t_1(s'), t_2(s')))$. But then $(t_1(s'), t_2(s')) \prec (t_1(s), t_2(s))$. So $t_1(s') < t_1(s)$. Contradiction. Hence $(t_1 \boxplus t_2)(s) \leq (t_1 \boxplus t_2)(s')$.

3. Let $t_1(s) = t_1(s') \neq n$ and $t_2(s) < t_2(s')$. Since t_1 is definite it follows that $t_1(s) = t_1(s') = 0$. Suppose that $(t_1 \boxplus t_2)(s) > (t_1 \boxplus t_2)(s')$. So $card(seg(t_1(s), t_2(s))) > card(seg(t_1(s'), t_2(s')))$. But then $(t_1(s'), t_2(s')) \prec (t_1(s), t_2(s))$. Since $t_1(s') = t_1(s)$, it follows that $t_2(s') < t_2(s)$. Contradiction. Suppose that $(t_1 \boxplus t_2)(s) = (t_1 \boxplus t_2)(s')$. Since $t_1(s) = t_1(s') = 0$ it follows that $t_2(s) = t_2(s')$. Contradiction. Hence $(t_1 \boxplus t_2)(s) < (t_1 \boxplus t_2)(s')$.

4. It must be shown that $bottom(t_1 \boxplus t_2) \subseteq bottom(t_1)$. Choose any $s \in bottom(t_1 \boxplus t_2)$. So $(t_1 \boxplus t_2)(s) = 0$. But then $card(seg(t_1(s), t_2(s))) = 0$. Since t_1 is definite it follows that $t_1(s) = 0$. So $s \in bottom(t_1)$. Since s was chosen arbitrarily it follows that $bottom(t_1 \boxplus t_2) \subseteq bottom(t_1)$.

5. It must be shown that $top(t_1) \subseteq top(t_1 \boxplus t_2)$. Choose any $s \in top(t_1)$. So $t_1(s) = n$. But then $(t_1 \boxplus t_2)(s) = n$, i.e. $s \in top(t_1 \boxplus t_2)$. Since s was chosen arbitrarily it follows that $top(t_1) \subseteq top(t_1 \boxplus t_2)$.

6. It must be shown that $get(t_1, 1, n) \subseteq get(t_1 \boxplus t_2, 1, n)$. Since t_1 is definite it follows that $get(t_1, 1, n) = top(t_1)$. By the previous condition it follows that $top(t_1) \subseteq top(t_1 \boxplus t_2)$. But $top(t_1 \boxplus t_2) \subseteq get(t_1 \boxplus t_2, 1, n)$. Hence $get(t_1, 1, n) \subseteq get(t_1 \boxplus t_2, 1, n)$.

7. It must be shown that $top(t_2) \subseteq top(t_1 \boxplus t_2)$. Choose any $s \in top(t_2)$. So $t_2(s) = n$. But then $(t_1 \boxplus t_2)(s) = n$, i.e. $s \in top(t_1 \boxplus t_2)$. Since s was chosen arbitrarily it follows that $top(t_2) \subseteq top(t_1 \boxplus t_2)$. ■

Proposition C.7 (5.8) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E. Then the templated revision operation $*_{\boxplus}$ satisfies postulates $(\mathbf{TR} * \mathbf{1})$ to $(\mathbf{TR} * \mathbf{8})$.*

Proof. Theorem 5.14 will be used to show that the templated revision operation $*_{\boxplus}$ satisfies postulates $(\mathbf{TR} * \mathbf{1})$ to $(\mathbf{TR} * \mathbf{8})$. For every sentence $\alpha \in L$, $Mod_M(\alpha) = bottom(t_\alpha)$ and $NMod_M(\alpha) = top(t_\alpha)$ for $t_\alpha \in T_D$ the definite t-ordering induced by α (since M is extensional). It will be shown that for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$, the following holds:

1. $Mod_M(bel(t_\alpha \boxplus t_E)) = Min_{t_E}(Mod_M(\alpha))$
2. $Mod_M(know(t_\alpha \boxplus t_E)) = Mod_M(know(t_E) \wedge \alpha)$
3. For every $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ iff $(t_\alpha \boxplus t_E)(s) \leq (t_\alpha \boxplus t_E)(s')$

1. Choose any $s \in Mod_M(bel(t_\alpha \boxplus t_E))$. So $s \in bottom(t_\alpha \boxplus t_E)$, i.e. $(t_\alpha \boxplus t_E)(s) = 0$. But if $(t_\alpha \boxplus t_E)(s) = 0$ then $card(seg(t_\alpha(s), t_E(s))) = 0$ and by proposition 5.7(1), $t_\alpha(s) = 0$. So $s \in bottom(t_\alpha)$, i.e. $s \in Mod_M(\alpha)$. Suppose there is some $s' \in Mod_M(\alpha)$ such that $t_E(s') \leq t_E(s)$ and $t_E(s) \not\leq t_E(s')$. But then $card(seg(t_\alpha(s), t_E(s))) \neq 0$. Contradiction. Hence, there can be no $s' \in Mod_M(\alpha)$ such that $t_E(s') \leq t_E(s)$, unless $t_E(s) \leq t_E(s')$. So $s \in Min_{t_E}(Mod_M(\alpha))$. Since s was chosen arbitrarily is follows that $Mod_M(bel(t_\alpha \boxplus t_E)) \subseteq Min_{t_E}(Mod_M(\alpha))$.

Conversely, choose any $s \in Min_{t_E}(Mod_M(\alpha))$. So $t_\alpha(s) = 0$ and $t_E(s) < n$ otherwise $know(t_E) \wedge \alpha$ is not M-satisfiable. But $s \in Min_{t_E}(Mod_M(\alpha))$ and hence there can be no $s' \in Mod_M(\alpha)$ such that $t_E(s') \leq t_E(s)$, unless $t_E(s) \leq t_E(s')$. But then $card(seg(t_\alpha(s), t_E(s))) = 0$ and since both $t_\alpha(s) < n$ and $t_E(s) < n$, it follows that $(t_\alpha \boxplus t_E)(s) = 0$. So $s \in bottom(t_\alpha \boxplus t_E)$, i.e. $s \in Mod_M(bel(t_\alpha \boxplus t_E))$. Since s was chosen arbitrarily is follows that $Min_{t_E}(Mod_M(\alpha)) \subseteq Mod_M(bel(t_\alpha \boxplus t_E))$. Hence $Mod_M(bel(t_\alpha \boxplus t_E)) = Min_{t_E}(Mod_M(\alpha))$.

2. Choose any $s \in Mod_M(know(t_\alpha \boxplus t_E))$. So $s \in get_\uparrow(t_\alpha \boxplus t_E, n-1)$, i.e. $(t_\alpha \boxplus t_E)(s) < n$. But then $card(seg(t_\alpha(s), t_E(s))) < n$. Since t_α is definite, it follows that $t_\alpha(s) = 0$ and $t_E(s) < n$, i.e. $s \in bottom(t_\alpha)$ and $s \in get_\uparrow(t_E, n-1)$, i.e. $s \in Mod_M(\alpha)$ and $s \in Mod_M(know(t_E))$, i.e. $s \in Mod_M(know(t_E) \wedge \alpha)$. Since s was chosen arbitrarily is follows that $Mod_M(know(t_\alpha \boxplus t_E)) \subseteq Mod_M(know(t_E) \wedge \alpha)$.

Conversely, chooses any $s \in Mod_M(know(t_E) \wedge \alpha)$, i.e. $s \in Mod_M(know(t_E))$ and $s \in Mod_M(\alpha)$, i.e. $s \in get_\uparrow(t_E, n-1)$ and $s \in bottom(t_\alpha)$, i.e. $t_E(s) < n$ and $t_\alpha(s) = 0$. But then $card(seg(t_\alpha(s), t_E(s))) < n$. So $(t_\alpha \boxplus t_E)(s) < n$, i.e. $s \in get_\uparrow(t_\alpha \boxplus t_E, n-1)$, i.e. $s \in Mod_M(know(t_\alpha \boxplus t_E))$. Since s was chosen arbitrarily is follows that $Mod_M(know(t_E) \wedge \alpha) \subseteq Mod_M(know(t_\alpha \boxplus t_E))$. Hence $Mod_M(know(t_\alpha \boxplus t_E)) = Mod_M(know(t_E) \wedge \alpha)$.

3. Choose any $s, s' \in Mod_M(\alpha)$. So $t_\alpha(s) = 0$ and $t_\alpha(s') = 0$. Suppose $t_E(s) \leq t_E(s')$. If $t_E(s') = n$ then $(t_\alpha \boxplus t_E)(s') = n$ and hence $(t_\alpha \boxplus t_E)(s) \leq (t_\alpha \boxplus t_E)(s')$. Otherwise $card(seg(t_\alpha(s), t_E(s))) \leq card(seg(t_\alpha(s'), t_E(s')))$ and thus $(t_\alpha \boxplus t_E)(s) \leq (t_\alpha \boxplus t_E)(s')$.

Conversely, suppose that $(t_\alpha \boxplus t_E)(s) \leq (t_\alpha \boxplus t_E)(s')$. If $(t_\alpha \boxplus t_E)(s') = n$ then either $t_E(s) = n$ or $t_E(s') = n$ and thus $t_E(s) \leq t_E(s')$. Otherwise $card(seg(t_\alpha(s), t_E(s))) \leq card(seg(t_\alpha(s'), t_E(s')))$. But $t_\alpha(s) = 0$ and $t_\alpha(s') = 0$ and hence $t_E(s) \leq t_E(s')$. ■

Proposition C.8 (5.9) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Every templated revision operation $*$ satisfies postulates* (**C1**)*,* (**C3**)*, and* (**C4**)*, provided $*$ is defined for the input sentence.*

Proof. Let $*$ be any templated revision operation, i.e. $*$ satisfies postulates (**TR $*$ 1**) to (**TR $*$ 8**). Then, by theorem 5.14, it follows that for every M-satisfiable sentence $\alpha \in L$ that is consistent with $know(t_E)$, it holds that

1. $Mod_M(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$,
2. $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E) \wedge \alpha)$, and
3. for every $s, s' \in Mod_M(\alpha)$, $t_E(s) \leq t_E(s')$ iff $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$.

(**C1**) Suppose that that $Mod_M(\beta) \subseteq Mod_M(\alpha)$. It must be shown that $Mod_M(bel(t_E* \beta)) = Mod_M(bel(t_E * \alpha) * \beta))$. Suppose that $know(t_E) \wedge \beta$ is M-satisfiable so that $t_E * \beta$ is defined. So $Mod_M(bel(t_E * \beta)) = Min_{t_E}(Mod_M(\beta))$ and for every $s, s' \in Mod_M(\beta)$, $t_E(s) \leq t_E(s')$ iff $(t_E * \beta)(s) \leq (t_E * \beta)(s')$. But $Mod_M(\beta) \subseteq Mod_M(\alpha)$ and thus, since $t_E(s) \leq t_E(s')$ iff $(t_E * \alpha)(s) \leq (t_E * \alpha)(s')$ for every $s, s' \in Mod_M(\alpha)$, it follows that t_E and $t_E * \alpha$ and order equivalent on $Mod_M(\beta)$. Hence $Mod_M(bel(t_E * \beta)) = Min_{t_E*\alpha}(Mod_M(\beta))$. But $Mod_M(know(t_E * \alpha)) = Mod_M(know(t_E)) \cap Mod_M(\alpha)$. Since $Mod_M(\beta) \subseteq Mod_M(\alpha)$ it follows that $Mod_M(know(t_E * \alpha)) \cap Mod_M(\beta) \neq \varnothing$. So $know(t_E*\alpha) \wedge \beta$ is M-satisfiable. But then $Mod_M(bel(t_E*\alpha)*\beta) = Min_{t_E*\alpha}(Mod_M(\beta))$. Hence $Mod_M(bel(t_E * \beta)) = Mod_M(bel(t_E * \alpha) * \beta)$.

(**C3**) Suppose that $Mod_M(bel(t_E * \beta)) \subseteq Mod_M(\alpha)$. It must be shown that Mod_M $(bel(t_E * \alpha) * \beta) \subseteq Mod_M(\alpha)$. Suppose that $know(t_E * \alpha) \wedge \beta$ is M-satisfiable so that $(t_E*\alpha)*\beta$ is defined. So $Mod_M(bel(t_E*\alpha)*\beta) = Min_{t_E*\alpha}(Mod_M(\beta))$ and for every $s, s' \in Mod_M(\beta)$, $t_E(s) \leq t_E(s')$ iff $((t_E*\alpha)*\beta)(s) \leq ((t_E*\alpha)*\beta)(s')$. Suppose that $know(t_E) \wedge \beta$ is M-satisfiable so that $t_E * \beta$ is defined. So $Mod_M(bel(t_E * \beta)) = Min_{t_E}(Mod_M(\beta))$ and for every $s, s' \in Mod_M(\beta)$, $t_E(s) \leq t_E(s')$ iff $(t_E*\beta)(s) \leq (t_E*\beta)(s')$. But then $(t_E*\alpha)*\beta$ and $t_E * \beta$ are order equivalent on $Mod_M(\beta)$. Suppose that $s \in Mod_M(bel(t_E * \alpha) * \beta)$, i.e. $s \in bottom((t_E * \alpha) * \beta)$, i.e. $((t_E * \alpha) * \beta)(s) = 0$. So there can be no $s' \in Mod_M(\beta)$ such that $((t_E * \alpha) * \beta)(s) < ((t_E * \alpha) * \beta)(s')$, i.e. there can be no $s' \in Mod_M(\beta)$ such that $(t_E * \beta)(s) < (t_E * \beta)(s')$. So $s \in Min_{t_E}(Mod_M(\beta))$. But since $Mod_M(bel(t_E * \beta)) = Min_{t_E}(Mod_M(\beta))$, it follows that $s \in bottom(t_E * \beta)$, i.e. $s \in Mod_M(bel(t_E * \beta))$. But $Mod_M(bel(t_E * \beta)) \subseteq Mod_M(\alpha)$ and thus $s \in Mod_M(\alpha)$. Since s was chosen arbitrarily it follows that $Mod_M(bel(t_E * \alpha) * \beta) \subseteq Mod_M(\alpha)$.

(**C4**) Suppose that that $Mod_M(bel(t_E * \beta)) \not\subseteq NMod_M(\alpha)$. It must be shown that $Mod_M(bel(t_E * \alpha) * \beta) \not\subseteq NMod_M(\alpha)$. Suppose that $know(t_E * \alpha) \wedge \beta$ is M-satisfiable so that $(t_E * \alpha) * \beta$ is defined. So $Mod_M(bel(t_E * \alpha) * \beta) = Min_{t_E*\alpha}(Mod_M(\beta))$ and for every $s, s' \in Mod_M(\beta)$, $t_E(s) \leq t_E(s')$ iff $((t_E * \alpha) * \beta)(s) \leq ((t_E * \alpha) * \beta)(s')$. Suppose that $know(t_E) \wedge \beta$ is M-satisfiable so that $t_E * \beta$ is defined. So $Mod_M(bel(t_E * \beta)) = Min_{t_E}(Mod_M(\beta))$ and for every $s, s' \in Mod_M(\beta)$, $t_E(s) \leq t_E(s')$ iff $(t_E * \beta)(s) \leq (t_E * \beta)(s')$. But then $(t_E * \alpha) * \beta$ and $t_E * \beta$ are order equivalent on $Mod_M(\beta)$. Suppose that $s \in Mod_M(bel(t_E * \alpha) * \beta)$, i.e. $s \in bottom(t_E * \alpha) * \beta)$, i.e. $((t_E * \alpha) * \beta)(s) = 0$. So there can be no $s' \in Mod_M(\beta)$ such that $((t_E * \alpha) * \beta)(s) < ((t_E * \alpha) * \beta)(s')$, i.e. there can be no $s' \in Mod_M(\beta)$ such that $(t_E * \beta)(s) < (t_E * \beta)(s')$. So $s \in Min_{t_E}(Mod_M(\beta))$. But since $Mod_M(bel(t_E * \beta)) = Min_{t_E}(Mod_M(\beta))$, it follows that $s \in bottom(t_E * \beta)$, i.e. $s \in Mod_M(bel(t_E * \beta))$. But $Mod_M(bel(t_E * \beta)) \not\subseteq NMod_M(\alpha)$. So either $s \notin NMod_M(\alpha)$, or $s \in NMod_M(\alpha)$ in which case there must be some $s' \in Mod_M(bel(t_E * \beta))$ such that $s' \notin NMod_M(\alpha)$. But then $s' \in Min_{t_E}(Mod_M(\beta))$. So $t_E(s') = t_E(s)$. But then $(t_E * \beta)(s') = (t_E * \beta)(s)$ and hence, since $(t_E * \alpha) * \beta$ and $t_E * \beta$ are order equivalent on $Mod_M(\beta)$, it follows that $((t_E * \alpha) * \beta)(s') = ((t_E * \alpha) * \beta)(s)$. Since $((t_E * \alpha) * \beta)(s) = 0$, it follows that $((t_E * \alpha) * \beta)(s') = 0$, i.e. $s' \in Mod_M(bel(t_E * \alpha) * \beta)$. Since s was chosen arbitrarily it follows that $Mod_M(bel(t_E * \alpha) * \beta) \not\subseteq NMod_M(\alpha)$. ■

Proposition C.9 (5.10) *Let $T = \langle S, t_E, l \rangle$ and $T' = \langle S, t_E * \alpha, l \rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and consistent with $know(t_E)$. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E * \alpha, \alpha)$
2. $dt(t_E * \alpha, \alpha) \leq dt(t_E, \alpha)$

Proof. 1. If α is T-valid, then $pl(t_E, \alpha) = pl(t_E * \alpha, \alpha) = n$. Assume that α is not T-valid. Now $Mod_T(know(t_E * \alpha)) = Mod_T(know(t_E)) \cap Mod_T(\alpha)$, i.e. $Mod_T(know(t_E * \alpha)) \subseteq Mod_T(\alpha)$, i.e. $get_{\uparrow}(t_E * \alpha, n-1) \subseteq Mod_T(\alpha)$. So $pl(t_E * \alpha, \alpha) = n - 1$. But $pl(t_E, \alpha) \leq n - 1$ and thus $pl(t_E, \alpha) \leq pl(t_E * \alpha, \alpha)$.

2. Since $know(t_E) \wedge \alpha$ is T-satisfiable it follows that $get_{\uparrow}(t_E, n-1) \cap Mod_T(\alpha) \neq \varnothing$. But then $dt(t_E, \alpha) < n$. But $Mod_T(bel(t_E * \alpha)) = Min_{t_E}(Mod_M(\alpha))$. Hence $bottom(t_E * \alpha) \cap Mod_T(\alpha) \neq \varnothing$. But then $dt(t_E * \alpha, \alpha) = 0$ and thus $dt(t_E * \alpha, \alpha) \leq dt(t_E, \alpha)$. ■

C.2 Proofs for Section 5.5

Proposition C.10 (5.11) *Let S be a non-empty set of states with $card(S) = n$ and let $t_1, t_2, t_3 \in T_N$. Then the following holds:*

1. $t_1 \bigtriangledown t_2 = t_2 \bigtriangledown t_1$
2. $t_1 \bigtriangledown (t_2 \bigtriangledown t_3) = (t_1 \bigtriangledown t_2) \bigtriangledown t_3$
3. $bottom(t_1 \bigtriangledown t_2) = bottom(t_1) \cup bottom(t_2)$
4. $Cont_D(t_1 \bigtriangledown t_2) = Cont_D(t_1) \cap Cont_D(t_2)$
5. $Cont_I(t_1 \bigtriangledown t_2) = Cont_I(t_1) \cap Cont_I(t_2)$

Proof. 1. From the definition it follows that $(t_1 \vee t_2)(s) = \min\{t_1(s), t_2(s)\} = \min\{t_2(s), t_1(s)\} = (t_2 \vee t_1)(s)$ for every $s \in S$. So $t_1 \vee t_2 = t_2 \vee t_1$. But then $g \circ (t_1 \vee t_2) = g \circ (t_2 \vee t_1)$, i.e. $t_1 \bigtriangledown t_2 = t_2 \bigtriangledown t_1$.

2. From the definition it follows that $(t_1 \vee (t_2 \vee t_3))(s) = \min\{t_1(s), \min\{t_2(s), t_3(s)\}\}$ and $((t_1 \vee t_2) \vee t_3)(s) = \min\{\min\{t_1(s), t_2(s)\}, t_3(s)\}$ for every $s \in S$. But $\min\{t_1(s), \min\{t_2(s), t_3(s)\}\} = \min\{\min\{t_1(s), t_2(s)\}, t_3(s)\}$ for every $s \in S$. Hence $t_1 \vee (t_2 \vee t_3) = (t_1 \vee t_2) \vee t_3$. But the normalise function $g \circ t$ is order equivalent to t for every $t \in T_S$ by proposition 3.16. Hence, $t_1 \bigtriangledown (t_2 \bigtriangledown t_3) = (t_1 \bigtriangledown t_2) \bigtriangledown t_3$.

3. If $bottom(t_1 \bigtriangledown t_2) = \varnothing$ then $t_1 \bigtriangledown t_2$ must be strongly contradictory, i.e. $g \circ (t_1 \vee t_2)$ is strongly contradictory. So $g((t_1 \vee t_2)(s)) = n$ for every $s \in S$. But then $(t_1 \vee t_2)(s) = n$ for every $s \in S$, i.e. $\min\{t_1(s), t_2(s)\} = n$ for every $s \in S$. So it must be the case that both

t_1 and t_2 are strongly contradictory. Hence $bottom(t_1) \cup bottom(t_2) = \varnothing$. Suppose that $bottom(t_1 \triangledown t_2) \neq \varnothing$. Choose any $s \in bottom(t_1 \triangledown t_2)$. So $(t_1 \triangledown t_2)(s) = 0$. If $(t_1 \vee t_2)(s) \neq 0$, then it must be the case that $t_1(s) \neq 0$ and $t_2(s) \neq 0$. But $t_1, t_2 \in T_N$ and thus, both t_1 and t_2 must be strongly contradictory. But then, by a similar argument as before, $(t_1 \triangledown t_2)(s) \neq 0$. Contradiction. So $(t_1 \vee t_2)(s) = 0$. But then either $t_1(s) = 0$ or $t_2(s) = 0$, i.e. $s \in bottom(t_1)$ or $s \in bottom(t_2)$, i.e. $s \in bottom(t_1) \cup bottom(t_2)$. Since s was chosen arbitrarily, it follows that $bottom(t_1 \triangledown t_2) \subseteq bottom(t_1) \cup bottom(t_2)$. Conversely, if $bottom(t_1) \cup bottom(t_2) = \varnothing$, then both t_1 and t_2 must be strongly contradictory since $t_1, t_2 \in T_N$. But then $t_1 \vee t_2$, and hence $t_1 \triangledown t_2$ would be strongly contradictory so that $bottom(t_1 \triangledown t_2) = \varnothing$. Suppose that $bottom(t_1) \cup bottom(t_2) \neq \varnothing$. Choose any $s \in bottom(t_1) \cup bottom(t_2)$. So $s \in bottom(t_1)$ or $s \in bottom(t_2)$, i.e. $t_1(s) = 0$ or $t_2(s) = 0$. But then $(t_1 \vee t_2)(s) = 0$ and thus $g((t_1 \vee t_2)(s)) = 0$, i.e. $(t_1 \triangledown t_2)(s) = 0$. So $s \in bottom(t_1 \triangledown t_2)$. Since s was chosen arbitrarily, it follows that $bottom(t_1) \cup bottom(t_2) \subseteq bottom(t_1 \triangledown t_2)$. Hence $bottom(t_1 \triangledown t_2) = bottom(t_1) \cup bottom(t_2)$.

4. It must be shown that $top(t_1 \triangledown t_2) = top(t_1) \cap top(t_2)$. If $top(t_1 \triangledown t_2) = \varnothing$ then $g((t_1 \vee t_2)(s)) < n$ for every $s \in S$. But then $(t_1 \vee t_2)(s) < n$ for every $s \in S$, i.e. $\min\{t_1(s), t_2(s)\} < n$ for every $s \in S$. So it must be the case that $top(t_1) \cap top(t_2) = \varnothing$. Suppose that $top(t_1 \triangledown t_2) \neq \varnothing$, i.e. $top(t_1 \vee t_2) \neq \varnothing$. But then $t_1(s) = n$ and $t_2(s) = n$, i.e. $s \in top(t_1)$ and $s \in top(t_2)$, i.e. $s \in top(t_1) \cap top(t_2)$. Since s was chosen arbitrarily, it follows that $top(t_1 \triangledown t_2) \subseteq top(t_1) \cap top(t_2)$. Conversely, if $top(t_1) \cap top(t_2) = \varnothing$, then it must be the case, using a similar argument as before, that $top(t_1 \vee t_2) = \varnothing$, i.e. $top(t_1 \triangledown t_2) = \varnothing$. Suppose that $top(t_1) \cap top(t_2) \neq \varnothing$. Choose any $s \in top(t_1) \cap top(t_2)$. So $s \in top(t_1)$ and $s \in top(t_2)$, i.e. $t_1(s) = n$ and $t_2(s) = n$. But then $\min\{t_1(s), t_2(s)\} = n$, i.e. $(t_1 \vee t_2)(s) = n$, i.e. $g((t_1 \vee t_2)(s)) = n$, i.e. $(t_1 \triangledown t_2)(s) = n$. So $s \in top(t_1 \triangledown t_2)$. Since s was chosen arbitrarily, it follows that $top(t_1) \cap top(t_2) \subseteq top(t_1 \triangledown t_2)$. Hence $top(t_1 \triangledown t_2) = top(t_1) \cap top(t_2)$.

5. It must be shown that $get(t_1 \bigtriangledown t_2, 1, n) = get(t_1, 1, n) \cap get(t_2, 1, n)$. But $get(t_1 \bigtriangledown t_2, 1, n) = S - bottom(t_1 \bigtriangledown t_2)$. But $bottom(t_1 \bigtriangledown t_2) = bottom(t_1) \cup bottom(t_2)$ from result 3. So $get(t_1 \bigtriangledown t_2, 1, n) = S - (bottom(t_1) \cup bottom(t_2))$. But $bottom(t_1) = S - get(t_1, 1, n)$ and $bottom(t_2) = S - get(t_2, 1, n)$. But then $get(t_1 \bigtriangledown t_2, 1, n) = get(t_1, 1, n) \cap get(t_2, 1, n)$.
■

Proposition C.11 (5.12) *Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. Then the templated update operation $\diamond_{\boxminus}$ satisfies postulates $(\mathbf{KM} \diamond \mathbf{1'})$ to $(\mathbf{KM} \diamond \mathbf{7'})$ for every satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, under the classical interpretation $M_0 = \langle S, l \rangle$ of L.*

Proof. $(\mathbf{KM} \diamond \mathbf{1'})$ It must be shown that $Mod(bel(t_E \diamond_{\boxminus} \alpha)) \subseteq Mod(\alpha)$, i.e. that $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \subseteq bottom(t_\alpha)$. By proposition 5.11(3), $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) = \cup_{s \in Mod(bel(t_E))} bottom(t_\alpha \boxplus \vec{F}(s))$. Since $t_\alpha \in T_D$ and $\vec{F}(s) \in T_E$ it follows by proposition 4.15(4), that $bottom(t_\alpha \boxminus \vec{F}(s)) \subseteq bottom(t_\alpha)$ for every $s \in Mod(bel(t_E))$. But then $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \subseteq bottom(t_\alpha)$.

$(\mathbf{KM} \diamond \mathbf{2'})$ Since α is not entailed by $bel(t_E)$, the result holds vacuously.

$(\mathbf{KM} \diamond \mathbf{3'})$ Since both $bel(t_E)$ and α are satisfiable, it must be shown that $Mod(bel(t_E \diamond_{\boxminus} \alpha)) \neq \varnothing$, i.e. that $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \neq \varnothing$. By proposition 4.15(4), $bottom(t_\alpha \boxminus \vec{F}(s)) \subseteq bottom(t_\alpha)$ for every $s \in Mod(bel(t_E))$. Since α is satisfiable, it follows that $bottom(t_\alpha \boxminus \vec{F}(s)) \neq \varnothing$ for every $s \in Mod(bel(t_E))$. But then, by proposition 5.11(3), $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \neq \varnothing$.

$(\mathbf{KM} \diamond \mathbf{4'})$ It must be shown that if $t_E = t_{E'}$ and $\alpha \equiv \alpha'$, then $Mod(bel(t_E \diamond_{\boxminus} \alpha)) = Mod(bel(t_{E'} \diamond_{\boxminus} \alpha'))$. Suppose that $t_E = t_{E'}$ and $\alpha \equiv \alpha'$. It must be shown that $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) = bottom(\bigtriangledown_{s' \in Mod(bel(t_{E'}))}(t_{\alpha'} \boxminus \vec{F}(s')))$. But $t_E = t_{E'}$ and thus $bottom(t_E) = bottom(t_{E'})$, i.e. $Mod(bel(t_E)) = Mod(bel(t_{E'}))$. But then

$\vec{F}(s) = \vec{F}(s')$ for every $s \in Mod(bel(t_E))$. Since $\alpha \equiv \alpha'$ it follows that $t_\alpha = t_{\alpha'}$. But then $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) = bottom(\bigtriangledown_{s' \in Mod(bel(t_{E'}))}(t_{\alpha'} \boxminus \vec{F}(s')))$.

(**KM** $\diamond$ **5′**) It must be shown that $Mod(bel(t_E \diamond_\boxminus \alpha)) \cap Mod(\beta) \subseteq Mod(bel(t_E \diamond_\boxminus (\alpha \wedge \beta)))$, i.e. that $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \cap bottom(t_\beta) \subseteq bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_{\alpha\wedge\beta} \boxminus \vec{F}(s)))$. Choose any $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \cap bottom(t_\beta)$. By proposition 5.11(3), $s' \in bottom(t_\alpha \boxminus \vec{F}(s))$ for some $s \in Mod(bel(t_E))$, say s'', and by proposition 4.15(4), $s' \in bottom(t_\alpha)$. But $s' \in bottom(t_\beta)$ and thus $s' \in bottom(t_{\alpha\wedge\beta})$. But since $bottom(t_{\alpha\wedge\beta}) \subseteq bottom(t_\alpha)$ and $s' \in bottom(t_\beta)$ it follows that $s' \in bottom(t_\alpha \boxminus \vec{F}(s''))$. But then $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_{\alpha\wedge\beta} \boxminus \vec{F}(s)))$. Since s' was chosen arbitrarily, it follows that $Mod(bel(t_E \diamond \alpha)) \cap Mod(\beta) \subseteq Mod(bel(t_E \diamond (\alpha \wedge \beta)))$.

(**KM** $\diamond$ **6′**) It must be shown that if $bel(t_E \diamond \alpha) \wedge \beta$ is satisfiable and $bel(t_E)$ is complete, then $Mod(bel(t_E \diamond (\alpha \wedge \beta))) \subseteq Mod(bel(t_E \diamond \alpha)) \cap Mod(\beta)$. Suppose that $bel(t_E \diamond \alpha) \wedge \beta$ is satisfiable, i.e. $Mod(bel(t_E \diamond \alpha)) \cap Mod(\beta) \neq \varnothing$, and that $bel(t_E)$ is complete, i.e. $bel(t_E)$ has exactly one model, say s. Choose any $s' \in Mod(bel(t_E \diamond (\alpha \wedge \beta)))$. So $s' \in bottom(t_{\alpha\wedge\beta} \boxminus \vec{F}(s))$. By proposition 4.15(4), $s' \in bottom(t_{\alpha\wedge\beta})$. So $s' \in bottom(t_\alpha)$ and $s' \in bottom(t_\beta)$. Suppose that $s' \notin Mod(bel(t_E \diamond \alpha)) \cap Mod(\beta)$. Since $s' \in Mod(\beta)$ it follows that $s' \notin Mod(bel(t_E \diamond_\boxminus \alpha))$, i.e. $s' \notin bottom(t_\alpha \boxminus \vec{F}(s))$. But $Mod(bel(t_E \diamond_\boxminus \alpha)) \cap Mod(\beta) \neq \varnothing$ and hence there must be some $s'' \in bottom(t_\alpha \boxminus \vec{F}(s))$ such that $s'' \in bottom(t_\beta)$. By proposition 4.15(4), $s'' \in bottom(t_\alpha)$. So $(t_\alpha \boxminus \vec{F}(s))(s'') < (t_\alpha \boxminus \vec{F}(s))(s')$. Since $s'' \in bottom(t_{\alpha\wedge\beta})$ and $bottom(t_{\alpha\wedge\beta}) \subseteq bottom(t_\alpha)$, it follows that $(t_{\alpha\wedge\beta} \boxminus \vec{F}(s))(s'') < (t_{\alpha\wedge\beta} \boxminus \vec{F}(s))(s')$. But then $s' \in bottom(t_{\alpha\wedge\beta} \boxminus \vec{F}(s))$. Contradiction. So $s' \in Mod(bel(t_E \diamond_\boxminus \alpha)) \cap Mod(\beta)$. Since s' was chosen arbitrarily, it follows that $Mod(bel(t_E \diamond_\boxminus (\alpha \wedge \beta))) \subseteq Mod(bel(t_E \diamond_\boxminus \alpha)) \cap Mod(\beta)$.

(**KM** $\diamond$ **7′**) It must be shown that $Mod(bel((t_E \bigtriangledown t_{E'}) \diamond_\boxminus \alpha) = Mod(bel(t_E \diamond_\boxminus \alpha)) \cup Mod(bel(t_{E'} \diamond_\boxminus \alpha))$, i.e. it must be shown that $bottom(\bigtriangledown_{s \in Mod(bel(t_E \bigtriangledown t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$ is equal to $bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) \cup bottom(\bigtriangledown_{s \in Mod(bel(t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$. Choose any $s' \in Mod(bel((t_E \bigtriangledown t_{E'}) \diamond_{\boxminus} \alpha))$. So $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E \bigtriangledown t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$. By proposition 5.11(3), $s' \in bottom(t_\alpha \boxminus \vec{F}(s))$ for some $s \in Mod(bel(t_E \bigtriangledown t_{E'}))$, say s''. By proposition 5.11(3) $bottom(t_E \bigtriangledown t_{E'}) = bottom(t_E) \cup bottom(t_{E'})$ and thus $Mod(bel(t_E \bigtriangledown t_{E'})) = Mod(bel(t_E)) \cup Mod(bel(t_{E'}))$. So $s'' \in Mod(bel(t_E))$ or $s'' \in Mod(bel(t_E))$. If $s'' \in Mod(bel(t_E))$, then $s' \in bottom(t_\alpha \boxminus \vec{F}(s''))$ and thus, by proposition 5.11(3), $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))$. Similarly, if $s'' \in Mod(bel(t_{E'}))$, then $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$. So $s' \in Mod(bel(t_E \diamond_{\boxminus} \alpha))$ or $s' \in Mod(bel(t_{E'} \diamond_{\boxminus} \alpha))$. Since s' was chosen arbitrarily, it follows that $Mod(bel((t_E \bigtriangledown t_{E'}) \diamond_{\boxminus} \alpha)) \subseteq Mod(bel(t_E \diamond_{\boxminus} \alpha)) \cup Mod(bel(t_{E'} \diamond_{\boxminus} \alpha))$.

Conversely, choose any $s' \in Mod(bel(t_E \diamond_{\boxminus} \alpha)) \cup Mod(bel(t_{E'} \diamond_{\boxminus} \alpha))$. If $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))$ then by proposition 5.11(3), $s' \in bottom(t_\alpha \boxminus \vec{F}(s))$ for some $s \in Mod(bel(t_E))$, say s''. Similarly, if $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$, then $s' \in bottom(t_\alpha \boxminus \vec{F}(s))$ for some $s \in Mod(bel(t_{E'}))$, say s'''. But $Mod(bel(t_E)) \cup Mod(bel(t_{E'})) = Mod(bel(t_E \bigtriangledown t_{E'}))$ by proposition 5.11(3). So $s' \in bottom(t_\alpha \boxminus \vec{F}(s''))$ for $s'' \in Mod(bel(t_E \bigtriangledown t_{E'}))$ or $s' \in bottom(t_\alpha \boxminus \vec{F}(s'''))$ for $s''' \in Mod(bel(t_E \bigtriangledown t_{E'}))$. But then, by proposition 5.11(3), $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E \bigtriangledown t_{E'}))}(t_\alpha \boxminus \vec{F}(s)))$. Since s' was chosen arbitrarily, it follows that $Mod(bel(t_E \diamond_{\boxminus} \alpha)) \cup Mod(bel(t_{E'} \diamond_{\boxminus} \alpha)) \subseteq Mod(bel((t_E \bigtriangledown t_{E'}) \diamond_{\boxminus} \alpha))$. But then $Mod(bel((t_E \bigtriangledown t_{E'}) \diamond_{\boxminus} \alpha)) = Mod(bel(t_E \diamond_{\boxminus} \alpha)) \cup Mod(bel(t_{E'} \diamond_{\boxminus} \alpha))$. ■

Theorem C.2 (5.15) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and $know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. A templated update operation $\diamond : T_E \times L \to T_E$ satisfies postulates $(\mathbf{TU} \diamond \mathbf{1})$ to $(\mathbf{TU} \diamond \mathbf{8})$ iff for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, the following holds:*

1. $Mod_M(bel(t_E \diamond \alpha)) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$

2. $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$

3. For every $s_1, s_2 \in Mod_M(\alpha)$, if $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$

Proof. ($\Longrightarrow$) Suppose that $\diamond$ satisfies postulates ($\mathbf{TU} \diamond \mathbf{1}$) to ($\mathbf{TU} \diamond \mathbf{7}$). It must be shown that conditions 1, 2, and 3 hold.

1. It must be shown that $Mod_M(bel(t_E \diamond \alpha)) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$. By postulate ($\mathbf{TU} \diamond \mathbf{2}$), $Mod_M(bel(t_E \diamond \alpha) \neq \varnothing$. Choose any $s' \in Mod_M(bel(t_E \diamond \alpha))$. Then by postulate ($\mathbf{TU} \diamond \mathbf{1}$), $s' \in Mod_M(\alpha)$. Suppose that $s' \notin Min_{\vec{F}(s)}(Mod_M(\alpha))$ for every $s \in Mod_M(bel(t_E))$. So for every $s \in Mod_M(bel(t_E))$ there must be some $s'' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ such that $\vec{F}(s)(s'') < \vec{F}(s)(s')$. Let β be a finite axiomatisation of $\{s'', s'\}$, i.e. $Mod_M(\beta) = \{s'', s'\}$. Now by postulate ($\mathbf{TU} \diamond \mathbf{1}$), $Mod_M(bel(t_E \diamond \beta)) \subseteq Mod_M(\beta)$. So let $s'' \in Mod_M(bel(t_E \diamond \beta))$ and $s' \notin Mod_M(bel(t_E \diamond \beta))$. Since $Mod_M(\beta) \cap Mod_M(\alpha) = Mod_M(\beta)$, it follows that $\alpha \wedge \beta \equiv_M \beta$. Thus, by postulate ($\mathbf{TU} \diamond \mathbf{3}$), $Mod_M(bel(t_E \diamond (\alpha \wedge \beta))) = Mod_M(bel(t_E \diamond \beta))$ and hence, by postulate ($\mathbf{TU} \diamond \mathbf{4}$), $Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E \diamond \beta))$. Since $s' \in Mod_M(bel(t_E \diamond \alpha))$ and $s' \in Mod_M(\beta)$ it follows that $s' \in Mod_M(bel(t_E \diamond \beta))$. Contradiction. So $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ for some $s \in Mod_M(bel(t_E))$. But then $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$. Since s' was chosen arbitrarily it follows that $Mod_M(bel(t_E \diamond \alpha)) \subseteq \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$.

Conversely, choose any $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$. So $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ for some $s \in Mod_M(bel(t_E))$, say s_1. So there is can be no $s'' \in Mod_M(\alpha)$ such that $\vec{F}(s_1)(s'') \leq \vec{F}(s_1)(s')$ unless $\vec{F}(s_1)(s') \leq \vec{F}(s_1)(s'')$. Suppose that $s' \notin Mod_M(bel(t_E \diamond \alpha))$. Let $Mod_M(bel(t_E)) = \{s_1, s_2, \ldots, s_k\}$ and let t_i be the definite t-ordering such that $Mod_M(bel((t_i)) = \{s_i\}$. Now $Mod_M(bel(t_E)) = Mod_M(bel(t_1)) \cup Mod_M(bel(t_2)) \cup \ldots \cup Mod_M(bel(t_k))$. But $Mod_M(bel(t)) = bottom(t)$ and thus, by proposition 5.11(3), $Mod_M(bel(t_E)) = Mod_M(bel(t_1 \bigtriangledown t_2 \bigtriangledown \ldots \bigtriangledown t_k))$. By postulate ($\mathbf{TU} \diamond \mathbf{6}$),

$Mod_M(bel(t_E \diamond \alpha)) = Mod_M(bel(t_1 \diamond \alpha)) \cup Mod_M(bel(t_2 \diamond \alpha)) \cup \ldots \cup Mod_M(bel(t_k \diamond \alpha))$. Since $s' \notin Mod_M(bel(t_E \diamond \alpha))$ it follows that $s' \notin Mod_M(bel(t_i \diamond \alpha))$ for every t_i. But, by postulate ($\mathbf{TU} \diamond \mathbf{2}$), $Mod_M(bel(t_E \diamond \alpha)) \neq \varnothing$. So there must be some $s'' \in Mod_M(bel(t_E \diamond \alpha))$, i.e. $s'' \in Mod_M(bel(t_i \diamond \alpha))$ for some t_i, say t_1. But then, by postulate ($\mathbf{TU} \diamond \mathbf{1}$), $s'' \in Mod_M(\alpha)$. Let β be a finite axiomatisation of $\{s'', s'\}$, i.e. $Mod_M(\beta) = \{s'', s'\}$. Since $Mod_M(\beta) \cap Mod_M(\alpha) = Mod_M(\beta)$, it follows that $\alpha \wedge \beta \equiv_M \beta$. Thus, by postulate ($\mathbf{TU} \diamond \mathbf{3}$), $Mod_M(bel(t_1 \diamond (\alpha \wedge \beta))) = Mod_M(bel(t_1 \diamond \beta))$ and hence, by postulate ($\mathbf{TU} \diamond \mathbf{4}$), $Mod_M(bel(t_1 \diamond \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_1 \diamond \beta))$. Since $s'' \in Mod_M(bel(t_1 \diamond \alpha))$ and $s'' \in Mod_M(\beta)$ it follows that $bel(t_1 \diamond \alpha) \wedge \beta$ is M-satisfiable, and hence, since $bel(t_1)$ is complete, it follows from postulate ($\mathbf{TU} \diamond \mathbf{5}$), that $Mod_M(bel(t_1 \diamond \beta)) \subseteq Mod_M(bel(t_1 \diamond \alpha)) \cap Mod_M(\beta)$. But then $Mod_M(bel(t_1 \diamond \beta)) = Mod_M(bel(t_1 \diamond \alpha)) \cap Mod_M(\beta)$. Since $s'' \in Mod_M(bel(t_1 \diamond \alpha))$ and $s' \notin Mod_M(bel(t_1 \diamond \alpha))$, it follows that $Mod_M(bel(t_1 \diamond \beta)) = \{s'\}$. So $(t_1 \diamond \beta)(s') = 0$ and $(t_1 \diamond \beta)(s'') > 0$. If $(t_1 \diamond \beta)(s'') = n$ then $s'' \notin Mod_M(know(t_1 \diamond \beta))$, and hence by postulate ($\mathbf{TU} \diamond \mathbf{7}$) it follows that $s'' \notin Mod_M(know(t_I))$, i.e. $\vec{F}(s_1)(s'') = n$. So $\vec{F}(s_1)(s') \leq \vec{F}(s_1)(s'')$. On the other hand, if $s \in Mod_M(know(t_1 \diamond \beta))$, then by postulate ($\mathbf{TU} \diamond \mathbf{8}$), $\vec{F}(s_1)(s') \leq \vec{F}(s_1)(s'')$. But $s' \in Min_{\vec{F}(s_1)}(Mod_M(\alpha))$ and thus $\vec{F}(s_1)(s'') \leq \vec{F}(s_1)(s')$. So $\vec{F}(s_1)(s') = \vec{F}(s_1)(s'')$. But $(t_1 \diamond \beta)(s') = 0$ and thus $(t_1 \diamond \beta)(s'') = 0$. Contradiction. So $s' \in Mod_M(bel(t_E \diamond \alpha))$. Since s' was chosen arbitrarily, it follows that $\bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha)) \subseteq Mod_M(bel(t_E \diamond \alpha))$. But then $Mod_M(bel(t_E \diamond \alpha)) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$.

2. It must be shown that $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$, which follows directly by postulate ($\mathbf{TU} \diamond \mathbf{7}$).

3. It must be shown that for every $s_1, s_2 \in Mod_M(\alpha)$, if $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. Choose any $s_1, s_2 \in Mod_M(\alpha)$. Suppose that $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$ but $\vec{F}(s)(s_1) \not\leq \vec{F}(s)(s_2)$ for every $s \in Mod_M(bel(t_E))$. Suppose that $(t_E \diamond \alpha)(s_2) = n$, i.e. $s_2 \notin Mod_M(know(t_E \diamond \alpha))$. But $Mod_M(know(t_E \diamond \alpha)) =$

$Mod_M(\alpha)$ by postulate ($\mathbf{TU} \diamond \mathbf{1}$) and hence $s_2 \notin Mod_M(\alpha)$. Contradiction. Suppose that $(t_E \diamond \alpha)(s_2) < n$. But then, by postulate ($\mathbf{TU} \diamond \mathbf{8}$), $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. Contradiction. Since s_1, s_2 was chosen arbitrarily, it follows that for every $s_1, s_2 \in Mod_M(\alpha)$, if $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$.

($\Longleftarrow$) Suppose that conditions 1, 2, and 3 hold. It must be shown that $\diamond$ satisfies postulates ($\mathbf{TU} \diamond \mathbf{1}$) to ($\mathbf{TU} \diamond \mathbf{8}$).

($\mathbf{TU} \diamond \mathbf{1}$) It must be shown that $Mod_M(bel(t_E \diamond \alpha)) \subseteq Mod_M(\alpha)$ and $Mod_M(know(t_E \diamond \alpha) \subseteq Mod_M(\alpha)$. Since $Min_{\vec{F}(s)}(Mod_M(\alpha)) \subseteq Mod_M(\alpha)$ for every $s \in Mod_M(bel(t_E))$ it follows that $\bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha)) \subseteq Mod_M(\alpha)$. But then $Mod_M(bel(t_E \diamond \alpha)$ $\subseteq Mod_M(\alpha)$.

($\mathbf{TU} \diamond \mathbf{2}$) It must be shown that $bel(t_E \diamond \alpha)$ and $know(t_E \diamond \alpha)$ are M-satisfiable.

Since α is M-satisfiable it follows that $Mod_M(\alpha) \neq \varnothing$. But then $\bigcup_{s \in Mod_M(bel(t_E))}$ $Min_{\vec{F}(s)}(Mod_M(\alpha)) \neq \varnothing$ from which it follows that $bel(t_E \diamond \alpha) \neq \varnothing$, i.e. $bel(t_E \diamond \alpha)$ is M-satisfiable. Furthermore, $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$. Since α is M-satisfiable, it follows that $know(t_E \diamond \alpha)$ is M-satisfiable.

($\mathbf{TU} \diamond \mathbf{3}$) Suppose that $t_E = t_{E'}$ and $\alpha \equiv_M \alpha'$. It must be shown that $t_E \diamond \alpha = t_{E'} \diamond \alpha'$. Let $X = Mod_M(\alpha)$ and $S - X = NMod_M(\alpha)$. Choose any $s_1, s_2 \in X$. Suppose that $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$. If $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. Since $t_E(s) = t_{E'}(s)$ for every $s \in S$, it follows that $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for $s \in Mod_M(bel(t_{E'}))$. So it is not the case that $\vec{F}(s)(s_1) \not\leq \vec{F}(s)(s_2)$ for every $s \in Mod_M(bel(t_{E'}))$. But then it is not the case that $(t_{E'} \diamond \alpha)(s_1) \not\leq (t_{E'} \diamond \alpha)(s_2)$. So $(t_{E'} \diamond \alpha)(s_1) \leq (t_{E'} \diamond \alpha)(s_2)$. Hence $t_E \diamond \alpha$ and $t_{E'} \diamond \alpha$ are order-equivalent on X. Now $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$ and thus $S - X \neq Mod_M(know(t_E \diamond \alpha))$, i.e. $S - X \neq get_{\uparrow}(t_E \diamond \alpha, n-1)$. So for every $s_1, s_2 \in S - X$, $(t_E \diamond \alpha)(s) = n$. Since

$Mod_M(know(t_{E'})) = Mod_M(know(t_E))$, it follows that for every $s_1, s_2 \in S - X$, $(t_{E'} \diamond \alpha)(s) = n$. Hence $t_E \diamond \alpha$ and $t_{E'} \diamond \alpha$ are order-equivalent on $S - X$. But $\alpha \equiv_M \alpha'$ and thus $X = Mod_M(\alpha')$ and $S - X = NMod_M(\alpha')$. But then $t_E \diamond \alpha$ and $t_{E'} \diamond \alpha'$ are order-equivalent. Since both $t_E \diamond \alpha$ and $t_{E'} \diamond \alpha'$ are regular, it follows by proposition 3.17 that $t_E \diamond \alpha = t_{E'} \diamond \alpha'$.

(**TU** $\diamond$ **4**) It must be shown that $Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E \diamond (\alpha \wedge \beta)))$. Choose any $s' \in Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta)$. So $s' \in Mod_M(bel(t_E \diamond \alpha))$ and $s' \in Mod_M(\beta)$, i.e. $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha)$, i.e. $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ for some $s \in Mod_M(bel(t_E))$, say s''. So $s' \in Mod_M(\alpha)$ and $s' \in Mod_M(\beta)$, i.e. $s' \in Mod_M(\alpha) \cap Mod_M(\beta)$, i.e. $s' \in Mod_M(\alpha \wedge \beta)$. But since $Mod_M(\alpha \wedge \beta) \subseteq Mod_M(\alpha)$ and $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ it follows that $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha \wedge \beta))$. But then $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha \wedge \beta)$, i.e. $s' \in Mod_M(bel(t_E \diamond (\alpha \wedge \beta)))$. Since s' was chosen arbitrarily it follows that $Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta) \subseteq Mod_M(bel(t_E \diamond (\alpha \wedge \beta)))$.

(**TU** $\diamond$ **5**) Suppose that $bel(t_E \diamond \alpha) \wedge \beta$ is M-satisfiable and $bel(t_E)$ complete. So $bel(t_E)$ has exactly one model, say s''. It must be shown that $Mod_M(bel(t_E \diamond (\alpha \wedge \beta))) \subseteq Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta)$. Choose any $s' \in Mod_M(bel(t_E \diamond (\alpha \wedge \beta)))$. So $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha \wedge \beta))$, i.e. $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha \wedge \beta))$. So $s' \in Mod_M(\alpha \wedge \beta)$, i.e. $s' \in Mod_M(\alpha)$ and $s' \in Mod_M(\beta)$. Suppose that $s' \notin Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta)$. Since $s' \in Mod_M(\beta)$ it follows that $s' \notin Mod_M(bel(t_E \diamond \alpha))$, i.e. $s' \notin Min_{\vec{F}(s'')}(Mod_M(\alpha))$. But $bel(t_E \diamond \alpha) \wedge \beta$ is M-satisfiable and hence there must be some $s''' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ such that $s''' \in Mod_M(\beta)$. So $\vec{F}(s'')(s''') < \vec{F}(s'')(s')$. But then $s''' \in Min_{\vec{F}(s'')}(Mod_M(\alpha \wedge \beta))$ and $s' \notin Min_{\vec{F}(s'')}(Mod_M(\alpha \wedge \beta))$. Contradiction. So $s' \in Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta)$. Since s' was chosen arbitrarily, it follows that $Mod_M(bel(t_E \diamond (\alpha \wedge \beta))) \subseteq Mod_M(bel(t_E \diamond \alpha)) \cap Mod_M(\beta)$.

(**TU** $\diamond$ **6**) It must be shown that $Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha)) = Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha))$. Choose any $s' \in Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha))$. So $s' \in \bigcup_{s \in Mod_M(bel(t_E \bigtriangledown t_{E'}))}$

$Min_{\vec{F}(s)}(Mod_M(\alpha))$, i.e. $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ for some $s'' \in Mod_M(bel(t_E \bigtriangledown t_{E'}))$. By proposition 5.11(3) $bottom(t_E \bigtriangledown t_{E'}) = bottom(t_E) \cup bottom(t_{E'})$, i.e. $Mod_M(bel(t_E \bigtriangledown t_{E'})) = Mod_M(bel(t_E)) \cup Mod_M(bel(t_{E'}))$. But then $s'' \in Mod_M(bel(t_E))$ or $s'' \in Mod_M(bel(t_{E'}))$. If $s'' \in Mod_M(bel(t_E))$ then $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ and thus $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$, i.e. $s' \in Mod_M(bel(t_E \diamond \alpha))$. If $s'' \in Mod_M(bel(t_{E'}))$ then $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ and thus $s' \in \bigcup_{s \in Mod_M(bel(t_{E'}))} Min_{\vec{F}(s)}(Mod_M(\alpha))$, i.e. $s' \in Mod_M(bel(t_{E'} \diamond \alpha))$. So $s' \in Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha))$. Since s' was chosen arbitrarily, it follows that $Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha)) \subseteq Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha))$.

Conversely, choose any $s' \in Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha))$. If $s' \in Mod_M(bel(t_E \diamond \alpha))$ then $s' \in Min_{\vec{F}(s'')}(Mod_M(\alpha))$ for some $s'' \in Mod_M(bel(t_E))$. On the other hand, if $s' \in Mod_M(bel(t_{E'} \diamond \alpha))$ then $s' \in Min_{\vec{F}(s''')}(Mod_M(\alpha))$ for some $s''' \in Mod_M(bel(t_{E'}))$. But $Mod_M(bel(t_E)) \cup Mod_M(bel(t_{E'})) = Mod_M(bel(t_E \bigtriangledown t_{E'}))$. So $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ for some $s \in Mod_M(bel(t_E \bigtriangledown t_{E'}))$, either s'' or s'''. Hence $s' \in Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha)$. Since s' was chosen arbitrarily, it follows that $Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha)) \subseteq Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha))$. But then $Mod_M(bel((t_E \bigtriangledown t_{E'}) \diamond \alpha)) = Mod_M(bel(t_E \diamond \alpha)) \cup Mod_M(bel(t_{E'} \diamond \alpha))$.

(**TU$\diamond$7**) It must be shown that $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$, which holds by condition 2.

(**TU$\diamond$8**) Suppose that $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2) < n$. It must be shown that $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. Since $(t_E \diamond \alpha)(s_1) \leq (t_E \diamond \alpha)(s_2) < n$, it follows that $s_1, s_2 \in Mod_M(know(t_E \diamond \alpha))$. But $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$ and thus $s_1, s_2 \in Mod_M(\alpha)$. But then, by condition 3, $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. ■

Proposition C.12 (5.13) *Let $M = \langle S, l \rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ be an epistemic state with $bel(t_E)$ the belief assertion and*

$know(t_E)$ the knowledge assertion associated with t_E and let $\vec{F}$ be an epistemic transition function. Then the templated update operation $\diamond_{\boxminus}$ satisfies postulates $(\mathbf{TU} \diamond \mathbf{1})$ to $(\mathbf{TU} \diamond \mathbf{8})$ for every satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$.

Proof. Theorem 5.15 will be used to show that the templated update operation $\diamond_{\boxminus}$ satisfies postulates $(\mathbf{TU} \diamond \mathbf{1})$ to $(\mathbf{TU} \diamond \mathbf{8})$. For every sentence $\alpha \in L$, $Mod_M(\alpha) = bottom(t_\alpha)$ and $NMod_M(\alpha) = top(t_\alpha)$ for $t_\alpha \in T_D$ the definite t-ordering induced by α (since M is extensional). It will be shown that for every M-satisfiable sentence $\alpha \in L$ that is inconsistent with $know(t_E)$, the following holds:

1. $Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$
2. $Mod_M(know(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) = Mod_M(\alpha)$
3. For every $s_1, s_2 \in Mod_M(\alpha)$, if $\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_1) \leq \bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$.

1. Choose any $s' \in Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))))$. So $s' \in bottom(\bigtriangledown_{s \in Mod(bel(t_E))} (t_\alpha \boxminus \vec{F}(s)))$. By proposition 5.11(3), $s' \in bottom(t_\alpha \boxminus \vec{F}(s))$ for some $s \in Mod(bel(t_E))$. So $(t_\alpha \boxminus \vec{F}(s))(s') = 0$. But if $(t_\alpha \boxminus \vec{F}(s))(s') = 0$ then $card(seg(t_\alpha(s'), \vec{F}(s)(s'))) = 0$ and by proposition 4.15(1), $t_\alpha(s') = 0$. So $s \in bottom(t_\alpha)$, i.e. $s \in Mod_M(\alpha)$. Suppose there is some $s'' \in Mod_M(\alpha)$ such that $\vec{F}(s)(s'') \leq \vec{F}(s)(s')$ and $\vec{F}(s)(s') \not\leq \vec{F}(s)(s'')$. But then $card(seg(t_\alpha(s'), \vec{F}(s)(s'))) \neq 0$. Contradiction. So there can be no $s'' \in Mod_M(\alpha)$ such that $\vec{F}(s)(s'') \leq \vec{F}(s)(s')$ unless $\vec{F}(s)(s') \leq \vec{F}(s)(s'')$. So $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$. But then $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$. Since s' was chosen arbitrarily, it follows that $Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) \subseteq \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$.

Conversely, choose any $s' \in \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$. So $s' \in Min_{\vec{F}(s)} (Mod_M(\alpha))$ for some $s \in Mod_M(bel(t_E))$. Since $s' \in Mod_M(\alpha)$, it follows that $t_\alpha(s') = 0$. But $s' \in Min_{\vec{F}(s)}(Mod_M(\alpha))$ and hence there can be no $s'' \in Mod_M(\alpha)$ such that $\vec{F}(s)(s'') \leq \vec{F}(s)(s')$ unless $\vec{F}(s)(s') \leq \vec{F}(s)(s'')$. But then $card(seg(t_\alpha(s'), \vec{F}(s)(s'))) = 0$ and thus $(t_\alpha \boxminus \vec{F}(s)))(s') = 0$. So $s' \in bottom(t_\alpha \boxminus \vec{F}(s)))$. But then, by proposition

5.11(3), $s' \in \mathit{bottom}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))$, i.e. $s' \in Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))$. Since s' was chosen arbitrarily, it follows that $\bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha)) \subseteq Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))$. Hence $Mod_M(bel(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$.

2. Choose any $s' \in Mod_M(\mathit{know}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))))$. So $s' \in \mathit{get}_\uparrow(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)), n-1)$. By proposition 5.11(4), $\mathit{top}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))) = \bigcap_{s \in Mod(bel(t_E))} \mathit{top}(t_\alpha \boxminus \vec{F}(s))$. But then $s' \in \mathit{get}_\uparrow(t_\alpha \boxminus \vec{F}(s), n-1)$ for some $s \in Mod(bel(t_E))$. So $(t_\alpha \boxminus \vec{F}(s)))(s') < n$. But then $\mathit{card}(\mathit{seg}(t_\alpha(s'), \vec{F}(s)(s'))) < n$. So $t_\alpha(s') < 0$. But t_α is definite and thus $t_\alpha(s') = 0$, i.e. $s' \in Mod_M(\alpha)$. Since s' was chosen arbitrarily, it follows that $Mod_M(\mathit{know}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) \subseteq Mod_M(\alpha)$.

Conversely, choose any $s' \in Mod_M(\alpha)$. So $s' \in \mathit{bottom}(t_\alpha)$, i.e. $t_\alpha(s') = 0$. But then $\mathit{card}(\mathit{seg}(t_\alpha(s'), \vec{F}(s)(s'))) < n$ for every $s \in Mod(bel(t_E))$. So $(t_\alpha \boxminus \vec{F}(s)))(s') < n$ for every $s \in Mod(bel(t_E))$, i.e. $s' \in \mathit{get}_\uparrow(t_\alpha \boxminus \vec{F}(s), n-1)$ for every $s \in Mod(bel(t_E))$. But then $s' \in \mathit{get}_\uparrow(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)), n-1)$, i.e. $s' \in Mod_M(\mathit{know}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))))$. But s' was chosen arbitrarily and thus $Mod_M(\alpha) \subseteq Mod_M(\mathit{know}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))))$. Hence $Mod_M(\mathit{know}(\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s)))) = Mod_M(\alpha)$.

3. Choose any $s_1, s_2 \in Mod_M(\alpha)$. So $s_1, s_2 \in \mathit{bottom}(t_\alpha)$. Suppose that $\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_1) \leq \bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_2)$. Suppose that $\vec{F}(s)(s_1) \not\leq \vec{F}(s)(s_2)$ for every $s \in Mod_M(bel(t_E))$. But $t_\alpha(s_1) = t_\alpha(s_2) = 0$ and hence $\mathit{card}(\mathit{seg}(t_\alpha(s_1), \vec{F}(s)(s_1))) \not\leq \mathit{card}(\mathit{seg}(t_\alpha(s_2), \vec{F}(s)(s_2))$ for every $s \in Mod_M(bel(t_E))$. But then $(t_\alpha \boxminus \vec{F}(s))(s_1) \not\leq (t_\alpha \boxminus \vec{F}(s))(s_2)$ for every $s \in Mod_M(bel(t_E))$. But then $\min\{(t_\alpha \boxminus \vec{F}(s))(s_1), (t_\alpha \boxminus \vec{F}(s))(s_2)\} = (t_\alpha \boxminus \vec{F}(s))(s_2)$ for every $s \in Mod_M(bel(t_E))$. Hence $\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_2) \leq \bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_1)$. Contradiction. So $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. Since s_1, s_2 was chosen arbitrarily, it follows that for every $s_1, s_2 \in Mod_M(\alpha)$, if $\bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_1) \leq \bigtriangledown_{s \in Mod(bel(t_E))}(t_\alpha \boxminus \vec{F}(s))(s_2)$, then $\vec{F}(s)(s_1) \leq \vec{F}(s)(s_2)$ for some $s \in Mod_M(bel(t_E))$. ∎

Proposition C.13 (5.14) *Let $T = \langle S, t_E, l\rangle$ and $T' = \langle S, t_E \diamond \alpha, l\rangle$ be templated interpretations of L such that l is injective and $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be T-satisfiable and inconsistent with $know(t_E)$. Let $\vec{F}$ be an epistemic transition function. Then the following holds:*

1. $pl(t_E, \alpha) \leq pl(t_E \diamond \alpha, \alpha)$
2. $dt(t_E \diamond \alpha, \alpha) \leq dt(t_E, \alpha)$

Proof. 1. If α is T-valid, then $pl(t_E, \alpha) = pl(t_E \diamond \alpha, \alpha) = n$. Assume that α is not T-valid. Since $know(t_E) \wedge \alpha$ is T-unsatisfiable, it follows that $pl(t_E, \alpha) = -1$. Now $Mod_T(know(t_E \diamond \alpha)) = Mod_T(\alpha)$, i.e. $get_\uparrow(t_E \diamond \alpha, n-1) = Mod_T(\alpha)$. So $pl(t_E \diamond \alpha, \alpha) = n - 1$. Hence $pl(t_E, \alpha) \leq pl(t_E \diamond \alpha, \alpha)$.

2. Since $know(t_E) \wedge \alpha$ is T-unsatisfiable, it follows that $dt(t_E, \alpha) = n$. But $Mod_T(bel(t_E \diamond \alpha)) = \bigcup_{s \in Mod_M(bel(t_E))} Min_{\vec{F}(s)}(Mod_M(\alpha))$ and thus $bottom(t_E \diamond \alpha) \cap Mod_T(\alpha) \neq \varnothing$. But then $dt(t_E \diamond \alpha, \alpha) = 0$. Hence $dt(t_E \diamond \alpha, \alpha) \leq dt(t_E, \alpha)$. ∎

Proposition C.14 *Let $M = \langle S, l\rangle$ be an extensional interpretation of L with $card(S) = n$. Let $t_E \in T_E$ and let $\alpha \in L$ be M-satisfiable and inconsistent with $know(t_E)$. Then $Cont_D(t_E) \not\subset Cont_D(t_E \diamond \alpha)$.*

Proof. Since α is M-satisfiable but inconsistent with $know(t_E)$ it follows that $get_\uparrow(t_\alpha, n-1) \cap Mod_M(know(t_E)) = \varnothing$, i.e. $Cont_D(t_\alpha) \subseteq Cont_D(t_E)$. By postulate (**TU $\diamond$ 7**) it follows that $Mod_M(know(t_E \diamond \alpha)) = Mod_M(\alpha)$, i.e. $Mod_M(know(t_E \diamond \alpha)) = get_\uparrow(t_\alpha, n-1)$. But then $Cont_D(t_E \diamond \alpha) \subseteq Cont_D(t_E)$, i.e. $Cont_D(t_E) \not\subset Cont_D(t_E \diamond \alpha)$. ∎

Appendix D

Proofs for Chapter 6

D.1 Proofs for Section 6.1

Proposition D.1 (6.2) *Let κ and λ be satisfiable knowledge bases of L. Postulate* ($\mathbf{KP} \triangle \mathbf{4}$) *is satisfied by every model-fitting operation of Revesz, not satisfied by every majority merging operation of Lin and Mendelzon, and not satisfied by every arbitration operation of Liberatore and Schaerf.*

Proof. Suppose that $\kappa \wedge \lambda$ is unsatisfiable.

Case 1: Let $\rhd$ be any model-fitting operation of Revesz, i.e. $\rhd$ satisfies postulates $(R \rhd 1)$ to $(R \rhd 7)$. Suppose that $Mod(\kappa \rhd \lambda) \subseteq Mod(\kappa)$. By postulate $(R \rhd 1)$, $Mod(\kappa \rhd \lambda) \subseteq Mod(\lambda)$. So $Mod(\kappa \rhd \lambda) \subseteq Mod(\kappa \wedge \lambda)$. But λ is satisfiable and thus, by postulate $(R \rhd 3)$, $\kappa \rhd \lambda$ is satisfiable. Since $Mod(\kappa \rhd \lambda) \subseteq Mod(\kappa \wedge \lambda)$ it follows that $\kappa \wedge \lambda$ must be satisfiable. Contradiction. So $Mod(\kappa \rhd \lambda) \not\subseteq Mod(\kappa)$, i.e. postulate $(KP \triangle 4)$ is satisfied.

Case 2: Let $\triangle$ be any majority merging operation of Lin and Mendelzon, i.e. $\triangle$ satisfies postulates ($\mathbf{LM} \triangle \mathbf{0}$) to ($\mathbf{LM} \triangle \mathbf{4}$). The proof is by counterexample. Suppose L is generated by $Atom = \{P(a), Q(a)\}$. Let $Mod(\kappa) = \{11\}$ and $Mod(\lambda) = \{10\}$ with α

the literal $P(a)$. So $Mod(\alpha) = \{11, 10\}$. But then both κ and λ entails α while neither κ nor λ entails $\neg\alpha$. Since both κ and λ entails $\neg\neg\alpha$, it follows from proposition 6.1 that $\kappa \not\approx \neg\alpha$ and $\gamma \not\approx \neg\alpha$. So the support for α is greater than the opposition to α together with the partial support for $\neg\alpha$. Hence, by postulate (**LM** $\triangle$ **4**), $Mod(\kappa \triangle \lambda) \subseteq Mod(\alpha)$, i.e. $Mod(\kappa \triangle \lambda) \subseteq \{11, 10\}$. For literal $\neg P(a)$, the support is 0, since neither κ nor λ entails $\neg P(a)$, while for literals $Q(a)$ and $\neg Q(a)$ both the support and opposition is 1, since $Q(a)$ is supported by κ but opposed by λ and vice versa for $\neg Q(a)$. So the support for these literals is not greater than the opposition to them. Hence there are no other restrictions on $Mod(\kappa \triangle \lambda)$ except that $Mod(\kappa \triangle \lambda) \neq \varnothing$, by postulate (**LM** $\triangle$ **1**). So it is possible that $Mod(\kappa \triangle \lambda) = \{11\}$. But then $Mod(\kappa \triangle \lambda) \subseteq Mod(\kappa)$, i.e. postulate (**KP** $\triangle$ **4**) is not satisfied.

Case 3: Let $\triangle$ be any arbitration operation of Liberatore and Schaerf, i.e. $\triangle$ satisfies postulates (**LS** $\triangle$ **1**) to (**LS** $\triangle$ **8**). Choose any $s \in Mod(\kappa \triangle \lambda)$. So $s \in Mod(\lambda \triangle \kappa)$ by postulate (**LS** $\triangle$ **1**). But, by postulate (**LS** $\triangle$ **7**), $s \in Mod(\kappa) \cup Mod(\lambda)$. So either $s \in Mod(\kappa)$ or $s \in Mod(\lambda)$, but not both since $\kappa \wedge \lambda$ is unsatisfiable. Suppose that $s \in Mod(\kappa)$. Since s was chosen arbitrarily, it follows that $Mod(\kappa \triangle \lambda) \subseteq Mod(\kappa)$, i.e. postulate (**KP** $\triangle$ **4**) is not satisfied. Suppose that $s \in Mod(\lambda)$. Since s was chosen arbitrarily, it follows that $Mod(\lambda \triangle \kappa) \subseteq Mod(\lambda)$, i.e. postulate (**KP** $\triangle$ **4**) is not satisfied.

■

Proposition D.2 (6.3) *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* (**KP** $\triangle$ **5**) *is satisfied by every arbitration operation of Liberatore and Schaerf but is not satisfied by every majority merging operation of Lin and Mendelzon.*

Proof. It must be shown that $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \subseteq Mod((\kappa \vee \kappa') \triangle \lambda)$. If $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) = \varnothing$ then the result holds vacuously. Suppose that $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \neq \varnothing$. Choose any $s \in Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda)$. So $s \in Mod(\kappa \triangle \lambda)$ and $s \in Mod(\kappa' \triangle \lambda)$.

Case 1: Let $\triangle$ be any arbitration operation of Liberatore and Schaerf, i.e. $\triangle$ satisfies postulates (**LS** $\triangle$ **1**) to (**LS** $\triangle$ **8**). Suppose that $s \notin Mod((\kappa \vee \kappa') \triangle \lambda)$. Then by postulate (**LS** $\triangle$ **1**), $s \notin Mod(\lambda \triangle (\kappa \vee \kappa'))$. But then, by postulate (**LS** $\triangle$ **6**), $s \notin Mod(\lambda \triangle \kappa)$ and $s \notin Mod(\lambda \triangle \kappa')$ and $s \notin Mod(\lambda \triangle \kappa) \cup Mod(\lambda \triangle \kappa')$. So $s \notin Mod(\kappa \triangle \lambda)$ and $s \notin Mod(\kappa' \triangle \lambda)$, by postulate (**LS** $\triangle$ **1**). Contradiction. Since s was chosen arbitrarily, it follows that $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \subseteq Mod((\kappa \vee \kappa') \triangle \lambda)$.

Case 2: Let $\triangle$ be any majority merging operation of Lin and Mendelzon, i.e. $\triangle$ satisfies postulates (**LM** $\triangle$ **0**) to (**LM** $\triangle$ **4**). The proof is by counterexample. Suppose L is generated by $Atom = \{P(a), Q(a)\}$. Let $Mod(\kappa) = \{11\}$, $Mod(\kappa') = \{00\}$, and $Mod(\lambda) = \{10\}$ with α the literal $P(a)$ and α' the literal $\neg Q(a)$. So $Mod(\alpha) = \{11, 10\}$ and $Mod(\alpha') = \{00, 10\}$. As shown by proposition 6.2(Case 2), $Mod(\kappa \triangle \lambda) \subseteq Mod(\alpha)$, i.e. $Mod(\kappa \triangle \lambda) \subseteq \{11, 10\}$, with no other restrictions on $Mod(\kappa \triangle \lambda)$ except that $Mod(\kappa \triangle \lambda) \neq \varnothing$, by postulate (**LM** $\triangle$ **1**). Similarly, it can be shown that $Mod(\kappa' \triangle \lambda) \subseteq Mod(\alpha')$, i.e. $Mod(\kappa' \triangle \lambda) \subseteq \{00, 10\}$, with no other restrictions on $Mod(\kappa' \triangle \lambda)$ except that $Mod(\kappa' \triangle \lambda) \neq \varnothing$, by postulate (**LM** $\triangle$ **1**). Suppose that $Mod(\kappa \triangle \lambda) = \{11, 10\}$ and $Mod(\kappa' \triangle \lambda) = \{00, 10\}$ so that $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) = \{10\}$. Now $Mod(\kappa \vee \kappa') = \{11, 00\}$. But neither $\kappa \vee \kappa'$ nor λ entails $\neg P(a)$ or $Q(a)$. However, λ entails both $\neg\neg P(a)$ and $\neg Q(a)$. So the support for $\neg P(a)$ and $Q(a)$ are in both cases less than the opposition to these literals. Turning to literals $P(a)$ and $\neg Q(a)$: $\kappa \vee \kappa'$ entails neither but λ entails both. Moreover, neither $\kappa \vee \kappa'$ nor λ entails $\neg\neg P(a)$ or $\neg Q(a)$. However, $\kappa \vee \kappa'$ partially supports both $\neg\neg P(a)$, since $Mod((\kappa \vee \kappa') \wedge \neg P(a)) = \{00\}$ and for $\{11\} \in Mod((\kappa \vee \kappa') \wedge P(a))$ it holds that $\{01\} \notin Mod((\kappa \vee \kappa')$, and $\neg Q(a)$, since $Mod((\kappa \vee \kappa') \wedge \neg\neg Q(a)) = \{11\}$ and for $\{00\} \in Mod((\kappa \vee \kappa') \wedge \neg Q(a))$ it holds that $\{01\} \notin Mod((\kappa \vee \kappa')$. So the support for $P(a)$ and $\neg Q(a)$ are in both cases not greater that the partial support for the negation of these literals. So postulate (**LM** $\triangle$ **4**) places no restriction on $Mod((\kappa \vee \kappa') \triangle \lambda)$. Suppose $Mod((\kappa \vee \kappa') \triangle \lambda) = \{11, 00\}$. But then

$Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \not\subseteq Mod((\kappa \vee \kappa') \triangle \lambda)$, i.e. postulate (**KP** $\triangle$ **5**) is not satisfied. ■

Proposition D.3 (6.4) *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* (**KP** $\triangle$ **5**) *is satisfied by every majority merging operation of Lin and Mendelzon, provided $\kappa \wedge \lambda$ or $\kappa' \wedge \lambda$ is satisfiable.*

Proof. Choose any $s \in Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda)$. So $s \in Mod(\kappa \triangle \lambda)$ and $s \in Mod(\kappa' \triangle \lambda)$. Suppose that $\kappa \wedge \lambda$ is satisfiable. Then by postulate (**LM** $\triangle$ **2**), $\kappa \triangle \lambda \equiv \kappa \wedge \lambda$. So $s \in Mod(\kappa) \cap Mod(\lambda)$, i.e. $s \in Mod(\kappa \vee \kappa') \cap Mod(\lambda)$. But if $\kappa \wedge \lambda$ is satisfiable, then $(\kappa \vee \kappa') \wedge \lambda$ is satisfiable and thus, by postulate (**LM** $\triangle$ **2**), $(\kappa \vee \kappa') \triangle \lambda \equiv (\kappa \vee \kappa') \wedge \lambda$. But then $s \in Mod((\kappa \vee \kappa') \triangle \lambda)$. Similarly if $\kappa' \wedge \lambda$ is satisfiable. Since s was chosen arbitrarily, it follows that $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \subseteq Mod((\kappa \vee \kappa') \triangle \lambda)$, i.e. postulate (**KP** $\triangle$ **5**) is satisfied. ■

Proposition D.4 (6.5) *Let κ, κ', and λ be satisfiable knowledge bases of L. Postulate* (**KP** $\triangle$ **6**) *is not satisfied by every model-fitting operation of Revesz, not satisfied by every arbitration operation of Liberatore and Schaerf, and not satisfied by every majority merging operation of Lin and Mendelzon.*

Proof. Let $Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda) \neq \varnothing$. It must be shown that $Mod((\kappa \vee \kappa') \triangle \lambda) \subseteq Mod(\kappa \triangle \lambda) \cap Mod(\kappa' \triangle \lambda)$.

Case 1: Let $\rhd$ be any model-fitting operation of Revesz, i.e. $\rhd$ satisfies postulates $(R \rhd 1)$ to $(R \rhd 7)$. The proof is by counterexample. Suppose L is generated by $Atom = \{P(a), Q(a)\}$. Let $Mod(\kappa) = \{11\}$, $Mod(\kappa') = \{10\}$, and $Mod(\lambda) = \{00, 01\}$. So $\kappa \wedge \lambda$, $\kappa' \wedge \lambda$, and $(\kappa \vee \kappa') \wedge \lambda$ are all unsatisfiable. By postulate (**R** $\rhd$ **1**), $Mod(\kappa \rhd \lambda) \subseteq \{00, 01\}$, $Mod(\kappa' \rhd \lambda) \subseteq \{00, 01\}$, and $Mod((\kappa \vee \kappa') \rhd \lambda) \subseteq \{00, 01\}$. Suppose that $Mod(\kappa \rhd \lambda) = \{00\}$, $Mod(\kappa' \rhd \lambda) = \{00\}$, and $Mod((\kappa \vee \kappa') \rhd \lambda) = \{00, 01\}$ so that $Mod(\kappa \rhd \lambda) \cap Mod(\kappa' \rhd \lambda) = \{00\}$ and postulate ($\mathbf{R} \rhd \mathbf{7}$) is satisfied. But then $Mod((\kappa \vee \kappa') \rhd \lambda) \nsubseteq Mod(\kappa \rhd \lambda) \cap Mod(\kappa' \rhd \lambda)$, i.e. postulate ($\mathbf{KP} \bigtriangleup \mathbf{6}$) is not satisfied.

Case 2: Let $\bigtriangleup$ be any arbitration operation of Liberatore and Schaerf, i.e. $\bigtriangleup$ satisfies postulates ($\mathbf{LS} \bigtriangleup \mathbf{1}$) to ($\mathbf{LS} \bigtriangleup \mathbf{8}$). The proof is by counterexample. Suppose L is generated by $Atom = \{P(a), Q(a)\}$. Let $Mod(\kappa) = \{11, 01\}$, $Mod(\kappa') = \{10, 01\}$, and $Mod(\lambda) = \{00\}$. So $\kappa \wedge \lambda$, $\kappa' \wedge \lambda$, and $(\kappa \vee \kappa') \wedge \lambda$ are all unsatisfiable. By postulate ($\mathbf{LS} \bigtriangleup \mathbf{7}$), $Mod(\kappa \bigtriangleup \lambda) \subseteq \{11, 01, 00\}$ and $Mod(\kappa' \bigtriangleup \lambda) \subseteq \{10, 01, 00\}$. Suppose that $Mod(\kappa \bigtriangleup \lambda) = \{01, 00\}$ and $Mod(\kappa' \bigtriangleup \lambda) = \{10, 00\}$ so that postulate ($\mathbf{LS} \bigtriangleup \mathbf{8}$) is satisfied. So $Mod(\kappa \bigtriangleup \lambda) \cap Mod(\kappa' \bigtriangleup \lambda) = \{00\}$. Then by postulates ($\mathbf{LS} \bigtriangleup \mathbf{1}$) and ($\mathbf{LS} \bigtriangleup \mathbf{6}$), $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) = \{01, 00\}$ or $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) = \{10, 00\}$ or $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) = \{01, 00, 10\}$. In all three cases, $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) \nsubseteq Mod(\kappa \bigtriangleup \lambda) \cap Mod(\kappa' \bigtriangleup \lambda)$, i.e. postulate ($\mathbf{KP} \bigtriangleup \mathbf{6}$) is not satisfied.

Case 3: Let $\bigtriangleup$ be any majority merging operation of Lin and Mendelzon, i.e. $\bigtriangleup$ satisfies postulates ($\mathbf{LM} \bigtriangleup \mathbf{0}$) to ($\mathbf{LM} \bigtriangleup \mathbf{4}$). The proof is by counterexample. Using the example from proposition 6.3(Case 2), the result is $Mod(\kappa \bigtriangleup \lambda) \cap Mod(\kappa' \bigtriangleup \lambda) = \{10\}$ and $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) = \{11, 00\}$. But then $Mod((\kappa \vee \kappa') \bigtriangleup \lambda) \nsubseteq Mod(\kappa \bigtriangleup \lambda) \cap Mod(\kappa' \bigtriangleup \lambda)$, i.e. postulate ($\mathbf{KP} \bigtriangleup \mathbf{6}$) is not satisfied. ■

D.2 Proofs for Section 6.5

Proposition D.5 (6.15) *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Every t-ordering produced by a basic templated merging operation $\bigtriangleup_\odot$ is regular.*

Proof. It must be shown that $\bigtriangleup_\odot(\mathbf{T})$ is in normal form and not strongly contradictory.

Suppose there is some $i \in B$ such that $i < n$ and $get_\rightarrow(\bigtriangleup_\odot(\mathbf{T}), i) = \varnothing$, but $get(\bigtriangleup_\odot(\mathbf{T}), i, n-1) \neq \varnothing$. So there can be no $s \in S$ such that $\bigtriangleup_\odot(\mathbf{T})(s) = i$ but there must be some $s', s'' \in S$ such that $\triangle_{\odot}(\mathbf{T})(s') < i$ and $i < \triangle_{\odot}(\mathbf{T})(s'') < n-1$. Since both $\triangle_{\odot}(\mathbf{T})(s') < n-1$ and $\triangle_{\odot}(\mathbf{T})(s'') < n-1$, it follows by definition 6.15 that $card(seg(f_{\odot}(s'))) < i$ and $card(seg(f_{\odot}(s''))) > i$. But $A_{\odot}$ is well-ordered and hence there must be some $f_{\odot}(s) \in A_{\odot}$ such that $card(seg(f_{\odot}(s))) = i$. From the linear order $\leq$ on $A_{\odot}$ it follows that $f_{\odot}(s) < f_{\odot}(s'') < k$. But then $\triangle_{\odot}(\mathbf{T})(s) = i$. Contradiction. So for each $i \in B$, it holds that if $i < n$ and $get_{\rightarrow}(\triangle_{\odot}(\mathbf{T}), i) = \varnothing$ then $get(\triangle_{\odot}(\mathbf{T}), i, n-1) = \varnothing$. But then $\triangle_{\odot}(\mathbf{T})$ is in normal form.

Suppose that $get_{\uparrow}(\triangle_{\odot}(\mathbf{T}), n-1) = \varnothing$. So $\triangle_{\odot}(\mathbf{T})(s) = n$ for every $s \in S$. But then, by definition 6.15, $f_{\odot}(s) = k$ for every $s \in S$ and $card(A_{\odot}) > 1$. But if $f_{\odot}(s) = k$ for every $s \in S$ then $ran(f_{\odot}) = \{k\}$ and thus $card(A_{\odot}) = 1$. Contradiction. So $get_{\uparrow}(\triangle_{\odot}(\mathbf{T}), n-1) \neq \varnothing$, i.e. $\triangle_{\odot}(\mathbf{T})$ is not strongly contradictory. ■

Lemma D.1 (6.1) *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Let $\triangle_{\odot}$ be a basic templated merging operation based on an indexing function $f_{\odot} : S \rightarrow \{0, 1, \ldots, k\}$ (with respect to $\mathbf{T}$). For every $s, s' \in S$,*

- *if $f_{\odot}(s) = f_{\odot}(s')$ then $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$ and*
- *if $f_{\odot}(s) < f_{\odot}(s')$ then $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$.*

Proof. Choose any $s, s' \in S$. Suppose that $f_{\odot}(s) = f_{\odot}(s')$. If $f_{\odot}(s) < k$ or $card(A_{\odot}) = 1$ then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$ and $\triangle_{\odot}(\mathbf{T})(s') = card(seg(f_{\odot}(s')))$. But $card(seg(f_{\odot}(s))) = card(seg(f_{\odot}(s')))$ and thus $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$. If $f_{\odot}(s) = k$ and $card(A_{\odot}) \neq 1$ then $\triangle_{\odot}(\mathbf{T})(s) = n$ and $\triangle_{\odot}(\mathbf{T})(s') = n$, i.e. $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$.

Suppose that $f_{\odot}(s) < f_{\odot}(s')$. So $card(A_{\odot}) \neq 1$ and $f_{\odot}(s) < k$. But then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$. If $f_{\odot}(s') = k$ then $\triangle_{\odot}(\mathbf{T})(s') = n$ and since $card(seg(f_{\odot}(s))) < n$ it follows that $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$. If $f_{\odot}(s') < k$ then $\triangle_{\odot}(\mathbf{T})(s') = card(seg(f_{\odot}(s')))$ and since $card(seg(f_{\odot}(s))) < card(seg(f_{\odot}(s')))$ by virtue of $f_{\odot}(s) < f_{\odot}(s')$, it follows that $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$. ■

Proposition D.6 (6.16) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Min}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{5})$*.*

Proof. Templated merging operation $\triangle_{Min}$ is based on indexing function $f_{Min} : S \to \{0, 1, \ldots, n\}$.

$(\mathbf{TM}\triangle\mathbf{1})$ It must be shown that $bottom(\triangle_{Min}(\mathbf{T})) \neq \varnothing$. By proposition 6.15, $\triangle_{Min}(\mathbf{T})$ is regular. But then $bottom(\triangle_{Min}(\mathbf{T})) \neq \varnothing$.

$(\mathbf{TM}\triangle\mathbf{2})$ It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{Min}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$, including t_1. But then $f_{Min}(s) = t_1(s) = 0$ and hence $\triangle_{Min}(\mathbf{T})(s) = card(seg(0))$. Since $card(seg(0)) = 0$ it follows that $s \in bottom(\triangle_{Min}(\mathbf{T}))$. Conversely, choose any $s \in bottom(\triangle_{Min}(\mathbf{T}))$, i.e. $\triangle_{Min}(\mathbf{T})(s) = 0$. So $card(seg(f_{Min}(s))) = 0$, i.e. $f_{Min}(s)$ is the least element in A_{Min}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $f_{Min}(s) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}\})$, i.e. $f_{Min}(s) \neq 0$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{Min}(s') = 0$, i.e. $f_{Min}(s')$ is the least element in A_{Min}. So $f_{Min}(s') = f_{Min}(s) = 0$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

$(\mathbf{TM}\triangle\mathbf{3})$ It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let $f_{Min} : S \to \{0, 1, \ldots, n\}$ be the indexing function with respect to $\mathbf{T}$ and $f'_{Min} : S \to \{0, 1, \ldots, n\}$ the indexing function with respect to $\mathbf{T}'$. From definition 6.16 it is clear that $f_{Min}(s) = f'_{Min}(s)$ for every $s \in S$. But then, $\triangle_{Min}(\mathbf{T}) = \triangle_{Min}(\mathbf{T}')$.

$(\mathbf{TM}\triangle\mathbf{4})$ It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{Min}(\mathbf{T})(s) \leq \triangle_{Min}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Consider the case where $t_i(s) = t_j(s)$ for every $t_i, t_j \in \mathbf{T}$. So $f_{Min}(s) = t_1(s)$. There are two subcases. *i)* If $t_i(s') = t_j(s')$ for every $t_i, t_j \in \mathbf{T}$ then $f_{Min}(s') = t_1(s')$. But $t_1(s) \leq t_k(s')$

and thus $f_{Min}(s) \leq f_{Min}(s')$. *ii)* If $t_i(s') \neq t_j(s')$ for some $t_i, t_j \in \mathbf{T}$, then $f_{Min}(s') = succ(\min\{t_i(s') \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s') \mid t_i \in \mathbf{T}\} = t_k(s')$. But $t_1(s) \leq t_k(s')$ and thus $t_1(s) \leq succ(t_k(s'))$, i.e. $f_{Min}(s) \leq f_{Min}(s')$.

Consider the case where $t_i(s) \neq t_j(s)$ for some $t_i, t_j \in \mathbf{T}$, in which case $f_{Min}(s) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s) \mid t_i \in \mathbf{T}\} = t_k(s)$. There are two subcases. *i)* If $t_i(s') = t_j(s')$ for every $t_i, t_j \in \mathbf{T}$ then $f_{Min}(s') = t_1(s')$. Since $t_i(s) \neq t_j(s)$ for some $t_i, t_j \in \mathbf{T}$, there must be some $t_l \in \mathbf{T}$ such that $t_k(s) < t_l(s)$. But $t_l(s) \leq t_l(s') = t_1(s')$ and thus $succ(t_k(s)) \leq t_1(s')$, i.e $f_{Min}(s) \leq f_{Min}(s')$. *ii)* If $t_i(s') \neq t_j(s')$ for some $t_i, t_j \in \mathbf{T}$, then $f_{Min}(s') = succ(\min\{t_i(s') \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s') \mid t_i \in \mathbf{T}\} = t_l(s')$. But $t_l(s) \leq t_l(s')$ and since $t_k(s) \leq t_l(s)$ it follows that $t_k(s) \leq t_l(s')$, i.e. $succ(t_k(s)) \leq succ(t_l(s'))$, i.e. $f_{Min}(s) \leq f_{Min}(s')$. So in all cases $f_{Min}(s) \leq f_{Min}(s')$ and thus, by lemma 6.1, $\triangle_{Min}(\mathbf{T})(s) \leq \triangle_{Min}(\mathbf{T})(s')$.

(**TM $\triangle$ 5**) It must be shown that if $\triangle_{Min}(\mathbf{T})(s) \leq \triangle_{Min}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{Min}(\mathbf{T})(s) \leq \triangle_{Min}(\mathbf{T})(s')$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. Consider the case where $t_i(s') = t_j(s')$ for every $t_i, t_j \in \mathbf{T}$. So $f_{Min}(s') = t_1(s')$. There are two subcases. *i)* If $t_i(s) = t_j(s)$ for every $t_i, t_j \in \mathbf{T}$ then $f_{Min}(s) = t_1(s)$. But $t_1(s') < t_1(s)$ and thus $f_{Min}(s') < f_{Min}(s)$. *ii)* If $t_i(s) \neq t_j(s)$ for some $t_i, t_j \in \mathbf{T}$, then $f_{Min}(s) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s) \mid t_i \in \mathbf{T}\} = t_k(s)$. But $t_1(s') < t_k(s)$ and thus $t_1(s') < succ(t_k(s))$, i.e. $f_{Min}(s') < f_{Min}(s)$.

Consider the case where $t_i(s') \neq t_j(s')$ for some $t_i, t_j \in \mathbf{T}$. So $f_{Min}(s') = succ(\min\{t_i(s') \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s') \mid t_i \in \mathbf{T}\} = t_k(s')$. There are two subcases. *i)* If $t_i(s) = t_j(s)$ for every $t_i, t_j \in \mathbf{T}$ then $f_{Min}(s) = t_1(s)$. Since $t_i(s') \neq t_j(s')$ for some $t_i, t_j \in \mathbf{T}$, there must be some $t_l \in \mathbf{T}$ such that $t_k(s') < t_l(s')$. But $t_l(s') < t_l(s) = t_1(s)$ and thus $succ(t_k(s')) < t_1(s)$, i.e $f_{Min}(s') < f_{Min}(s)$. *ii)* If $t_i(s) \neq t_j(s)$ for some $t_i, t_j \in \mathbf{T}$, then $f_{Min}(s) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}\})$. Let $\min\{t_i(s) \mid t_i \in \mathbf{T}\} = t_l(s)$. But $t_l(s') < t_l(s)$ and since $t_k(s') \leq t_l(s')$ it follows that $t_k(s') < t_l(s)$, i.e. $succ(t_k(s')) < succ(t_l(s\prime))$, i.e. $f_{Min}(s') < f_{Min}(s)$. So in all cases $f_{Min}(s') < f_{Min}(s)$ and thus, by lemma 6.1, $\triangle_{Min}(\mathbf{T})(s') < \triangle_{Min}(\mathbf{T})(s)$. Contradiction. But then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.7 (6.17) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Max}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{5})$.

Proof. Templated merging operation $\triangle_{Max}$ is based on indexing function $f_{Max} : S \to \{0, 1, \ldots, n\}$.

$(\mathbf{TM}\triangle\mathbf{1})$ It must be shown that $bottom(\triangle_{Max}(\mathbf{T})) \neq \varnothing$. By proposition 6.15, $\triangle_{Max}(\mathbf{T})$ is regular. But then $bottom(\triangle_{Max}(\mathbf{T})) \neq \varnothing$.

$(\mathbf{TM}\triangle\mathbf{2})$ It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{Max}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $\max\{t_i(s) \mid t_i \in \mathbf{T}\} = 0$, i.e. $f_{Max}(s) = 0$ and hence $\triangle_{Max}(\mathbf{T})(s) = card(seg(0))$. Since $card(seg(0)) = 0$ it follows that $s \in bottom(\triangle_{Max}(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_{Max}(\mathbf{T})$, i.e. $\triangle_{Max}(\mathbf{T})(s) = 0$. So $card(seg(f_{Max}(s)) = 0$, i.e. $f_{Max}(s)$ is the least element in A_{Max}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $\max\{t_i(s) \mid t_i \in \mathbf{T}\} > 0$, i.e. $f_{Max}(s) \neq 0$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{Max}(s') = 0$, i.e. $f_{Max}(s')$ is the least element in A_{Max}. So $f_{Max}(s') = f_{Max}(s) = 0$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

$(\mathbf{TM}\triangle\mathbf{3})$ It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let $f_{Max} : S \to \{0, 1, \ldots, n\}$ be the indexing function with respect to $\mathbf{T}$ and $f'_{Max} : S \to \{0, 1, \ldots, n\}$ the indexing function with respect to $\mathbf{T}'$. From definition 6.16 it is clear that $f_{Max}(s) = f'_{Max}(s)$ for every $s \in S$. But then, $\triangle_{Max}(\mathbf{T}) = \triangle_{Max}(\mathbf{T}')$.

$(\mathbf{TM}\triangle\mathbf{4})$ It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{Max}(\mathbf{T})(s) \leq \triangle_{Max}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Let

$\max\{t_i(s) \mid t_i \in \mathbf{T}\} = t_k(s)$. Since $t_k(s) \leq t_k(s')$ it follows that $t_k(s) \leq \max\{t_i(s') \mid t_i \in \mathbf{T}\}$. But then $f_{Max}(s) \leq f_{Max}(s')$. By lemma 6.1, it follows that $\triangle_{Max}(\mathbf{T})(s) \leq \triangle_{Max}(\mathbf{T})(s')$.

(**TM** $\triangle$ **5**) It must be shown that if $\triangle_{Max}(\mathbf{T})(s) \leq \triangle_{Max}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{Max}(\mathbf{T})(s) \leq \triangle_{Max}(\mathbf{T})(s')$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. Let $\max\{t_i(s') \mid t_i \in \mathbf{T}\} = t_k(s')$. Since $t_k(s') < t_k(s)$ it follows that $t_k(s') < \max\{t_i(s) \mid t_i \in \mathbf{T}\}$. But then $f_{Max}(s') < f_{Max}(s)$. By lemma 6.1, it follows that $\triangle_{Max}(\mathbf{T})(s') < \triangle_{Max}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.8 (6.18) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_\Sigma$ *satisfies properties* (**TM** $\triangle$ **1**) *to* (**TM** $\triangle$ **5**).

Proof. Templated merging operation $\triangle_\Sigma$ is based on indexing function $f_\Sigma : S \rightarrow \{0, 1, \ldots, m \times n\}$.

(**TM** $\triangle$ **1**) It must be shown that $bottom(\triangle_\Sigma(\mathbf{T})) \neq \varnothing$. By proposition 6.15, $\triangle_\Sigma(\mathbf{T})$ is regular. But then $bottom(\triangle_\Sigma(\mathbf{T})) \neq \varnothing$.

(**TM** $\triangle$ **2**) It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_\Sigma(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $\sum\{t_i(s) \mid t_i \in \mathbf{T}\} = 0$, i.e. $f_\Sigma(s) = 0$ and hence $\triangle_\Sigma(\mathbf{T})(s) = card(seg(0))$. Since $card(seg(0)) = 0$ it follows that $s \in bottom(\triangle_\Sigma(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_\Sigma(\mathbf{T})$, i.e. $\triangle_\Sigma(\mathbf{T})(s) = 0$. So $card(seg(f_\Sigma(s)) = 0$, i.e. $f_\Sigma(s)$ is the least element in A_Σ. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $\sum\{t_i(s) \mid t_i \in \mathbf{T}\} > 0$, i.e. $f_\Sigma(s) \neq 0$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_\Sigma(s') = 0$, i.e. $f_\Sigma(s')$ is the least element in A_Σ. So $f_\Sigma(s') = f_\Sigma(s) = 0$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

(**TM** $\triangle$ **3**) It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let $f_\Sigma : S \rightarrow \{0, 1, \ldots, m \times n\}$ be the indexing function with respect to $\mathbf{T}$ and

$f'_\Sigma : S \rightarrow \{0, 1, \dots, m \times n\}$ the indexing function with respect to $\mathbf{T}'$. From definition 6.16 it is clear that $f_\Sigma(s) = f'_\Sigma(s)$ for every $s \in S$. But then, $\triangle_\Sigma(\mathbf{T}) = \triangle_\Sigma(\mathbf{T}')$.

(**TM** $\triangle$ **4**) It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_\Sigma(\mathbf{T})(s) \leq \triangle_\Sigma(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. But then $\sum\{t_i(s) \mid t_i \in \mathbf{T}\} \leq \sum t_i(s') \mid t_i \in \mathbf{T}$, i.e. $f_\Sigma(s) \leq f_\Sigma(s')$. But then, by lemma 6.1, it follows that $\triangle_\Sigma(\mathbf{T})(s) \leq \triangle_\Sigma(\mathbf{T})(s')$.

(**TM** $\triangle$ **5**) It must be shown that if $\triangle_\Sigma(\mathbf{T})(s) \leq \triangle_\Sigma(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_\Sigma(\mathbf{T})(s) \leq \triangle_\Sigma(\mathbf{T})(s')$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $\sum t_i(s') \mid t_i \in \mathbf{T} < \sum\{t_i(s) \mid t_i \in \mathbf{T}\}$, i.e. $f_\Sigma(s') < f_\Sigma(s)$. But then, by lemma 6.1, it follows that $\triangle_\Sigma(\mathbf{T})(s') < \triangle_\Sigma(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.9 (6.19) *Let $\mathbf{T} = (t_1, t_2, \dots, t_m)$ be an information tuple. Every t-ordering produced by a refined templated merging operation $\triangle_{R\odot}$ is regular.*

Proof. It must be shown that $\triangle_{R\odot}(\mathbf{T})$ is in normal form and not strongly contradictory.

Suppose there is some $i \in B$ such that $i < n$ and $get_\rightarrow(\triangle_{R\odot}(\mathbf{T}), i) = \varnothing$, but $get(\triangle_{R\odot}(\mathbf{T}), i, n-1) \neq \varnothing$. So there can be no $s \in S$ such that $\triangle_{R\odot}(\mathbf{T})(s) = i$ but there must be some $s', s'' \in S$ such that $\triangle_{R\odot}(\mathbf{T})(s') < i$ and $i < \triangle_{R\odot}(\mathbf{T})(s'') < n-1$. Since both $\triangle_{R\odot}(\mathbf{T})(s') < n-1$ and $\triangle_{R\odot}(\mathbf{T})(s'') < n-1$, it follows by definition 6.18 that $card(seg(f_{R\odot}(s'))) < i$ and $card(seg(f_{R\odot}(s''))) > i$. But $A_{R\odot}$ is well-ordered and hence there must be some $f_{R\odot}(s) \in A_\odot$ such that $card(seg(f_{R\odot}(s))) = i$. From the lexicographic ordering $\preceq$ on $A_{R\odot}$ it follows that $first(f_{R\odot}(s)) < first(f_{R\odot}(s'')) < k$. But then $\triangle_{R\odot}(\mathbf{T})(s) = i$. Contradiction. So for each $i \in B$, it holds that if $i < n$ and $get_\rightarrow(\triangle_{R\odot}(\mathbf{T}), i) = \varnothing$ then $get(\triangle_{R\odot}(\mathbf{T}), i, n-1) = \varnothing$. But then $\triangle_{R\odot}(\mathbf{T})$ is in normal form.

Suppose that $get_{\uparrow}(\triangle_{R\odot}(\mathbf{T}), n-1) = \varnothing$. So $\triangle_{R\odot}(\mathbf{T})(s) = n$ for every $s \in S$. But then, by definition 6.18, $first(f_{R\odot}(s)) = k$ for every $s \in S$ and $first(\min(A_{R\odot})) \neq k$. But if $first(f_{R\odot}(s)) = k$ for every $s \in S$ then $first(\min(A_{R\odot})) = k$. Contradiction. So $get_{\uparrow}(\triangle_{R\odot}(\mathbf{T}), n-1) \neq \varnothing$, i.e. $\triangle_{R\odot}(\mathbf{T})$ is not strongly contradictory. ■

Lemma D.2 (6.2) *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Let $\triangle_{R\odot}$ be a refined templated merging operation based on a refined indexing function $f_{R\odot} : S \to \{0, 1, \ldots, k\}^l$ (with respect to $\mathbf{T}$). For every $s, s' \in S$,*

- *if $f_{R\odot}(s) = f_{R\odot}(s')$ then $\triangle_{R\odot}(\mathbf{T})(s) = \triangle_{R\odot}(\mathbf{T})(s')$*
- *if $f_{R\odot}(s) \prec f_{R\odot}(s')$ and ($first(f_{\odot}(s)) < k$ or $first(f_{\odot}(s')) < k$ or $first(\min(A_{R\odot})) = k$) then $\triangle_{R\odot}(\mathbf{T})(s) < \triangle_{R\odot}(\mathbf{T})(s')$*
- *if $f_{R\odot}(s) \prec f_{R\odot}(s')$ and ($first(f_{\odot}(s)) = k$ and $first(f_{\odot}(s')) = k$ and $first(\min(A_{R\odot})) \neq k$) then $\triangle_{R\odot}(\mathbf{T})(s) = \triangle_{R\odot}(\mathbf{T})(s')$*

Proof. Choose any $s, s' \in S$. Suppose that $f_{\odot}(s) = f_{\odot}(s')$. If $first(f_{\odot}(s)) < k$ or $first(\min(A_{R\odot})) = k$ then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$ and $\triangle_{\odot}(\mathbf{T})(s') = card(seg(f_{\odot}(s')))$. But $card(seg(f_{\odot}(s))) = card(seg(f_{\odot}(s')))$ and thus $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$. If $first(f_{\odot}(s)) = k$ and $first(\min(A_{R\odot})) \neq k$ then $\triangle_{\odot}(\mathbf{T})(s) = n$ and $\triangle_{\odot}(\mathbf{T})(s') = n$, i.e. $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$.

Suppose that $f_{\odot}(s) \prec f_{\odot}(s')$ and that $first(f_{\odot}(s)) < k$ or $first(f_{\odot}(s')) < k$ or $first(\min(A_{R\odot})) = k$. If $first(\min(A_{R\odot})) = k$ then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$ and $\triangle_{\odot}(\mathbf{T})(s') = card(seg(f_{\odot}(s')))$. But since $f_{\odot}(s) \prec f_{\odot}(s')$ it follows that $card(seg(f_{\odot}(s))) < card(seg(f_{\odot}(s')))$. But then $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$. Assume $first(\min(A_{R\odot})) \neq k$. If $first(f_{\odot}(s)) < k$ and $first(f_{\odot}(s')) < k$ then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$ and $\triangle_{\odot}(\mathbf{T})(s') = card(seg(f_{\odot}(s')))$. But $f_{\odot}(s) \prec f_{\odot}(s')$ and thus $card(seg(f_{\odot}(s))) < card(seg(f_{\odot}(s')))$. Hence $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$. If $first(f_{\odot}(s)) < k$ and $first(f_{\odot}(s')) =$

k then $\triangle_{\odot}(\mathbf{T})(s) = card(seg(f_{\odot}(s)))$ and $\triangle_{\odot}(\mathbf{T})(s') = n$. But $card(seg(f_{\odot}(s))) < n$ and hence $\triangle_{\odot}(\mathbf{T})(s) < \triangle_{\odot}(\mathbf{T})(s')$.

Suppose that $f_{\odot}(s) \prec f_{\odot}(s')$ and that $first(f_{\odot}(s)) = k$ and $first(f_{\odot}(s')) = k$ and $first(\min(A_{R\odot})) \neq k$. But then $\triangle_{\odot}(\mathbf{T})(s) = n$ and $\triangle_{\odot}(\mathbf{T})(s') = n$, i.e. $\triangle_{\odot}(\mathbf{T})(s) = \triangle_{\odot}(\mathbf{T})(s')$. ■

Proposition D.10 (6.20) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RMin}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{5})$.

Proof. Templated merging operation $\triangle_{RMin}$ is based on indexing function $f_{RMin} : S \to \{0, 1, \ldots, n\}^m$.

$(\mathbf{TM} \triangle \mathbf{1})$ It must be shown that $bottom(\triangle_{RMin}(\mathbf{T})) \neq \varnothing$. By proposition 6.19, $\triangle_{RMin}(\mathbf{T})$ is regular. But then $bottom(\triangle_{RMin}(\mathbf{T})) \neq \varnothing$.

$(\mathbf{TM} \triangle \mathbf{2})$ It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{RMin}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RMin}(s) = (0, 0, \ldots, 0)$ and hence $\triangle_{RMin}(\mathbf{T})(s) = card(seg(0, 0, \ldots, 0))$. Since $card(seg(0, 0, \ldots, 0)) = 0$ it follows that $s \in bottom(\triangle_{RMin}(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_{RMin}(\mathbf{T})$, i.e. $\triangle_{RMin}(\mathbf{T})(s) = 0$. So $card(seg(f_{RMin}(s)) = 0$, i.e. $f_{RMin}(s)$ is the least element in A_{RMin}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $f_{RMin}(s) \neq (0, 0, \ldots, 0)$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RMin}(s') = (0, 0, \ldots, 0)$, i.e. $f_{RMin}(s')$ is the least element in A_{RMin}. So $f_{RMin}(s') = f_{RMin}(s) = (0, 0, \ldots, 0)$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

$(\mathbf{TM} \triangle \mathbf{3})$ It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let f_{RMin} be the refined indexing function with respect to $\mathbf{T}$ and f'_{RMin}

the refined indexing function with respect to $\mathbf{T}'$. From definition 6.19 it is clear that $f_{RMin}(s) = f'_{RMin}(s)$ for every $s \in S$. But then, $\triangle_{RMin}(\mathbf{T}) = \triangle_{RMin}(\mathbf{T}')$.

($\mathbf{TM} \triangle \mathbf{4}$) It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{RMin}(\mathbf{T})(s) \leq \triangle_{RMin}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Let $f_{RMin}(s) = (x_1, x_2, \ldots, x_m)$ and $f_{RMin}(s') = (y_1, y_2, \ldots, y_m)$. Since both $f_{RMin}(s)$ and $f_{RMin}(s')$ are ordered increasingly it follows that $x_1 \leq x_2 \leq \ldots \leq x_m$ and $y_1 \leq y_2 \leq \ldots \leq y_m$. But $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. So $x_1 \leq y_i$ for every $i = 1, 2, \ldots, m$. If $x_1 = y_1$ then $x_2 \leq y_i$ for every $i = 2, \ldots, m$. And so on. But then $f_{RMin}(s) \preceq f_{RMin}(s')$. Hence, by lemma 6.2, it follows that $\triangle_{RMin}(\mathbf{T})(s) \leq \triangle_{RMin}(\mathbf{T})(s')$.

($\mathbf{TM} \triangle \mathbf{5}$) It must be shown that if $\triangle_{RMin}(\mathbf{T})(s) \leq \triangle_{RMin}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{RMin}(\mathbf{T})(s) \leq \triangle_{RMin}(\mathbf{T})(s')$. Let $f_{RMin}(s) = (x_1, x_2, \ldots, x_m)$ and $f_{RMin}(s') = (y_1, y_2, \ldots, y_m)$. Since both $f_{RMin}(s)$ and $f_{RMin}(s')$ are ordered increasingly it follows that $x_1 \leq x_2 \leq \ldots \leq x_m$ and $y_1 \leq y_2 \leq \ldots \leq y_m$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $y_1 < x_1$. So $f_{RMin}(s') \prec f_{RMin}(s)$. But then, by lemma 6.2, it follows that $\triangle_{RMin}(\mathbf{T})(s') \leq \triangle_{RMin}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.11 (6.21) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RMax}$ *satisfies properties* ($\mathbf{TM} \triangle \mathbf{1}$) *to* ($\mathbf{TM} \triangle \mathbf{5}$).

Proof. Templated merging operation $\triangle_{RMax}$ is based on indexing function $f_{RMax} : S \rightarrow \{0, 1, \ldots, n\}^m$.

($\mathbf{TM} \triangle \mathbf{1}$) It must be shown that $bottom(\triangle_{RMax}(\mathbf{T})) \neq \varnothing$. By proposition 6.19, $\triangle_{RMax}(\mathbf{T})$ is regular. But then $bottom(\triangle_{RMax}(\mathbf{T})) \neq \varnothing$.

($\mathbf{TM} \triangle \mathbf{2}$) It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{RMax}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any

$s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RMax}(s) = (0, 0, \ldots, 0)$. So $card(seg(f_{RMax}(s))) = 0$, i.e. $s \in bottom(\triangle_{RMax}(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_{RMax}(\mathbf{T})$, i.e. $\triangle_{RMax}(\mathbf{T})(s) = 0$. So $card(seg(f_{RMax}(s)) = 0$, i.e. $f_{RMax}(s)$ is the least element in A_{RMax}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $f_{RMax}(s) \neq (0, 0, \ldots, 0)$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RMax}(s') = (0, 0, \ldots, 0)$, i.e. $f_{RMax}(s')$ is the least element in A_{RMax}. So $f_{RMax}(s') = f_{RMax}(s) = (0, 0, \ldots, 0)$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

(**TM** $\triangle$ **3**) It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let f_{RMax} be the refined indexing function with respect to $\mathbf{T}$ and f'_{RMax} the refined indexing function with respect to $\mathbf{T}'$. From definition 6.19 it is clear that $f_{RMax}(s) = f'_{RMax}(s)$ for every $s \in S$. But then, $\triangle_{RMax}(\mathbf{T}) = \triangle_{RMax}(\mathbf{T}')$.

(**TM** $\triangle$ **4**) It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{RMax}(\mathbf{T})(s) \leq \triangle_{RMax}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Let $f_{RMax}(s) = (x_1, x_2, \ldots, x_m)$ and $f_{RMax}(s') = (y_1, y_2, \ldots, y_m)$. Since both $f_{RMax}(s)$ and $f_{RMax}(s')$ are ordered decreasingly it follows that $x_1 \geq x_2 \geq \ldots \geq x_m$ and $y_1 \geq y_2 \geq \ldots \geq y_m$. But $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. So $x_1 \leq y_1$. If $x_1 = y_1$ then $x_2 \leq y_2$. And so on. But then $f_{RMax}(s) \preceq f_{RMax}(s')$. Hence, by lemma 6.2, it follows that $\triangle_{RMax}(\mathbf{T})(s) \leq \triangle_{RMax}(\mathbf{T})(s')$.

(**TM** $\triangle$ **5**) It must be shown that if $\triangle_{RMax}(\mathbf{T})(s) \leq \triangle_{RMax}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{RMax}(\mathbf{T})(s) \leq \triangle_{RMax}(\mathbf{T})(s')$. Let $f_{RMax}(s) = (x_1, x_2, \ldots, x_m)$ and $f_{RMax}(s') = (y_1, y_2, \ldots, y_m)$. Since both $f_{RMax}(s)$ and $f_{RMax}(s')$ are ordered decreasingly it follows that $x_1 \geq x_2 \geq \ldots \geq x_m$ and $y_1 \geq y_2 \geq \ldots \geq y_m$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $y_1 < x_1$. So $f_{RMax}(s') \prec f_{RMax}(s)$. But then, by lemma 6.2, it follows that $\triangle_{RMax}(\mathbf{T})(s') \leq \triangle_{RMax}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.12 (6.22) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{R\Sigma}$ *satisfies properties* $(\mathbf{TM} \triangle \mathbf{1})$ *to* $(\mathbf{TM} \triangle \mathbf{5})$.

Proof. Templated merging operation $\triangle_{R\Sigma}$ is based on indexing function $f_{R\Sigma} : S \to \{0, 1, \ldots, m \times n\}^n$.

$(\mathbf{TM} \triangle \mathbf{1})$ It must be shown that $bottom(\triangle_{R\Sigma}(\mathbf{T})) \neq \varnothing$. By proposition 6.19, $\triangle_{R\Sigma}(\mathbf{T})$ is regular. But then $bottom(\triangle_{R\Sigma}(\mathbf{T})) \neq \varnothing$.

$(\mathbf{TM} \triangle \mathbf{2})$ It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{R\Sigma}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $f_{R\Sigma}(s) = (0, 0, \ldots, 0)$. So $card(seg(f_{R\Sigma}(s))) = 0$, i.e. $s \in bottom(\triangle_{R\Sigma}(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_{R\Sigma}(\mathbf{T})$, i.e. $\triangle_{R\Sigma}(\mathbf{T})(s) = 0$. So $card(seg(f_{R\Sigma}(s)) = 0$, i.e. $f_{R\Sigma}(s)$ is the least element in $A_{R\Sigma}$. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $f_{R\Sigma}(s) \neq (0, 0, \ldots, 0)$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{R\Sigma}(s') = (0, 0, \ldots, 0)$, i.e. $f_{R\Sigma}(s')$ is the least element in A_{RMaj}. So $f_{R\Sigma}(s') = f_{R\Sigma}(s) = (0, 0, \ldots, 0)$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

$(\mathbf{TM} \triangle \mathbf{3})$ It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let $f_{R\Sigma}$ be the refined indexing function with respect to $\mathbf{T}$ and $f'_{R\Sigma}$ the refined indexing function with respect to $\mathbf{T}'$. From definition 6.19 it is clear that $f_{R\Sigma}(s) = f'_{R\Sigma}(s)$ for every $s \in S$. But then, $\triangle_{R\Sigma}(\mathbf{T}) = \triangle_{R\Sigma}(\mathbf{T}')$.

$(\mathbf{TM} \triangle \mathbf{4})$ It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{R\Sigma}(\mathbf{T})(s) \leq \triangle_{R\Sigma}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Let $f_{R\Sigma}(s) = (x_1, x_2, \ldots, x_n)$ and $f_{R\Sigma}(s') = (y_1, y_2, \ldots, y_n)$. But $x_j = \sum\{t_i(s) \mid t_i \in \mathbf{T} \text{ and } t_i(s) > j - 1\}$ and $y_j = \sum\{t_i(s') \mid t_i \in \mathbf{T} \text{ and } t_i(s') > j - 1\}$. Hence $x_1 \geq x_2 \geq \ldots \geq x_n$ and $y_1 \geq y_2 \geq \ldots \geq y_n$. But $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. So $x_1 \leq y_1$. If $x_1 = y_1$ then $x_2 \leq y_2$. And so on. But then $f_{R\Sigma}(s) \preceq f_{R\Sigma}(s')$. Hence, by lemma 6.2, it follows that

$\triangle_{R\Sigma}(\mathbf{T})(s) \leq \triangle_{R\Sigma}(\mathbf{T})(s')$.

(**TM** $\triangle$ **5**) It must be shown that if $\triangle_{R\Sigma}(\mathbf{T})(s) \leq \triangle_{R\Sigma}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{R\Sigma}(\mathbf{T})(s) \leq \triangle_{R\Sigma}(\mathbf{T})(s')$. Let $f_{R\Sigma}(s) = (x_1, x_2, \ldots, x_n)$ and $f_{R\Sigma}(s') = (y_1, y_2, \ldots, y_n)$. But $x_j = \sum\{t_i(s) \mid t_i \in \mathbf{T} \text{ and } t_i(s) > j-1\}$ and $y_j = \sum\{t_i(s') \mid t_i \in \mathbf{T} \text{ and } t_i(s') > j-1\}$. Hence $x_1 \geq x_2 \geq \ldots \geq x_n$ and $y_1 \geq y_2 \geq \ldots \geq y_n$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $y_1 < x_1$. So $f_{R\Sigma}(s') \prec f_{R\Sigma}(s)$. But then, by lemma 6.2, it follows that $\triangle_{R\Sigma}(\mathbf{T})(s') \leq \triangle_{R\Sigma}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. ■

Proposition D.13 (6.23) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *be an information tuple. Property* (**TM** $\triangle$ **6**) *is satisfied by templated merging operation* $\triangle_\Sigma$ *and* $\triangle_{R\Sigma}$ *but not satisfied by* $\triangle_{Min}$, $\triangle_{Max}$, $\triangle_{RMin}$, *and* $\triangle_{RMax}$.

Proof. Case 1: It must be shown that $\triangle_\Sigma$ and $\triangle_{R\Sigma}$ satisfy property (**TM** $\triangle$ **6**).

$\triangle_\Sigma$: It must be shown that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_\Sigma(\mathbf{T}_{t_j,k})(s) \leq \triangle_\Sigma(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$. Choose any $s, s' \in S$. If it holds that if $\triangle_\Sigma(\mathbf{T}_{t_j,1})(s) \leq \triangle_\Sigma(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) \leq t_j(s')$ then it will hold for every k. So let $k(s,s') = 1$. Otherwise, it holds that if $\triangle_\Sigma(\mathbf{T}_{t_j,1})(s) \leq \triangle_\Sigma(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) > t_j(s')$. Let $k(s,s') = (\triangle_\Sigma(\mathbf{T}_{t_j,k})(s') - \triangle_\Sigma(\mathbf{T}_{t_j,k})(s)) + 2$. But then $\triangle_\Sigma(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_\Sigma(\mathbf{T}_{t_j,k(s,s')})(s')$. So there exists a $k(s,s')$ such that, for $s, s' \in S$, it holds that if $\triangle_\Sigma(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_\Sigma(\mathbf{T}_{t_j,k(s,s')})(s')$ then $t_j(s) > t_j(s')$. By taking k to be the maximum of $k(s,s')$ for all $\forall s, s' \in S$, the result is obtained.

$\triangle_{R\Sigma}$: It must be shown that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{R\Sigma}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{R\Sigma}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$. Choose any $s, s' \in S$. If it holds that if $\triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) \leq t_j(s')$ then it will hold for every k. So let $k(s,s') = 1$. Otherwise, it holds that if $\triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) > t_j(s')$. Let $k(s,s') = (\triangle_{R\Sigma}(\mathbf{T}_{t_j,k})(s') - \triangle_{R\Sigma}(\mathbf{T}_{t_j,k})(s)) + 2$. But then $\triangle_{R\Sigma}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{R\Sigma}(\mathbf{T}_{t_j,k(s,s')})(s')$. So there exists a

$k(s, s')$ such that, for $s, s' \in S$, it holds that if $\triangle_{R\Sigma}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{R\Sigma}(\mathbf{T}_{t_j,k(s,s')})(s')$ then $t_j(s) > t_j(s')$. By taking k to be the maximum of $k(s, s')$ for all $\forall s, s' \in S$, the result is obtained.

Case 2: It must be shown that $\triangle_{Min}$, $\triangle_{Max}$, $\triangle_{RMin}$ and $\triangle_{RMax}$ do not satisfy property (**TM** $\triangle$ **6**).

$\triangle_{Min}$: Let $\mathbf{T} = (t_1, t_2)$ be an info tuple with $t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$ and $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$. Let $t_j = \{(11, 4), (10, 0), (01, 4), (00, 4)\}$. So $f_{Min}(11) = succ(\min\{t_i(11) \mid t_i \in \mathbf{T}_{t_j,k}\}) = 1$ with respect to $\mathbf{T}_{t_j,k}$ for every k and $f_{Min}(10) = succ(\min\{t_i(10) \mid t_i \in \mathbf{T}_{t_j,k}\}) = 1$ with respect to $\mathbf{T}_{t_j,k}$ for every k. Hence $f_{Min}(11) = f_{Min}(10)$ with respect to $\mathbf{T}_{t_j,k}$ for every k. But then, by lemma 6.1, $\triangle_{Min}(\mathbf{T}_{t_j,k})(11) = \triangle_{Min}(\mathbf{T}_{t_j,k})(10)$ for every k. But $t_j(11) = 4$ and $t_j(10) = 0$, i.e. $t_j(11) > t_j(10)$. So it does not hold that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{Min}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{Min}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$.

$\triangle_{Max}$: Let $\mathbf{T} = (t_1, t_2)$ be an info with $t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$ and $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$. Let $t_j = \{(11, 1), (10, 1), (01, 0), (00, 2)\}$. So $f_{Max}(00) = \max\{t_i(00) \mid t_i \in \mathbf{T}_{t_j,k}\} = 3$ with respect to $\mathbf{T}_{t_j,k}$ for every k and $f_{Max}(01) = \max\{t_i(01) \mid t_i \in \mathbf{T}_{t_j,k}\} = 4$ with respect to $\mathbf{T}_{t_j,k}$ for every k. Hence $f_{Max}(00) < f_{Max}(01)$ with respect to $\mathbf{T}_{t_j,k}$ for every k. But then, by lemma 6.1, $\triangle_{Max}(\mathbf{T}_{t_j,k})(00) < \triangle_{Max}(\mathbf{T}_{t_j,k})(01)$ for every k. But $t_j(00) = 2$ and $t_j(01) = 0$, i.e. $t_j(00) > t_j(01)$. So it does not hold that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{Max}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{Max}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$.

$\triangle_{RMin}$: Let $\mathbf{T} = (t_1, t_2)$ be an info tuple with $t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$ and $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$. Let $t_j = \{(11, 4), (10, 2), (01, 0), (00, 1)\}$. So $f_{RMin}(11) = (0, 0, 4)$ with respect to $\mathbf{T}_{t_j,1}$ and $f_{RMin}(10) = (0, 1, 2)$ with respect to $\mathbf{T}_{t_j,1}$. Since $t_j(11) = \max(f_{RMin}(11))$ and $t_j(10) = \max(f_{RMin}(10))$ it follows that $f_{RMin}(11) \prec f_{RMin}(10)$ with respect to $\mathbf{T}_{t_j,k}$ for every k. But $first(f_{RMin}(11)) < 4$ and thus, by lemma 6.2, $\triangle_{RMin}(\mathbf{T}_{t_j,k})(11) < \triangle_{RMin}(\mathbf{T}_{t_j,k})(10)$ for every k. But $t_j(11) = 4$ and $t_j(10) = 0$, i.e. $t_j(11) > t_j(10)$. So it does not hold that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{RMin}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{RMin}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$.

$\triangle_{RMax}$: Let $\mathbf{T} = (t_1, t_2)$ be an info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$ and $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$. Let $t_j = \{(11,4),(10,2),(01,0),(00,1)\}$. So $f_{RMax}(00) = (3,1,1)$ with respect to $\mathbf{T}_{t_j,1}$ and $f_{RMax}(01) = (4,2,0)$ with respect to $\mathbf{T}_{t_j,1}$. Since $t_j(00) = \min(f_{RMax}(00))$ and $t_j(01) = \min(f_{RMax}(01))$ it follows that $f_{RMax}(00) \prec f_{RMax}(01)$ with respect to $\mathbf{T}_{t_j,k}$ for every k. But $first(f_{RMax}(00)) < 4$ and thus, by lemma 6.2, $\triangle_{RMax}(\mathbf{T}_{t_j,k})(00) < \triangle_{RMax}(\mathbf{T}_{t_j,k})(01)$ for every k. But $t_j(00) = 1$ and $t_j(01) = 0$, i.e. $t_j(11) > t_j(10)$. So it does not hold that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{RMax}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{RMax}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$. ■

Proposition D.14 (6.24) *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Property* (**TM** $\triangle$ **7**) *is satisfied by templated merging operations $\triangle_{Min}$ and $\triangle_{Max}$, but not satisfied by $\triangle_{\Sigma}$, $\triangle_{RMin}$, $\triangle_{RMax}$, and $\triangle_{\Sigma}$.*

Proof. Case 1: It must be shown that $\triangle_{Min}$ and $\triangle_{Max}$ satisfy property (**TM** $\triangle$ **7**).

$\triangle_{Min}$: t must be shown that $\forall t_j \in T_E\ \forall k\ \triangle_{Min}(\mathbf{T}_{t_j,k}) = \triangle_{Min}(\mathbf{T}_{t_j,1})$. Let $f_{Min,1}(s)$ and $f_{Min,k}(s)$ be the indexing functions with respect to information tuples $\mathbf{T}_{t_j,1}$ and $\mathbf{T}_{t_j,k}$ respectively. Choose any $t_j \in T_E$. If $t_j(s) = t_i(s)$ for every $t_i \in \mathbf{T}$ then $f_{Min,1}(s) = t_1(s) = t_j(s)$. But then $f_{Min,k}(s) = t_1(s) = t_j(s)$ for every k, i.e. $f_{Min,1}(s) = f_{Min,k}(s)$ for every k. Otherwise $f_{Min,1}(s) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,1}\})$. But $\min\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,1}\} = \min\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,k}\}$ for every k and thus $succ(\min\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,1}\}) = succ(\min\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,k}\})$ for every k, i.e. $f_{Min,1}(s) = f_{Min,k}(s)$ for every k. So for every $s \in S$, $f_{Min,1}(s) = f_{Min,k}(s)$ for every k. But then $_{Min}(\mathbf{T}_{t_j,k}) = \triangle_{Min}(\mathbf{T}_{t_j,1})$ for every k. Since t_j was chosen arbitrarily, it follows that $\triangle_{Min}$ satisfies property (**TM** $\triangle$ **7**).

$\triangle_{Max}$: It must be shown that $\forall t_j \in T_E\ \forall k\ \triangle_{Max}(\mathbf{T}_{t_j,k}) = \triangle_{Max}(\mathbf{T}_{t_j,1})$ for $1 \leq k \leq n+1$. Let $f_{Max,1}(s)$ and $f_{Max,k}(s)$ be the indexing functions with respect to information tuples $\mathbf{T}_{t_j,1}$ and $\mathbf{T}_{t_j,k}$ respectively. Choose any $t_j \in T_E$. Now $f_{Max,1}(s) = \max\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,1}\}$. But $\max\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,1}\} = \max\{t_i(s) \mid t_i \in \mathbf{T}_{t_j,k}\}$ for every k and thus $f_{Max,1}(s) = f_{Max,k}(s)$ for every k. So for every $s \in S$, $f_{Max,1}(s) = f_{Max,k}(s)$ for every k. But then $\triangle_{Max}(\mathbf{T}_{t_j,k}) = \triangle_{Max}(\mathbf{T}_{t_j,1})$ for every k. Since t_j was chosen arbitrarily, it follows that $\triangle_{Max}$ satisfies property ($\mathbf{TM} \triangle \mathbf{7}$).

Case 2: It must be shown that $\triangle_\Sigma$, $\triangle_{RMin}$, $\triangle_{RMax}$, and $\triangle_{RMaj}$ do not satisfy property ($\mathbf{TM} \triangle \mathbf{7}$).

$\triangle_\Sigma$: Let $\mathbf{T} = (t_1, t_2)$ be an info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$ and $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$. Let $t_j = \{(11,1),(10,2),(01,2),(00,0)\}$. So $f_\Sigma(11) = \sum t_i(11) \mid t_i \in \mathbf{T}_{t_j,1} = 1$ with respect to $\mathbf{T}_{t_j,1}$ and $f_\Sigma(00) = \sum t_i(00) \mid t_i \in \mathbf{T}_{t_j,1} = 4$ with respect to $\mathbf{T}_{t_j,1}$. Hence $f_\Sigma(11) < f_\Sigma(00)$ with respect to $\mathbf{T}_{t_j,1}$. But then, by lemma 6.1, $\triangle_\Sigma(\mathbf{T}_{t_j,1})(11) < \triangle_\Sigma(\mathbf{T}_{t_j,1})(00)$. For $k = 5$ it holds that $f_\Sigma(11) = \sum t_i(11) \mid t_i \in \mathbf{T}_{t_j,5} = 5$ with respect to $\mathbf{T}_{t_j,5}$ and $f_\Sigma(00) = \sum t_i(00) \mid t_i \in \mathbf{T}_{t_j,5} = 4$ with respect to $\mathbf{T}_{t_j,5}$. Hence $f_\Sigma(00) < f_\Sigma(11)$ with respect to $\mathbf{T}_{t_j,5}$. But then, by lemma 6.1, $\triangle_\Sigma(\mathbf{T}_{t_j,5})(00) < \triangle_\Sigma(\mathbf{T}_{t_j,5})(11)$. So for $k = 5$ it holds that $\triangle_\Sigma(\mathbf{T}_{t_j,5}) \neq \triangle_\Sigma(\mathbf{T}_{t_j,1})$.

$\triangle_{RMin}$: Let $\mathbf{T} = (t_1, t_2, t_3)$ be the info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$, $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$, and $t_3 = \{(11,0),(10,2),(01,4),(00,1)\}$. Let $t_j = \{(11,1),(10,2),(01,2),(00,0)\}$. Since $f_{RMin}(00) = (0,1,1,3)$ and $f_{RMin}(11) = (0,0,0,1)$ with respect to $\mathbf{T}_{t_j,1}$ it follows that $f_{RMin}(11) \prec f_{RMin}(00)$. But $first(f_{RMin}(11)) < 4$ and thus, by lemma 6.2, $\triangle_{RMin}(\mathbf{T}_{t_j,1})(11) < \triangle_{RMin}(\mathbf{T}_{t_j,1})(00)$. For $k = 4$ it is the case that $f_{RMin}(00) = (0,0,0,0,1,1,3)$ and $f_{RMin}(11) = (0,0,0,1,1,1,1)$ with respect to $\mathbf{T}_{t_j,4}$. But then $f_{RMin}(00) \prec f_{RMin}(11)$ and hence, since $first(f_{RMin}(00)) < 4$, it follows by lemma 6.2 that $\triangle_{RMin}(\mathbf{T}_{t_j,4})(00) < \triangle_{RMin}(\mathbf{T}_{t_j,4})(11)$. So for $k = 4$ it holds that $\triangle_{RMin}(\mathbf{T}_{t_j,4}) \neq \triangle_{RMin}(\mathbf{T}_{t_j,1})$.

$\triangle_{RMax}$: Let $\mathbf{T} = (t_1, t_2, t_3)$ be the info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$, $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$, and $t_3 = \{(11,0),(10,2),(01,4),(00,1)\}$. Let $t_j =$

$\{(11,3),(10,2),(01,1),(00,0)\}$. Since $f_{RMax}(00) = (3,1,1,0)$ and $f_{RMax}(11) = (3,0,0,0)$ with respect to $\mathbf{T}_{t_j,1}$ it follows that $f_{RMax}(11) \prec f_{RMax}(00)$. But $first(f_{RMin}(11)) < 4$ and thus, by lemma 6.2, $\triangle_{RMax}(\mathbf{T}_{t_j,1})(11) < \triangle_{RMax}(\mathbf{T}_{t_j,1})(00)$. For $k = 2$ it is the case that $f_{RMax}(00) = (3,1,1,0,0)$ and $f_{RMax}(11) = (3,3,0,0,0)$ with respect to $\mathbf{T}_{t_j,2}$. But then $f_{RMax}(00) \prec f_{RMax}(11)$ and hence, since $first(f_{RMax}(00)) < 4$, it follows by lemma 6.2 that $\triangle_{RMax}(\mathbf{T}_{t_j,2})(00) < \triangle_{RMax}(\mathbf{T}_{t_j,2})(11)$. So for $k = 2$ it holds that $\triangle_{RMax}(\mathbf{T}_{t_j,2}) \neq \triangle_{RMax}(\mathbf{T}_{t_j,1})$.

$\triangle_{R\Sigma}$: Let $\mathbf{T} = (t_1, t_2, t_3)$ be the info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$, $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$, and $t_3 = \{(11,0),(10,2),(01,4),(00,1)\}$. Let $t_j = \{(11,4),(10,1),(01,2),(00,0)\}$. Since $f_{R\Sigma}(11) = (4,4,4,4)$ and $f_{R\Sigma}(00) = (5,3,3,0)$ with respect to $\mathbf{T}_{t_j,1}$ it follows that $f_{R\Sigma}(11) \prec f_{R\Sigma}(00)$. But $first(f_{R\Sigma}(11)) < 16$ with respect to $\mathbf{T}_{t_j,1}$ and thus, by lemma 6.2, $\triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(11) < \triangle_{R\Sigma}(\mathbf{T}_{t_j,1})(00)$. For $k = 2$ it is the case that $f_{R\Sigma}(11) = (8,8,8,8)$ and $f_{R\Sigma}(00) = (5,3,3,0)$ with respect to $\mathbf{T}_{t_j,2}$. So $f_{R\Sigma}(00) \prec f_{R\Sigma}(11)$. But $first(f_{R\Sigma}(00)) < 20$ with respect to $\mathbf{T}_{t_j,2}$ and thus, by lemma 6.2, $\triangle_{R\Sigma}(\mathbf{T}_{t_j,2})(00) < \triangle_{R\Sigma}(\mathbf{T}_{t_j,2})(11)$. So for $k = 2$ it holds that $\triangle_{R\Sigma}(\mathbf{T}_{t_j,2}) \neq \triangle_{R\Sigma}(\mathbf{T}_{t_j,1})$. ■

D.3 Proofs for Section 6.6

Proposition D.15 (6.25) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{Cont}$ *satisfies properties* $(\mathbf{TM}\triangle\mathbf{1})$ *to* $(\mathbf{TM}\triangle\mathbf{6})$ *but fails to satisfy property* $(\mathbf{TM}\triangle\mathbf{7})$*.*

Proof. Templated merging operation $\triangle_{Cont}$ is based on indexing function $f_{Cont} : S \to \{0, 1, \ldots, m\}$.

$(\mathbf{TM}\triangle\mathbf{1})$ It must be shown that $bottom(\triangle_{Cont}(\mathbf{T})) \neq \varnothing$. By proposition 6.15, $\triangle_{Cont}(\mathbf{T})$ is regular. But then $bottom(\triangle_{Cont}(\mathbf{T})) \neq \varnothing$.

($\mathbf{TM} \triangle \mathbf{2}$) It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{Cont}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) = 0$, i.e. $f_{Cont}(s) = 0$ and hence $\triangle_{Cont}(\mathbf{T})(s) = card(seg(0))$. Since $card(seg(0)) = 0$ it follows that $s \in bottom(\triangle_{Cont}(\mathbf{T}))$. Conversely, choose any $s \in bottom(\triangle_{Cont}(\mathbf{T})$, i.e. $\triangle_{Cont}(\mathbf{T})(s) = 0$. So $card(seg(f_{Cont}(s)) = 0$, i.e. $f_{Cont}(s)$ is the least element in A_{Cont}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) > 0$, i.e. $f_{Cont}(s) > 0$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{Cont}(s') = 0$, i.e. $f_{Cont}(s')$ is the least element in A_{Min}. So $f_{Cont}(s') = f_{Cont}(s) = 0$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

($\mathbf{TM} \triangle \mathbf{3}$) It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let $f_{Cont} : S \rightarrow \{0, 1, \ldots, m\}$ be the content-based indexing function with respect to $\mathbf{T}$ and $f'_{Cont} : S \rightarrow \{0, 1, \ldots, m\}$ the content-based indexing function with respect to $\mathbf{T}'$. From definition 6.22 it is clear that $f_{Cont}(s) = f'_{Cont}(s)$ for every $s \in S$. But then, $\triangle_{Cont}(\mathbf{T}) = \triangle_{Cont}(\mathbf{T}')$.

($\mathbf{TM} \triangle \mathbf{4}$) It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{Cont}(\mathbf{T})(s) \leq \triangle_{Cont}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. But then $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) \leq card(\{t_i \in \mathbf{T} \mid t_i(s') > 0\})$, i.e. $f_{Cont}(s) \leq f_{Cont}(s')$. But then, by lemma 6.1, it follows that $\triangle_{Cont}(\mathbf{T})(s) \leq \triangle_{Cont}(\mathbf{T})(s')$.

($\mathbf{TM} \triangle \mathbf{5}$) It must be shown that if $\triangle_{Cont}(\mathbf{T})(s) \leq \triangle_{Cont}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{Cont}(\mathbf{T})(s) \leq \triangle_{Cont}(\mathbf{T})(s')$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $card(\{t_i \in \mathbf{T} \mid t_i(s') > 0\})\} < card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\})$, i.e. $f_{Cont}(s') < f_{Cont}(s)$. But then, by lemma 6.1, it follows that $\triangle_{Cont}(\mathbf{T})(s') < \triangle_{Cont}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$.

($\mathbf{TM} \triangle \mathbf{6}$) It must be shown that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{Cont}(\mathbf{T}_{t_j,k})(s)$

$\leq \triangle_{Cont}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$. Choose any $s, s' \in S$. If it holds that if $\triangle_{Cont}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{Cont}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) \leq t_j(s')$ then it will hold for every k. So let $k(s,s') = 1$. Otherwise, it holds that if $\triangle_{Cont}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{Cont}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) > t_j(s')$. Let $k(s,s') = (\triangle_{Cont}(\mathbf{T}_{t_j,k})(s') - \triangle_{Cont}(\mathbf{T}_{t_j,k})(s)) + 2$. But then $\triangle_{Cont}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{Cont}(\mathbf{T}_{t_j,k(s,s')})(s')$. So there exists a $k(s,s')$ such that, for $s, s' \in S$, it holds that if $\triangle_{Cont}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{Cont}(\mathbf{T}_{t_j,k(s,s')})(s')$ then $t_j(s) > t_j(s')$. By taking k to be maximum of $k(s,s')$ for all $\forall s, s' \in S$, the result is obtained.

(**TM** $\triangle$ **7**) It must be shown that it is not the case that that $\forall t_j \in T_E \; \forall k \; \triangle_{Cont}(\mathbf{T}_{t_j,k}) = \triangle_{Cont}(\mathbf{T}_{t_j,1})$. Let $\mathbf{T} = (t_1, t_2)$ be an info tuple with $t_1 = \{(11,0),(10,1),(01,2),(00,3)\}$ and $t_2 = \{(11,0),(10,0),(01,4),(00,1)\}$. Let $t_j = \{(11,1),(10,2),(01,2),(00,0)\}$. So $f_{Cont}(11) = card(\{t_i \in \mathbf{T}_{t_j,1} \mid t_i(11) > 0\} = 1$ with respect to $\mathbf{T}_{t_j,1}$ and $f_{Cont}(00) = card(\{t_i \in \mathbf{T}_{t_j,1} \mid t_i(00) > 0\} = 2$ with respect to $\mathbf{T}_{t_j,1}$. Hence $f_{Cont}(11) < f_{Cont}(00)$ with respect to $\mathbf{T}_{t_j,1}$. But then, by lemma 6.1, $\triangle_{Cont}(\mathbf{T}_{t_j,1})(11) < \triangle_{Cont}(\mathbf{T}_{t_j,1})(00)$. For $k = 5$ it holds that $f_{Cont}(11) = card(\{t_i \in \mathbf{T}_{t_j,5} \mid t_i(11) > 0\} = 5$ with respect to $\mathbf{T}_{t_j,5}$ and $f_{Cont}(00) = card(\{t_i \in \mathbf{T}_{t_j,5} \mid t_i(00) > 0\} = 2$ with respect to $\mathbf{T}_{t_j,5}$. Hence $f_{Cont}(11) > f_{Cont}(00)$ with respect to $\mathbf{T}_{t_j,5}$. But then, by lemma 6.1, $\triangle_{Cont}(\mathbf{T}_{t_j,5})(11) > \triangle_{Cont}(\mathbf{T}_{t_j,5})(00)$. So for $k = 5$ it holds that $\triangle_{Cont}(\mathbf{T}_{t_j,5}) \neq \triangle_{Cont}(\mathbf{T}_{t_j,1})$. ∎

Proposition D.16 (6.27) *Let* $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ *and* $\mathbf{T}' = (t'_1, t'_2, \ldots, t'_m)$ *be information tuples. Templated merging operation* $\triangle_{RCont}$ *satisfies properties* (**TM** $\triangle$ **1**) *to* (**TM** $\triangle$ **6**) *but fails to satisfy property* (**TM** $\triangle$ **7**).

Proof. Templated merging operation $\triangle_{RCont}$ is based on indexing function $f_{RCont} : S \to \{0, 1, \ldots, m\}^n$.

(**TM** $\triangle$ **1**) It must be shown that $bottom(\triangle_{RCont}(\mathbf{T})) \neq \varnothing$. By proposition 6.19, $\triangle_{RCont}(\mathbf{T})$ is regular. But then $bottom(\triangle_{RCont}(\mathbf{T})) \neq \varnothing$.

(**TM** $\triangle$ **2**) It must be shown that if $\mathbf{T}$ is satisfiable, then $bottom(\triangle_{RCont}(\mathbf{T})) = \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. Assume that $\mathbf{T}$ is satisfiable, i.e. $\bigcap_{t_i \in \mathbf{T}} bottom(t_i) \neq \varnothing$. Choose any

$s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RCont}(s) = (0, 0, \ldots, 0)$. So $card(seg(f_{RCont}(s))) = 0$, i.e. $s \in bottom(\triangle_{RCont}(\mathbf{T})$. Conversely, choose any $s \in bottom(\triangle_{RCont}(\mathbf{T})$, i.e. $\triangle_{RCont}(\mathbf{T})(s) = 0$. So $card(seg(f_{RCont}(s)) = 0$, i.e. $f_{RCont}(s)$ is the least element in A_{RCont}. Suppose that $s \notin bottom(t_i)$ for some $t_i \in \mathbf{T}$. So $f_{RCont}(s) \neq (0, 0, \ldots, 0)$. But $\mathbf{T}$ is satisfiable and hence there must be some $s' \neq s$ such that $t_i(s') = 0$ for every $t_i \in \mathbf{T}$. But then $f_{RCont}(s') = (0, 0, \ldots, 0)$, i.e. $f_{RCont}(s')$ is the least element in A_{RCont}. So $f_{RCont}(s') = f_{RCont}(s) = (0, 0, \ldots, 0)$. Contradiction. So $s \in bottom(t_i)$ for every $t_i \in \mathbf{T}$, i.e. $s \in \bigcap_{t_i \in \mathbf{T}} bottom(t_i)$.

(**TM $\triangle$ 3**) It must be shown that if $\mathbf{T} \approx \mathbf{T}'$ then $\triangle(\mathbf{T}) = \triangle(\mathbf{T}')$. Assume that $\mathbf{T} \approx \mathbf{T}'$. Let f_{RCont} be the refined content-based indexing function with respect to $\mathbf{T}$ and f_{RCont} the refined content-based indexing function with respect to $\mathbf{T}'$. From definition 6.19 it is clear that $f_{RCont}(s) = f_{RCont}(s)$ for every $s \in S$. But then, $\triangle_{RCont}(\mathbf{T}) = \triangle_{RCont}(\mathbf{T}')$.

(**TM $\triangle$ 4**) It must be shown that if $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$ then $\triangle_{RCont}(\mathbf{T})(s) \leq \triangle_{RCont}(\mathbf{T})(s')$. Choose any $s, s' \in S$ such that $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. Let $f_{RCont}(s) = (x_1, x_2, \ldots, x_n)$ and $f_{RCont}(s') = (y_1, y_2, \ldots, y_n)$. But $x_j = card(\{t_i \in \mathbf{T} \mid t_i(s) > j - 1\})$ and $y_j = card(\{t_i \in \mathbf{T} \mid t_i(s') > j - 1\})$. Hence $x_1 \geq x_2 \geq \ldots \geq x_n$ and $y_1 \geq y_2 \geq \ldots \geq y_n$. But $t_i(s) \leq t_i(s')$ for every $t_i \in \mathbf{T}$. So $x_1 \leq y_1$. If $x_1 = y_1$ then $x_2 \leq y_2$. And so on. But then $f_{RCont}(s) \preceq f_{RCont}(s')$. Hence, by lemma 6.2, it follows that $\triangle_{RCont}(\mathbf{T})(s) \leq \triangle_{RCont}(\mathbf{T})(s')$.

(**TM $\triangle$ 5**) It must be shown that if $\triangle_{RCont}(\mathbf{T})(s) \leq \triangle_{RCont}(\mathbf{T})(s')$ then $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$. Assume that $\triangle_{RCont}(\mathbf{T})(s) \leq \triangle_{RCont}(\mathbf{T})(s')$. Let $f_{RCont}(s) = (x_1, x_2, \ldots, x_n)$ and $f_{RCont}(s') = (y_1, y_2, \ldots, y_n)$. But $x_j = card(\{t_i \in \mathbf{T} \mid t_i(s) > j - 1\})$ and $y_j = card(\{t_i \in \mathbf{T} \mid t_i(s') > j - 1\})$. Hence $x_1 \geq x_2 \geq \ldots \geq x_n$ and $y_1 \geq y_2 \geq \ldots \geq y_n$. Suppose that $t_i(s') < t_i(s)$ for every $t_i \in \mathbf{T}$. But then $y_1 < x_1$. So $f_{RCont}(s') \prec f_{RCont}(s)$. But then, by lemma 6.2, it follows that $\triangle_{RCont}(\mathbf{T})(s') \leq$

$\triangle_{RCont}(\mathbf{T})(s)$. Contradiction. So $t_i(s) \leq t_i(s')$ for some $t_i \in \mathbf{T}$.

(**TM** $\triangle$ **6**) It must be shown that $\exists k$ s.t. $\forall s, s' \in S$, if $\triangle_{RCont}(\mathbf{T}_{t_j,k})(s) \leq \triangle_{RCont}(\mathbf{T}_{t_j,k})(s')$ then $t_j(s) \leq t_j(s')$. Choose any $s, s' \in S$. If it holds that if $\triangle_{RCont}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{RCont}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) \leq t_j(s')$ then it will hold for every k. So let $k(s, s') = 1$. Otherwise, it holds that if $\triangle_{RCont}(\mathbf{T}_{t_j,1})(s) \leq \triangle_{RCont}(\mathbf{T}_{t_j,1})(s')$ then $t_j(s) > t_j(s')$. Let $k(s, s') = (\triangle_{RCont}(\mathbf{T}_{t_j,k})(s') - \triangle_{RCont}(\mathbf{T}_{t_j,k})(s)) + 2$. But then $\triangle_{RCont}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{RCont}(\mathbf{T}_{t_j,k(s,s')})(s')$. So there exists a $k(s, s')$ such that, for $s, s' \in S$, it holds that if $\triangle_{RCont}(\mathbf{T}_{t_j,k(s,s')})(s) > \triangle_{RCont}(\mathbf{T}_{t_j,k(s,s')})(s')$ then $t_j(s) > t_j(s')$. By taking k to be the maximum of $k(s, s')$ for all $\forall s, s' \in S$, the result is obtained.

(**TM** $\triangle$ **7**) Let $\mathbf{T} = (t_1, t_2, t_3)$ be the info tuple with $t_1 = \{(11, 0), (10, 1), (01, 2), (00, 3)\}$, $t_2 = \{(11, 0), (10, 0), (01, 4), (00, 1)\}$, and $t_3 = \{(11, 0), (10, 2), (01, 4), (00, 1)\}$. Let $t_j = \{(11, 4), (10, 1), (01, 2), (00, 0)\}$. Since $f_{RCont}(11) = (1, 1, 1, 1)$ and $f_{RCont}(00) = (3, 1, 1, 0)$ with respect to $\mathbf{T}_{t_j,1}$ it follows that $f_{RCont}(11) \prec f_{RCont}(00)$. But $first(f_{RCont}(11)) < 4$ with respect to $\mathbf{T}_{t_j,1}$ and thus, by lemma 6.2, $\triangle_{RCont}(\mathbf{T}_{t_j,1})(11) < \triangle_{RCont}(\mathbf{T}_{t_j,1})(00)$. For $k = 4$ it is the case that $f_{RCont}(11) = (4, 4, 4, 4)$ and $f_{RCont}(11) = (3, 1, 1, 0)$ with respect to $\mathbf{T}_{t_j,4}$. So $f_{RCont}(00) \prec f_{RCont}(11)$. But $first(f_{RCont}(00)) < 4$ with respect to $\mathbf{T}_{t_j,4}$ and thus, by lemma 6.2, $\triangle_{RCont}(\mathbf{T}_{t_j,4})(00) < \triangle_{RCont}(\mathbf{T}_{t_j,4})(11)$. So for $k = 4$ it holds that $\triangle_{RCont}(\mathbf{T}_{t_j,4}) \neq \triangle_{RCont}(\mathbf{T}_{t_j,1})$. ■

Proposition D.17 (6.28) *Let $\mathbf{T} = (t_1, t_2, \ldots, t_m)$ be an information tuple. Then the following holds for each $s \in S$:*

1. *$s \in Cont_D(\triangle_{Cont}(\mathbf{T}))$ iff $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$*
2. *if $s \in Cont_0(\triangle_{Cont}(\mathbf{T}))$ then $s \in Cont_0(t_i)$ for some $t_i \in \mathbf{T}$*
3. *$s \in Cont_D(\triangle_{RCont}(\mathbf{T}))$ iff $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$*
4. *if $s \in Cont_0(\triangle_{RCont}(\mathbf{T}))$ then $s \in Cont_0(t_i)$ for some $t_i \in \mathbf{T}$*

Proof. 1. Choose any $s \in Cont_D(\triangle_{Cont}(\mathbf{T}))$. Suppose there is some $t_i \in \mathbf{T}$, say t_k, such that $s \notin Cont_0(t_k)$. Since $Cont_D(\triangle_{Cont}(\mathbf{T})) = top(\triangle_{Cont}(\mathbf{T}))$ it follows that $\triangle_{Cont}(\mathbf{T})(s) = n$. So $f_{Cont}(s) = m$ and $card(A_{Cont}) \neq 1$. But $f_{Cont}(s) = card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\})$ and thus $t_k(s) > 0$. But $Cont_0(t_k) = get(t_k, 1, n)$, i.e. $s \in Cont_0(t_k)$. Contradiction. So $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$. Conversely choose any $s \in S$ such that $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$. But then $t_i(s) > 0$ for every $t_i \in \mathbf{T}$, i.e. $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) = m$, i.e. $f_{Cont}(s) = m$. Now $card(A_{Cont}) \neq 1$ because otherwise $f_{Cont}(s') = m$ for every $s' \in S$, i.e. no $t_i \in \mathbf{T}$ would be regular. So $\triangle_{Cont}(\mathbf{T})(s) = n$, i.e. $s \in top(\triangle_{Cont}(\mathbf{T}))$, i.e. $s \in Cont_D(\triangle_{Cont}(\mathbf{T}))$.

2. Choose any $s \in Cont_0(\triangle_{Cont}(\mathbf{T}))$. Suppose for every $t_i \in \mathbf{T}$, it holds that $s \notin Cont_0(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) = 0$, i.e. $f_{Cont}(s) = 0$. Since $card(seg(f_{Cont}(s))) = 0$ it follows that $\triangle_{Cont}(\mathbf{T})(s) = 0$. But then $s \notin Cont_0(\triangle_{Cont}(\mathbf{T}))$. Contradiction. So there must be some $t_i \in \mathbf{T}$ such that $s \in Cont_0(t_i)$.

3. Choose any $s \in Cont_D(\triangle_{RCont}(\mathbf{T}))$. Suppose there is some $t_i \in \mathbf{T}$, say t_k, such that $s \notin Cont_0(t_k)$. Since $Cont_D(\triangle_{RCont}(\mathbf{T})) = top(\triangle_{RCont}(\mathbf{T}))$ it follows that $\triangle_{RCont}(\mathbf{T})(s) = n$. So $first(f_{RCont}(s)) = m$ and $first(\min(A_{RCont})) \neq m$. But $first(f_{RCont}(s)) = card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\})$. So $t_k(s) > 0$. But $Cont_0(t_k) = get(t_k, 1, n)$, i.e. $s \in Cont_0(t_k)$. Contradiction. So $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$. Conversely choose any $s \in S$ such that $s \in Cont_0(t_i)$ for every $t_i \in \mathbf{T}$. But then $t_i(s) > 0$ for every $t_i \in \mathbf{T}$, i.e. $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) = m$, i.e. $first(f_{RCont}(s)) = m$. Now $first(\min(A_{RCont})) \neq m$ because otherwise $first(f_{RCont}(s')) = m$ for every $s' \in S$, i.e. no $t_i \in \mathbf{T}$ would be regular. So $\triangle_{RCont}(\mathbf{T})(s) = n$, i.e. $s \in top(\triangle_{RCont}(\mathbf{T}))$, i.e. $s \in Cont_D(\triangle_{RCont}(\mathbf{T}))$.

4. Choose any $s \in Cont_0(\triangle_{RCont}(\mathbf{T}))$. Suppose for every $t_i \in \mathbf{T}$, it holds that $s \notin Cont_0(t_i)$. So $t_i(s) = 0$ for every $t_i \in \mathbf{T}$. But then $card(\{t_i \in \mathbf{T} \mid t_i(s) > 0\}) = 0$,

i.e. $f_{RCont}(s) = (0, 0, \ldots, 0)$. So $first(f_{RCont}(s)) = 0$. Since $card(seg(f_{RCont}(s))) = 0$ it follows that $\triangle_{RCont}(\mathbf{T})(s) = 0$. But then $s \notin Cont_0(\triangle_{RCont}(\mathbf{T}))$. Contradiction. So there must be some $t_i \in \mathbf{T}$ such that $s \in Cont_0(t_i)$. $\blacksquare$

Bibliography

Aho, A.V. & Ullman, J.D. (1992) *Foundations of Computer Science.* Computer Science Press.

Alchourrón, C.E., Gärdenfors, P., & Makinson, D. (1985) On the logic of theory change: Partial meet functions for contractions and revision. *Journal of Symbolic Logic*, 50, pp.510-530.

Alchourrón, C.E. & Makinson, D. (1981) Hierarchies of regulations and their logic. In: Hilpinen, R. ed. *New studies in deontic logic*. Dordrecht, D. Reidel, pp.125-148.

Alchourrón, C.E. & Makinson, D. (1982) On the logic of theory change: Contraction functions and their associated revision functions. *Theoria*, 48, pp.14-37.

Alchourrón, C.E. & Makinson, D. (1985) On the logic of theory change: Safe contraction. *Studia Logica*, 44, pp.405-422.

Andréka, H., Ryan, M., & Schobbens, P-Y. (2002) Operators and Laws for Combining Preference Relations, *Journal of Logic and Computation*, 12(1), pp.13-53.

Åqvist, L. (1984) Deontic Logic. In: Gabbay, D.M. & Guenthner, F. eds. *Handbook of Philosophical Logic*, Volume 2: Extensions of Classical Logic. Dordrecht, D. Reidel, pp.605-714.

Åqvist, L. (2002) Deontic Logic. In: Gabbay, D.M. & Guenthner, F. eds. *Handbook of Philosophical Logic, 2nd Edition,* Volume 8. Kluwer Academic Publishers, pp.147-264.

Arrow, K.J. (1951) *Social Choice and Individual Values.* 1st Edition. New York, Wiley.

Arrow, K.J. (1959) Rational choice functions and orderings. *Economica*, 26, pp.121-127.

Arrow, K.J. (1963) *Social Choice and Individual Values.* 2nd Edition. New York, Wiley.

Arrow, K.J., Sen, A.K., & Suzumura, K. eds. (2002) *Handbook of Social Choice and Welfare*, Volume 1. Handbooks in Economics, 19. North-Holland, Elsevier.

Aumann, R.J. (1976) Agreeing to disagree. *Annals of Statistics*, 4(6), pp.1236-1239.

Austin, J.L. (1962) *How to do things with words.* New York, Oxford University Press.

Baral, C., Kraus, S., & Minker, J. (1991) Combining Multiple Knowledge Bases. *IEEE Transactions on Knowledge and Data Engineering*, 3(2), pp.208-220.

Baral, C., Kraus, S., Minker, J., & Subrahmanian, V.S. (1992) Combining knowledge bases consisting of first-order theories. *Computational Intelligence*, 8(1), pp.45-71.

Bar-Hillel, Y & Carnap, R. (1953) Semantic information. *British Journal for the Philosophy of Science*, 4, pp.147-157.

Bar-Hillel, Y. (1955) An Examination of Information Theory. *Philosophy of Science*, 22, pp.86-105.

Reprinted as Chapter 16 in: (1964) Bar-Hillel, Y. ed. *Language and information*, Addison-Wesley.

Ben-Naim, J. (2005) Preferential and preferential-discriminative consequence relations. *Journal of Logic and Computation*, 15(3), pp.263-294.

Benferhat, S., Kaci, S., Le Berre, D. & Williams, M-A. (2004) Weakening conflicting information for iterated revision and knowledge integration. *Artificial Intelligence*, 153, pp.339-371.

A preliminary version appears in: (2001) Nebel, B. eds. *Proceedings of the 17th Inter-*

national Joint Conference on Artificial Intelligence (IJCAI-01), 4-10 August, Seattle, Washington, USA. Morgan Kaufmann, pp.109-115.

Benferhat, S., Dubois, D., Kaci, S. & Prade, H. (2002) Possibilistic merging and distance-based fusion of propositional information. *Annals of Mathematics and Artificial Intelligence*, 34(1-3), pp.217-252.
A preliminary version appears in: (2000) Horn, W. eds. *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI'00)*, 20-25 August, Berlin, Germany. IOS Press, pp.3-7, under the title "Encoding information fusion in possibilistic logic: a general framework for rational syntactic merging.".

Benferhat, S., Dubois, D., Kaci, S. & Prade, H. (2006) Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. *Information Fusion*, 7, pp.135-150.
A preliminary version appears in: (2002) Fensel, D., Guinchiglia, F., & McGuinness, D. eds. *Proceedings of the 8th International Conference on Principles of Knowledge Representation and Reasoning (KR 2002)*, 22-25 April, Toulouse, France. Morgan Kaufmann, pp.421-432, under the title "Bipolar representation and fusion of preferences in the possibilistic logic framework.".

Black, D. (1948) On the rationale of group decision-making. *Journal of Political Economy*, 56, pp.23-34.

Black, D. (1958) *The Theory of Committees and Elections*. Cambridge, Cambridge University Press.

Bochman, A. (2001) *A Logical Theory of Nonmonotonic Inference and Belief Change*. Berlin, Springer.

Borda (J.-C. de Borda) (1781) Mémoire sur les élections par scrutin. In: *Mémoires de l'Académie Royale des Sciences année*. pp.657-665. Translated in English by de Grazia,

A. (1953) Mathematical derivation of an election system, *Isys*, 44, pp.42-51.

Borgida, A. (1985) Language features for flexible handling of exception in information systems. *ACM Transactions in Database Systems*, 10, pp.563-603.

Borgida, A. (1988) *Updates in propositional databases.* Technical Report DCS-TR-222. Department of Computer Science, Rutgers University.

Borgida, A. & Imielinski, T. (1984) Decision making in committees: A framework for dealing with inconsistencies and non-monotonicity. In: *Proceedings of the 1st Non-Monotonic Reasoning Workshop*, 17-19 October, New Paltz, New York, USA. AAAI, pp.21-32.

Boutilier, C. (1993) Revision sequences and nested conditionals. In: Bajcsy, R. ed. *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, 28 August-3 September, Chambéry, France. Morgan Kaufmann, pp.519-525.

Boutilier, C. (1996a) Iterated revision and minimal revision of conditional beliefs. *Journal of Philosophical Logic*, 25(3), pp.262-305.

Boutilier, C. (1996b) Abduction to plausible causes: and event-based model of belief update. *Artificial Intelligence*, 83, pp.143-166.

Boutilier, C. (1998) A unified model of qualitative belief change: a dynamical systems perspective. *Artificial Intelligence*, 98(1-2), pp.281-316.
A preliminary version appears in: (1995) Mellish, C.S. ed. *Proceeding of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, 20-25 August 1995, Montreal, Quebec. San Mateo, CA, Morgan Kaufmann, pp.1550-1556, under the title "Generalized update: Belief change in dynamic settings.".

Bradshaw, J.M. ed. (1997) *Software Agents.* Menlo Park, CA, AAAI/MIT Press.

Bratman, M.E. (1987) *Intentions, Plans, and Practical Reasoning.* Cambridge, MA, Harvard University Press.

Brink, C. & Heidema, J. (1989) *A Versimilar Ordering of Propositional Theories: the Infinite Case.* Technical Report 4/89. Department of Mathematics, Rand Afrikaans University.

Brink, C. & Heidema, J. (1987) A versimilar ordering of theories phrased in a propositional language. *The British Journal for the Philosophy of Science*, 38, pp.533-549.

Burger, I.C. & Heidema, J. (1995) *Generalized Lexicographical Orderings.* Research Report 206/95 (12). Department of Mathematics, Applied Mathematics and Astronomy, University of South Africa.

Cadoli, M., Donini, F.M., Liberatore, P., & Schaerf, M. (1999) The size of a revised knowledge base. *Artificial Intelligence*,115, pp.25-64.
An earlier version appears in: (1995) *Proceedings of the 14th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Databases (PODS'95),* 22-25 May, San Jose, California, USA. ACM Press, pp.151-162.

Campbell, D.E. & Kelly, J.S. (2002) Impossibility Theorems in the Arrovian Framework. In: Arrow, K.J., Sen, A.K., & Suzumura, K. eds. *Handbook of Social Choice and Welfare,* Volume 1. Handbooks in Economics, 19. North-Holland, Elsevier, pp.35-94.

Cantwell, J. (1999) Some logics for iterated belief change. *Studia Logica*, 63, pp.49-84.

Carnap, R. (1950) *Logical Foundations of Probability.* University of Chicago Press.

Carnap, R. & Bar-Hillel, Y. (1952) *An Outline of a Theory of Semantic Information.* Technical Report 247. The Research Laboratory of Electronics, MIT.
Reprinted as Chapter 15 of Bar-Hillel, Y. (1964) *Language and information*, Addison-Wesley.

Castelfranchi, C. (1998) Modeling Social Action for AI Agents. *Artificial Intelligence*, 103, pp.157-182.

Chang, C.C. & Keisler, H.J. (1990) *Model Theory.* 3rd Edition. North Holland Publishing Co.

Chopra, S., Ghose, A., & Meyer, T. (2003) Non-prioritized ranked belief change. *Journal of Philosophical Logic*, 32(4), pp.417-443.
A preliminary version appears in: (2001) van Benthem, J. ed. *Proceedings of the 8th Conference on Theoretical Aspects of Reasoning and Knowledge (TARK VIII)*, 8-10 July, Certosa di Pontignano, University of Siena, Italy. Morgan Kaufmann, pp.151-162, by the authors "Meyer, T., Ghose, A., & Chopra, S.".

Chopra, S., Ghose, A., & Meyer, T. (2006) Social choice theory, belief merging, and strategy-proofness. *Information Fusion*, 7, pp.61-79.
A preliminary version appears in: (2001) Benferhat, S. & Besnard, P. eds. *Proceedings of the 6th European Conference on Symbolic and Quantitative Approached to Reasoning with Uncertainty (ECSQARU 2001)*, 19-21 September, Toulouse, France. Lecture Notes in Artificial Intelligence, 2143. Springer-Verlag, pp.466-477, under the title "Social choice, merging and elections." by the authors "Meyer, T., Ghose, A., & Chopra, S.".

Chopra, S., Meyer, T., & Wong, K-S. (2006) Iterated belief change and the recovery axiom. *To appear in Journal of Philosophical Logic.*
A preliminary version appears in: (2002) van Harmelen, F. ed. *Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'2002)*, 21-26 July, Lyon, France. Frontiers in Artificial Intelligence and Applications, 77. IOS Press, pp.541-545, under the title "Iterated revision and the axiom of recovery: A unified treatment via epistemic states." by the authors "Chopra, S., Ghose, A., & Meyer, T.".

Cohen, P.R. & Levesque, H.J. (1990) Intention Is Choice with Commitment. *Artificial Intelligence*, 42, pp.213-261.

Cohen, P.R. & Levesque, H.J. (1995) Communicative Actions for Artificial Agents. In: Lesser, V.R., & Gasser, L. eds. *Proceedings of the 1st International Conference on Multi-Agent Systems (ICMAS'95)*, 12-14 June, San Francisco, California, USA. The MIT Press, pp.65-72.

Cohen, P.R. & Levesque, H.J. (1997) Communicative Actions for Artificial Agents. In: Bradshaw, J.M. ed. *Software Agents*. Menlo Park, CA, AAAI/MIT Press, pp.419-436.

Condorcet (M.J.A.N. de Condorcet) (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* Paris, Imprimerie Royale. Facsimile published in 1972 by Chelsea Publishing Company, New York.

Dalal, M. (1988) Investigations into a theory of knowledge base revision. In: *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence (AAAI'88)*, 21-26 August, St. Paul, Minnesota, USA. AAAI Press/The MIT Press, pp.475-479.

Damasio, A.R. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY, Gosset/Putnam Press.

Darwiche, A. & Pearl, J. (1994) On the logic of iterated belief revision. In: Fagin, R. ed. *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning about Knowledge (TARK V)*, 13-16 March, Pacific Grove, California, USA. Morgan Kaufmann, pp.5-23.

Darwiche, A. & Pearl, J. (1997) On the logic of iterated belief revision. *Artificial Intelligence*, 89, pp.1-29.

Davis, E. (1990) *Representations of Commonsense Knowledge*. Morgan Kaufmann.

Delgrande, J., Dubois, D., & and Lang, J. (2006) Iterated revision as prioritised merging. In: Doherty, P., Mylopoulos, J., & Welty, C.A. eds. *Proceedings of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR 2006)*, 2-5 June, Lake District of the United Kingdom. AAAI Press, pp.210-220.

Delgrande, J., Lang, J., Rott, H., & Tallon, J-M. (2005) Executive summary. In: Delgrande, J., Lang, J., Rott, H., & Tallon, J-M. eds. *Belief Change in Rational Agents: Perspectives from Artificial Intelligence, Philosophy, and Economics - Dagstuhl Seminar Proceedings 05321*, [Internet] 8-12 August, Schloss Dagstuhl, Germany. Available from: <http://drops.dagstuhl.de/opus/volltexte/2005/357>

Dennett, D.C. (1987) *The Intentional Stance.* Cambridge, MA, MIT Press.

Doherty, P., Lukaszewicz, W., & Madalińska-Bugaj, E. (1998) The PMA and relativizing minimal change for action update. In: Cohn, A.G., Schubert, L.K., & Shapiro, S.C. eds. *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, 2-5 June, Trento, Italy. Morgan Kaufmann, pp.258-269.

Doyle, J. & Wellman, M.P. (1991) Impediments to Universal Preference-Based Default Theories, *Artificial Intelligence*, 49(1-3), pp.97-128.
A preliminary version appears in: (1989) Brachman, R.J., Levesque, H.J., & Reiter, R. eds. *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, 15-18 May, Toronto, Canada. Morgan Kaufmann, pp.94-102.

Dubois, D., Lang, J., & Prade, H. (1994) Possibilistic logic. In: Gabbay, D.M., Hogger, C.J., Robinson, J.A. & Nute, D. eds. *Handbook of Logic in Artificial Intelligence and Logic Programming*, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning. Oxford University Press, pp.439-513.

Dubois, D. & Prade, H. (1991) Epistemic entrenchment and possibilistic logic. *Artificial Intelligence*, 50, pp.223-239.

Dubois, D. & Prade, H. (1992) Belief change and possibility theory. In: Gärdenfors, P. ed. *Belief Revision*, Cambridge Tracts in Theoretical Computer Science, 29. Cambridge, UK, Cambridge University Press, pp.142-182.

Dubois, D. & Prade, H. (1998) Introduction: Revising, Updating and Combining Knowledge. In: Gabbay, D.M. & Smets, P. eds. *Handbook of Defeasible Reasoning and Uncertainty Management Systems,* Volume 3: Belief Change. Dordrecht, Kluwer Academic Publishers, pp.1-15.

Dubois, D. & Prade, H. (2004) On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems*, 142, pp.143-161.

Eiter, T. & Gottlob, G. (1992) On the complexity of propositional knowledge base revision, updates, and counterfactuals. *Artificial Intelligence*, 57, pp.227-270.
An earlier version appears in: (1992) *Proceedings of the 11th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Databases (PODS'92),* 2-4 June, San Diego, California, USA. ACM Press.

Fagin, R. & Halpern, J.Y. (1994) Reasoning about knowledge and probability. *Journal of the ACM*, 41(2), pp.340-367.
A preliminary version appears in: (1988) Vardi, M.Y. ed. *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning About Knowledge (TARK II)*. 7-9 March, Pacific Grove, California, USA. Morgan Kaufmann, pp.277-293.

Fagin, R., Halpern, J.Y., Moses, Y., & Vardi, M.Y. (1995) *Reasoning about Knowledge.* Cambridge, MIT Press.

Festinger, L. (1957) *Cognitive Dissonance.* Stanford, CA, Stanford University Press.

Finin, T., Fritzson, R., McKay, D., & McEntire, R. (1994) KQML as an Agent Communication Language. In: *Proceedings of the 3rd International Conference on Information and Knowledge Management (CIKM'94)*, 29 November-2 December, Gaithersburg, Maryland, USA. ACM Press, pp.456-463.

Finin, T., Labrou, Y., & Mayfield, J. (1997) KQML as an Agent Communication Language. In: Bradshaw, J.M. ed. *Software Agents*. Menlo Park, CA, AAAI/MIT Press, pp.291-316.

Fitting, M. (1993) Basic Modal Logic. In: Gabbay, D.M., Hogger, C.J., & Robinson, J.A. eds. *Handbook of Logic in Artificial Intelligence and Logic Programming*, Volume 1: Logical Foundations. Oxford, Clarendon Press, pp.268-448.

Ferguson, D. & Labuschagne, W.A. (2001) A preferential semantics for epistemic logic. *OUCS-2001-09*, Department of Computer Science, University of Otago.

Fermé, E. & Rott, H. (2004) Revision by comparison. *Artificial Intelligence*, 157, pp.5-47.

Forbus, K.D. (1989) Introducing actions into qualitative simulation. In: Sridhanan, N. ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, August, Detroit, Michigan, USA. Morgan Kaufmann, pp.1273-1278.

Freund, M., Lehmann, D., & Morris P. (1991) Rationality, Transitivity, and Contraposition. *Artificial Intelligence*, 52, pp.191-203.

Freund, M. & Lehmann, D. (1994) *Belief revision and rational inference*. Technical Report TR-94-16. The Leibniz Center for Research in Computer Science, Institute of Computer Science, The Hebrew University of Jerusalem.

Friedman, N. & Halpern, J.Y. (1997) Modeling Belief in Dynamic Systems. Part I: Foundations. *Artificial Intelligence*, 95(2), pp.257-316.
A preliminary version appears in: (1994) Fagin, R. ed. *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning about Knowledge (TARK V)*, 13-16 March,

Pacific Grove, California, USA. Morgan Kaufmann, pp.44-64, under the title "A knowledge-based framework for belief change, Part I: Foundations.".

Friedman, N. & Halpern, J.Y. (1999a) Belief Revision: A Critique. *Journal of Logic, Language, and Information*, 8, pp.401-420.
A preliminary version appears in: (1996) Aiello, L.C., Doyle, J., & Shapiro, S.C. eds. *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*, 5-8 November, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.421-431.

Friedman, N. & Halpern, J.Y. (1999b) Modeling Belief in Dynamic Systems. Part II: Revision and Update. *Journal of Artificial Intelligence Research*, 10, pp.117-167.
A preliminary version appears in: (1994) Doyle, J., Sandewall, E., & Torasso, P. eds. *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, 24-27 May, Bonn, Germany. Morgan Kaufmann, pp.190-201, under the title "A knowledge-based framework for belief change, Part II: revision and update.".

Fuhrmann, A. (1991) Theory contraction through base contraction. *Journal of Philosophical Logic*, 20, pp.175-203.

Gabbay, D.M. (1985) Theoretical foundations for non-monotonic reasoning in expert systems. In: Apt, K.R. ed. *Proceeding NATO Advanced Study Institute on Logics and Models of Concurrent Systems*, La Colle-sur-Loop, France. Berlin, Springer, pp.439-457.

Gaertner, W. (2002) Domain Restrictions. In: Arrow, K.J., Sen, A.K., & Suzumura, K. eds. *Handbook of Social Choice and Welfare*, Volume 1. Handbooks in Economics, 19. North-Holland, Elsevier, pp.131-170.

Gärdenfors, P. (1978) Conditionals and changes of belief. In: Niiniluoto, I. & Tuomela, R. eds. *The Logic and Epistemology of Scientific Change. Acta Philosophica Fennica*, 30. pp.381-404.

Gärdenfors, P. (1982) Rules for rational changes of belief. In: Pauli, T. ed. *320311: Philosophical Essays Dedicated to Lennart Åqvist on his Fiftieth Birthday.* Philosophical Studies, 34. Department of Philosophy, University of Uppsala, Uppsala, pp.88-101.

Gärdenfors, P. (1984) Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy*, 62, pp.136-157.

Gärdenfors, P. (1988) *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* Cambridge, MA, MIT Press.

Gärdenfors, P. (1990) The Dynamics of Belief Systems: Foundations vs. Coherence Theories. *Revue International de Philosophie*, 44, pp.24-46.

Gärdenfors, P. & Makinson, D. (1988) Revisions of knowledge systems using epistemic entrenchment. In: Vardi, M.Y. ed. *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning About Knowledge (TARK II)*. 7-9 March, Pacific Grove, California, USA. Morgan Kaufmann, pp.83-95.

Gärdenfors, P. & Makinson, D. (1994) Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65, pp.197-245.

Gärdenfors, P. & Rott, H. (1995) Belief Revision. In: Gabbay, D.M., Hogger, C.J., & Robinson, J.A. eds. *Handbook of Logic in Artificial Intelligence and Logic Programming*, Volume 4: Epistemic and Temporal Reasoning. Oxford, Clarendon Press, pp.35-132.

Genesereth, M.R. & Fikes, R.E. (1992) *Knowledge Interchange Format Version 3, Reference Manual.* Technical Report Logic-92-1. Stanford University Logic Group.

Genesereth, M.R. & Ketchpel, S.P. (1994) Software Agents. *Communications of the ACM*, 37(7), July, pp.48-53.

Gigerenzer, G. (2002) *Calculated Risks. How to Know When Number Deceive You.* New York, Simon and Schuster.

Girle, R.A. (1998) Logical Fiction: Real vs. Ideal. In: Lee, H-Y. & Motodo, H. eds. *Topics in Artificial Intelligence, Proceedings of the 5th Pacific Rim International Conference on Artificial Intelligence (PRICAI'98)*, 22-27 November, Singapore. Lecture Notes in Artificial Intelligence, 1531. Springer-Verlag, pp.542-552.

Goldszmidt, M. & Pearl, J. (1992) Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and action. In: Nebel, B., Rich, C., & Swartout, W.R. eds. *Proceedings of the 3th International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, 25-29 October, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.661-672.

Grahne, G. (1991) Updates and counterfactuals. In: Allen, J.F., Fikes, R., & Sandewall, E. eds. *Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning (KR'91)*, 22-25 April, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.269-276.

Grégoire, E. & Konieczny, S. (2006) Logic-based approaches to information fusion. *Information Fusion*, 7, pp.4-18.

Grice, H.P. (1957) "Meaning", *The Philosophical Review*, 66, pp.377-388.

Grove, A. (1988) Two modellings for theory change. *Journal of Philosophical Logic*, 17, pp.157-170.

Gulbransen, E.A. (1975) "Not Safe Enough". *Bulletin of the Atomic Scientists*, June, 5.

Guttag, J. (1986) Notes on Type Abstraction. In: Gehani, N. & McGettrick, A.D. eds. *Software Specification Techniques.* Wokingham, England, Addison-Wesley Publishing Company, pp.55-74.

Haidt, J. (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), pp.814-834.

Halpern, J.Y. (1986) Reasoning about Knowledge: An Overview. In: Halpern, J.Y. ed. *Proceedings of the 1st Conference on Theoretical Aspects of Reasoning about Knowledge (TARK I)*, 19-22 March, Monterey, California, USA. Morgan Kaufmann, pp.1-17.

Halpern, J.Y. (1987) Using reasoning about knowledge to analyse distributed systems. In: Traub, J.F., Grosz, B.J., Lampson, B.W., & Nilsson, N.J. eds. *Annual Review of Computer Science,* Volume 2. Palo Alto, Annual Reviews Inc., pp.37-68.

Halpern, J.Y. and Fagin, R. (1989) Modeling knowledge and action in distributed systems. *Distributed Computing*, 3(4), pp.159-179.

Halpern, J.Y. (1995) Reasoning About Knowledge: A Survey. In: Gabbay, D.M., Hogger, C.J., & Robinson, J.A. eds. *Handbook of Logic in Artificial Intelligence and Logic Programming*, Volume 4: Epistemic and Temporal Reasoning. Oxford, Clarendon Press, pp.1-34.

Halpern, J.Y. (1996) Should knowledge entail belief? *Journal of Philosophical Logic*, 25:5, pp.483-494.

Hanson, B. (1969): An analysis of some deontic logics. *Noûs*, 3, pp.373-398.
Reprinted in Hilpinen, R. ed. (1971) *Deontic logic: Introductory and systematic readings.* Dordrecht, Reidel, pp.121-147.

Hansson, B. (1973) The independence condition in the theory of social choice. *Theory and Decision*, 4, pp.25-49.

Hansson, S.O. (1992) In defense of base contraction. *Synthese*, 91, pp.239-245.

Hansson, S.O. (1997) Semi-revision. *Journal of Applied Non-Classical Logic*, 7(1-2), pp.151-175.

Hansson, S.O. (1998) Revision of Belief Sets and Belief Bases. In: Gabbay, D.M. & Smets, P. eds. *Handbook of Defeasible Reasoning and Uncertainty Management Systems,* Volume 3: Belief Change. Dordrecht, Kluwer Academic Publishers, pp.17-75.

Hansson, S.O. (1999a) *A Textbook of Belief Dynamics - Theory Change and Database Updating.* Applied Logic Series, 11. Dordrecht, Kluwer Academic Press.

Hansson, S.O. (1999b) A survey of non-prioritized belief revision. *Erkenntnis*, 50(2-3), pp.413-427.

Hansson, S.O. (2006) Coherence in epistemology and belief revision. *Philosophical Studies*, 128, pp.93-108.

Hansson, S.O. & Olsson, E.J. (1999) Providing foundations for coherentism. *Erkenntnis*, 51, pp.243-265.

Harman, G. (1986) *Change in view.* Cambridge MA, A Bradford Book, MIT Press.

Harnad, S. (1990) The Symbol Grounding Problem. *Physica D*, 42, pp.335-346.

Harper, W.L. (1976) Ramsey test conditionals and iterated belief change. In: Harper, W.L. & Hooker, C.A. eds. *Foundations of Probability Theory and Statistical Theories of Sciences*, Volume 1. Norwell, MA, D. Reidel, pp.117-135.

Harper, W.L. (1977) Rational conceptual change. In: *Proceedings of the Meeting of the Philosophy of Science Association (PSA 1976),* Volume 2. East Lansing, Philosophy of Science Association, pp.462-494.

Herzig, A. (1996) The PMA Revisited. In: Aiello, L.C., Doyle, J., & Shapiro, S.C. eds. *Proceedings of the 5th International Conference on Principles of Knowledge Repre-*

sentation and Reasoning (KR'96), 5-8 November, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.40-50.

Herzig, A. (1998) Logics for Belief Base Updating. In: Gabbay, D.M. & Smets, P. eds. *Handbook of Defeasible Reasoning and Uncertainty Management Systems,* Volume 3: Belief Change. Dordrecht, Kluwer Academic Publishers, pp.189-231.

Herzig, A. & Rifi, O. (1999) Propositional belief base update and minimal change. *Artificial Intelligence*, 115(1), pp.107-138.

Hewitt, C. (1977) Viewing control structures as patterns of passing messages. *Artificial Intelligence*, 8, pp.323-364.

Hintikka, J. (1962) *Knowledge and Belief.* Ithaca, New York, Cornell University Press.

Hintikka, J. (1970) On semantic information. In: Yourgrau, W. & Breck, A.D. eds. *Physics, Logic, and History.* New York, Plenum Press, pp.147-172.

Hughes, G.E. & Cresswell, M.J. (1968) *An introduction to modal logic.* London, Methuen.

Hughes, G.E. & Cresswell, M.J. (1996) *A new introduction to modal logic.* London, Routledge.

Huhns, M.N. & Singh, M.P. eds. (1998) *Readings in Agents.* San Francisco, CA, Morgan Kaufmann Publishers.

Jin, Y. & Thielscher, M. (2007) Iterated belief revision, revised. *Artificial Intelligence*, 171(1), pp.1-18.
A preliminary version appears in: (2005) Kaelbling, L.P. & Saffiotti, A. eds. *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05)*, 30 July-5 August, Edinburgh, Scotland, UK. Professional Book Centre, pp.478-483.

Kahneman, D., Slovic, P., & Tversky, A. (1982) *Judgement under uncertainty: Heuristics and biases.* Cambridge, NY, Cambridge University Press.

Kahneman, D. & Tversky, A. (1979) Prospect theory: An analysis of decision under risk, *Econometrica*, 47, pp.263-291.

Kahneman, D. & Tversky, A. (2000) *Choices, Values, and Frames*. Cambridge, NY, Cambridge University Press.

Katsuno, H. & Mendelzon, A.O. (1989) A unified view of propositional knowledge base updates. In: Sridhanan, N. ed. *Proceedings of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, August, Detroit, Michigan, USA. Morgan Kaufmann, pp.1413-1419.

Katsuno, H. & Mendelzon, A.O. (1991) Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52, pp.263-294.

Katsuno, H. & Mendelzon, A.O. (1992) On the Difference between Updating a Knowledge Base and Revising it. In: Gärdenfors, P. ed. *Belief Revision*, Cambridge Tracts in Theoretical Computer Science, 29. Cambridge, UK, Cambridge University Press, pp.183-203.
A preliminary version appears in: (1991) Allen, J.F., Fikes, R., & Sandewall, E. eds. *Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning (KR'91)*, 22-25 April, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.387-394.

Keller, A.M, & Winslett Wilkins, M. (1985) On the use of an extended relational model to handle changing incomplete information. *IEEE Transactions on Software Engineering*, SE-11:7, pp.620-633.

Kemeny, J.G. (Chairman). (1979) *The Need for Change: The Legacy of TMI.* Report of the President's Commission on the Accident at Three Mile Island. Washington D.C., Government Printing Office.

Konieczny, S., Lang, J., & Marquis, P. (2004) DA^2 Merging Operators. *Artificial Intelligence*, 157(1-2), pp.49-79.

Konieczny, S., Lang, J., & Marquis, P. (2002) Distance Based Merging: A General Framework and some Complexity Results. In: Fensel, D., Guinchiglia, F., & McGuinness, D. eds. *Proceedings of the 8th International Conference on Principles of Knowledge Representation and Reasoning (KR 2002)*, 22-25 April, Toulouse, France. Morgan Kaufmann, pp.97-108.

Konieczny, S. & Pino-Pérez, R. (1998) On the logic of merging. In: Cohn, A.G., Schubert, L.K., & Shapiro, S.C. eds. *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR'98)*, 2-5 June, Trento, Italy. Morgan Kaufmann, pp.488-498.

Konieczny, S. & Pino-Pérez, R. (1999) Merging with integrity constraints. In: Hunter, A. & Parsons, S. eds. *Proceedings of the 5th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'99)*, 5-9 July, London, UK. Lecture Notes in Artificial Intelligence, 1638. Springer-Verlag, pp.233-244.

Konieczny, S. & Pino-Pérez, R. (2002) Merging Information under Constraints: A Logical Framework. *Journal of Logic and Computation*, 12(5), pp.773-808.

Konolige, K. (1986) *A Deduction Model of Belief.* Research Notes in Artificial Intelligence. London, Pitman.

Kraus, S. & Lehmann, D. (1988) Knowledge, Belief and Time. *Theoretical Computer Science*, 58, pp.155-174.

Kraus, S., Lehmann, D., & Magidor, M. (1990) Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, 44, pp.167-207.

Kripke, S.A. (1963) Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16, pp.83-94.

Kyburg, H.E. Jr. (1961) *Probability and the Logic of Rational Belief*. Middletown, Wesleyan University Press.

Labuschagne, W.A. & Ferguson, D. (2002) Information-theoretic semantics for epistemic logic. In: Bonanno, G., Colombatto, E., & van der Hoek, W. eds. *Proceedings of the Fifth Conference on Logic and the Foundations of Game and Decision Theory (LOFT 5)*, 28-30 June, Torino, Italy. Informal proceedings.

Labuschagne, W.A., Ferguson, D., Heidema, J., Meyer, T.A., & van der Westhuizen, P.L. (2002) Information templates - A useful data structure for applied logic. *LAMAS: the Otago Workshop on Logic and Multi-Agent Systems*, October 2002.

Labuschagne, W.A. & Heidema, J. (2001) Knowledge and belief - The agent-oriented view. In: Roux, A.P.J. & Coetzee, P.H. eds. *Culture in Retrospect*, Pretoria, UNISA Press, pp.194-214.

Lang, J., Marquis, P. & Williams, M-A. (2001) Updating Epistemic states. In: Stumptner, M., Corbett, D., & Brooks, M. eds. *AI 2001: Advance in Artificial Intelligence, Proceedings of the 14th Australian Joint Conference on Artificial Intelligence*, 10-14 December, Adelaide, Australia. Lecture Notes in Artificial Intelligence, 2256. Springer, pp.297-308.

LeDoux, J.E. (1996) *The Emotional Brain.* New York, Simon & Schuster.

Lehmann, D. & Magidor, M. (1992) 'What does a conditional knowledge base entail?'. *Artificial Intelligence*, 55, pp.1-60.

A preliminary version appears in: (1989) Brachman, R.J., Levesque, H.J., & Reiter, R. eds. *Proceedings of the 1st International Conference on Principles of Knowledge*

Representation and Reasoning (KR'89), 15-18 May, Toronto, Canada. Morgan Kaufmann, pp.212-222, by the author "Lehmann, D.".

Lehmann, D. (1995) Belief revision, revised. In: Mellish, C.S. ed. *Proceeding of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, 20-25 August 1995, Montreal, Quebec. San Mateo, CA, Morgan Kaufmann, pp.1534-1540.

Lehmann, D. (2001) Nonmonotonic Logics and Semantics. *Journal of Logic and Computation*, 11(2), pp.229-256.

Lenzen, W. (1978) Recent work in epistemic logic. *Acta Philosophica Fennica*, 30, pp.1-219.

Levi, I. (1977) Subjunctives, dispositions, and chances. *Synthese*, 34, pp.423-455.

Levi, I. (1980) *The Enterprise of Knowledge*. Cambridge, MA, MIT Press.

Lewis, D.K. (1973) Counterfactuals. *Journal of Philosophy*, 70, pp.556-567.

Liberatore, P. (1997) The Complexity of Iterated Belief Revision. In: Afrati, F.N. & Kolaitis, P.G. eds. *Proceedings of the 6th International Conference on Database Theory (ICDT'97)*, 8-10 January, Delphi, Greece. Lecture Notes in Artificial Intelligence, 1186. Springer, pp.276-290.

Liberatore, P. & Schaerf, M. (2001) Belief Revision and Update: Complexity of Model Checking. *Journal of Computer and System Sciences*, 62(1), pp.43-72.
An earlier version appears in: (1996) The complexity of model checking for belief revision and update. In: *Proceedings of the 13th National Conference of the American Association for Artificial Intelligence (AAAI'96)*, 4-8 August, Portland, Oregon, USA. AAAI Press/The MIT Press, pp.556-561, under the title "The Complexity of Model Checking for Belief Revision and Update.".

Liberatore, P. & Schaerf, M. (1998) Arbitration (or How to Merge Knowledge Bases). *IEEE Transactions on Knowledge and Data Engineering*, 10(1), pp.76-90.
A preliminary version appears in: (1995) De Glas, M. & Pawlak, Z. eds. *Proceedings of the 2nd World Conference on the Fundamentals of Artificial Intelligence (WOC-FAI'95)*, 3-7 July, Paris, France. Angkor, France, pp.217-228, under the title "Arbitration: A Commutative Operator for Belief Revision.".

Liberatore, P. & Schaerf, M. (2000) BReLS: A System for the Integration of Knowledge Bases. In: Cohn, A.G., Guinchiglia, F., & Selman, B. eds. *Proceedings of the 7th International Conference on Principles of Knowledge Representation and Reasoning (KR 2000)*, 12-15 April, Breckenridge, Colorado, USA. Morgan Kaufmann, pp.145-152.

Lin, J. (1996) Integration of weighted knowledge bases. *Artificial Intelligence*, 83(2), pp.363-378.

Lin, J. & Mendelzon, A.O. (1999) Knowledge base merging by majority. In: Pareschi, R. & Fronhoefer, B. eds. *Dynamic Worlds: From the Frame Problem to Knowledge Management*. Kluwer, pp.195-218.

Loeckx, J., Ehrich, H., & Wolf, M. (1996) *Specification of Abstract Data Types.* Chichester, Wiley Teubner.

Makinson, D. (1987) On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic*, 16, pp.383-394.

Makinson, D. (1993) Five Faces of Minimality. *Studia Logica*, 52, pp.339-379.

Makinson, D. (1997a) Screened revision. *Theoria*, 63, pp.14-23.

Makinson, D. (1997b) On the force of some apparent counterexamples to recovery. In: Valdés, E.G., Krawietz, W., von Wright, G.H., and Zimmerling, R. eds. *Normative Systems in Legal and Moral Theory.* Berlin, Dunker and Humboldt, pp.475-481.

Makinson, D. (1999) On a fundamental problem of deontic logic. In: McNamara, P. & Prakken, H. eds. *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science*. IOS Press, pp.29-54.

Makinson, D. (2003a) Ways of Doing Logic: What was Different about AGM 1985? *Journal of Logic and Computation*, 13, pp.3-13.

Makinson, D. (2003b) Bridges between classical and nonmonotonic logic. *Logic Journal of the IGPL*, 11(1), pp.69-96.

Makinson, D. (2005) *Bridges from Classical to Nonmonotonic Logic*. Texts in Computing, 5. London, King's College.

Makinson, D. & Gärdenfors, P. (1991) Relations between the logic of theory change and nonmonotonic logic. In: Fuhrman, A. & Morreau, M. eds. *The Logic of Theory Change*, Lecture Notes in Artificial Intelligence, 465. Springer-Verlag, pp.185-205.

Maynard-Reid II, P. & Shoham, Y. (2001) Belief Fusion: Aggregating Pedigreed Belief States. *Journal of Logic, Language and Information*, 10(2), pp.183-209.
A preliminary version appears in: (1998) Bonanno, G., Colombatto, E., Kaneko, M., & van der Hoek, W. eds. *Proceedings of the 3rd Conference on Logic and the Foundations of Game and Decision Theory (LOFT 3)*, 17-20 December, Torino, Italy. Informal proceedings, under the names "Maynard-Reid II, P. & Shoham, Y.".

Maynard-Zhang, P. & Lehmann, D. (2003) Representing and Aggregating Conflicting Beliefs. *Journal of Artificial Intelligence Research*, 19, pp.155-203.
A preliminary version appears in: (2000) Cohn, A.G., Guinchiglia, F., & Selman, B. eds. *Proceedings of the 7th International Conference on Principles of Knowledge Representation and Reasoning (KR 2000)*, 12-15 April, Breckenridge, Colorado, USA. Morgan Kaufmann, pp.153-164, under the names "Maynard-Reid II, P. & Lehmann, D.".

McCarthy, J. (1979) *Ascribing mental qualities to machines.* Technical Report Memo 326. Stanford AI Lab.

McCarthy, J. (1980) Circumscription – a form of non-monotonic reasoning. *Artificial Intelligence*, 13, pp.25-39.

Meyer, T.A. (1999) *Semantic belief change*. Ph.D. thesis, University of South Africa.

Meyer, T. (2001a) On the semantics of combination operations. *Journal of Applied Non-Classical Logics*, 11(1-2), pp.59-84.

A preliminary version appears in: (2000) Mizoguchi, R. & Slaney, J. eds. *Proceedings of the 6th Pacific Rim International Conference on Artificial Intelligence (PRICAI2000)*, 28 August-1 September, Melbourne, Australia. Lecture Notes in Artificial Intelligence, 1886. Springer-Verlag, pp.286-296, under the title "Merging Epistemic States.".

Meyer, T. (2001b) Basic Infobase Change. *Studia Logica*, 67, pp.215-242.

Meyer, T., Labuschagne, W.A., & Heidema, J. (2000a) Refined epistemic entrenchment. *Journal of Logic, Language, and Information*, 9(2), pp.237-259.

Meyer, T., Labuschagne, W.A., & Heidema, J. (2000b) Infobase Change: A First Approximation. *Journal of Logic, Language, and Information*, 9(3), pp.353-377.

Moore, S. & Oaksford, M. (2002) *Emotional Cognition: From brain to behaviour*, Advances in Consciousness Research, 44. John Benjamins Publishing Company.

Mostowski, A. (1957) On a generalization of quantifiers. *Fundamenta Mathematicae*, 44, pp.12-36.

Moulin, H. (1985) Choice functions over a finite set: A summary. *Social Choice and Welfare*, 2, pp.147-160.

Moses, Y. & Shoham, Y. (1993) Belief as defeasible knowledge. *Artificial Intelligence*, 64, pp.299-321.

Nayak, A.C., Pagnucco, M., & Peppas, P. (2003) Dynamic belief reivison operators. *Artificial Intelligence*, 146, pp.192-228.

Nayak, A., Foo, N., Pagnucco, M., & Sattar, A. (1996) Changing conditional beliefs unconditionally. In: Shoham, Y. ed. *Proceedings of the 6th Conference on Theoretical Aspects of Reasoning and Knowledge (TARK VI)*, 17-20 March, De Zeeuwse Stromen, The Netherlands. Morgan Kaufmann, pp.119-135.

Nayak, A., Goebel, R., Orgun, M., & Pham, T. (2005) Iterated Belief Change and the Levi Identity. In: Delgrande, J., Lang, J., Rott, H., & Tallon, J-M. eds. *Belief Change in Rational Agents: Perspectives from Artificial Intelligence, Philosophy, and Economics - Dagstuhl Seminar Proceedings 05321*, [Internet] 8-12 August, Schloss Dagstuhl, Germany. Available from: <http://drops.dagstuhl.de/opus/volltexte/2005/331>

Nebel, B. (1989) A knowledge level analysis of belief revision. In: Brachman, R.J., Levesque, H.J., & Reiter, R. eds. *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, 15-18 May, Toronto, Canada. Morgan Kaufmann, pp.301-311.

Nebel, B. (1992) Syntax based approaches to belief revision. In: Gärdenfors, P. ed. *Belief Revision.* Cambridge Tracts in Theoretical Computer Science, 29. Cambridge, UK, Cambridge University Press, pp.52-88.

Nebel, B. (1998) How Hard is it to Revise a Belief Base? In: Gabbay, D.M. & Smets, P. eds. *Handbook of Defeasible Reasoning and Uncertainty Management Systems,* Volume 3: Belief Change. Dordrecht, Kluwer Academic Publishers, pp.77-145.

Neches R., Fikes, R.E., Finin, T., Gruber, T., Patil, R., Senator, T., & Swartout, W.R. (1991) Enabling Technology for Knowledge Sharing. *AI Magazine*, 12(3), Fall, pp.36-56.

Oatley, K. & Johnson-Laird, P.N. (1987) Towards a Cognitive Theory of Emotions. *Cognition and Emotion*, 1, pp.29-50.

Ortony, A., Clore, G., & Collins, A. (1988). *The Cognitive Structure of Emotions.* Cambridge University Press.

Palladino, N.J. (1976) "Defends Zirconium". *Bulletin of the Atomic Scientists*, March, 5.

Papadimitriou, C.H. (1994) *Computational Complexity.* Addison-Wesley.

Papini, O. (2001) Iterated revision operations stemming from the history of an agent's observations. In: Williams, M. & Rott, H. eds. *Frontiers in Belief Revision.* Applied Logic Series, 22. Dordrecht, Kluwer Academic Publishers, pp.279-301.

Parikh, R. (1985) The logic of games and its applications. *Annals of Discrete Mathematics*, 24, pp.111-140.

Pattanaik, P.K. (2002) Positional Rules of Collective Decision-Making. In: Arrow, K.J., Sen, A.K., & Suzumura, K. eds. *Handbook of Social Choice and Welfare,* Volume 1. Handbooks in Economics, 19. North-Holland, Elsevier, pp.361-394.

Perrow, C. (1984) *Normal Accidents: Living with High-Risk Technologies.* New York, Basic Books Inc. Publishers.

Piattelli-Palmarini, M. (1994) *Inevitable Illusions: How Mistakes of Reason Rule Our Minds.* New York, Wiley.

Picard, R.W. (1997) *Affective computing.* MIT Press.

Pigozzi, G. (2006) Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgement aggregation. *Forthcoming in Synthese.*

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976) Basic Objects in Natural Categories, *Cognitive Psychology*, 8, pp.382-439.

Pollock, J.L. (1995) *Cognitive Carpentry*, MIT Press.

Poole, D. (1989) What the lottery paradox tells us about default reasoning. *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, 15-18 May, Toronto, Canada. Morgan Kaufmann, pp.333-340.

Poole, D. (1991) The effect of knowledge on belief: conditioning, specificity and the lottery paradox in default reasoning. *Artificial Intelligence*, 49, pp.281-307.

Popper, K.R. (1934) *The Logic of Scientific Discovery.* Harper & Row Publishers.

Rao, A.S. & Georgeff, M.P. (1991) Modeling Rational Agents within a BDI-Architecture. In: Allen, J.F., Fikes, R. & Sandewall, E. eds. *Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning (KR'91)*, 22-25 April, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.473-484.

Rao, A.S. & Georgeff, M.P. (1998) Decision procedures for BDI logics. *Journal of Logic and Computation*, 8(3), pp.293-344.

Rasmussen, J. (1986) *Information Processing and Human-machine Interaction: An Approach to Cognitive Engineering.* Elsevier Science Publishers.

Reilly, W.S., & Bates, J. (1992) *Building emotional agents.* Technical Report CMU-CS-92-143. Carnegie Mellon University.

Reiter, R. (1980) A logic for default reasoning. *Artificial Intelligence*, 13, pp.81-132.

Revesz, P.Z. (1993) On the Semantics of Theory Change: Arbitration between Old and New Information. In: *Proceedings of the 12th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Databases (PODS'93),* 25-28 May, Washington DC, New York, USA. ACM Press, pp.71-92.

Revesz, P.Z. (1997) On the Semantics of Arbitration. *International Journal of Algebra and Computation*, 7(2), pp.133-160.

Rott, H. (1998) Logic and Choice. In: Gilboa, I. ed. *Proceedings of the 7th Conference on Theoretical Aspects of Reasoning and Knowledge (TARK VII)*, 22-24 July, Evanston, Illinois, USA. Morgan Kaufmann, pp.235-248.

Rott, H. (2001) *Change, choice and inference: A study of belief revision and non-monotonic reasoning.* Oxford Logic Guides, 42. Oxford, Clarendon Press.

Rozin, P. (1997) Moralization. In: Brandt, A. & Rozin, P. eds. *Morality and Health.* New York, Routledge, pp.379-401.

Russell, S.J. & Norvig, P. (1995) *Artificial Intelligence: A Modern Approach.* 1st Edition. Upper Saddle River, NJ, Prentice Hall.

Russell, S.J. & Norvig, P. (2003) *Artificial Intelligence: A Modern Approach.* 2nd Edition. Upper Saddle River, NJ, Prentice Hall.

Ryan, M. & Schobbens, P-Y. (1997) Counterfactuals and Updates as Inverse Modalities. *Journal of Logic, Language, and Information*, 6, pp.123-146.

Ryan, M., Schobbens, P-Y., & Rodrigues, O. (1996) Counterfactuals and Updates as Inverse Modalities. In: Shoham, Y. ed. *Proceedings of the 6th Conference on Theoretical Aspects of Reasoning and Knowledge (TARK VI)*, 17-20 March, De Zeeuwse Stromen, The Netherlands. Morgan Kaufmann, pp.163-173.

Schmitt, F.F. (1992) *Knowledge and Belief.* London, Routledge.

Searle, J.R. (1969) *Speech acts: An essay in the philosophy of language.* New York, Cambridge University Press.

Searle, J.R. (1979) *Expression and Meaning: Studies in the Theory of Speech Acts.* New York, Cambridge University Press.

Sen, A.K. (1969) Quasi-transitivity, rational choice and collective decisions. *Review of Economic Studies*, 36, pp.381-393.

Sen, A.K. & Pattanaik, P.K. (1969) Necessary and sufficient conditions for rational choice under majority decision. *Journal of Economic Theory*, 1, pp.178-202.

Sen, A.K. (1970) *Collective Choice and Social Welfare*. Mathematical Economics Texts, 5. San Francisco, Holden-Day.

Sen, A.K. (1986) Social Choice Theory. In: Arrow, K.J. & Intrriligator, M.D. eds. *Handbook of Mathematical Economics*, Volume III, Chapter 22. Elsevier Science Publishers, pp.1073-1181.

Shannon, C.E. (1949) *The Mathematical Theory of Communication.* University of Illinois Press.

Shoham, Y. & Cousins, S.B. (1994) Logics of Mental Attitudes in AI. In: Lakemeyer, G. & Nebel, B. eds. *Foundations of Knowledge Representation and Reasoning.* Lecture Notes in Artificial Intelligence, 810. Springer-Verlag, pp.296-309.

Shoham, Y. (1987) A semantical approach to nonmonotonic logics. In: Ginsberg, M.L. ed. *Readings in Nonmonotonic Reasoning.* San Mateo, CA, Morgan Kaufmann, pp.227-250.

Shoham, Y. (1988) *Reasoning about Change - Time and Causation from the Standpoint of Artificial Intelligence.* Cambridge, MA, MIT Press.

Shoham, Y. (1990) *Agent-oriented programming*. Technical Report STAN-CS-1335-90. Computer Science Department, Stanford University.

Shoham, Y. (1993) Agent-oriented programming. *Artificial Intelligence*, 60, pp.51-92.

Shoham, Y. (1997) An overview of agent-oriented programming. In: Bradshaw, J.M. ed. *Software Agents.* Menlo Park, CA, AAAI/MIT Press, pp.271-290.

Simon, H.A. (1967) Motivational and emotional controls of cognition. *Psychological Review*, 74(1), pp.29-39.

Singh, M.P. (1998) Agent communication languages: Rethinking the principles. *IEEE Computer*, 31(12), pp.40-47.

Sosa, E. (1980) The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge. *Midwest Studies in Philosophy*, 5, pp.3-25.

Spohn, W. (1988) Ordinal conditional functions: A dynamic theory of epistemic states. In: Harper, W.L. & Skyrms, B. eds. *Causation in Decision, Belief Change, and Statistics: Proceedings of the Irvine Conference on Probability and Causation*, Volume II, University of California, Irvine. The University of Western Ontario Series in Philosophy of Science, 42. Dordrecht, Reidel, pp.105-134.

Spohn, W. (1990) A general non-probabilistic theory of inductive reasoning. In: Shachter, R.D., Levitt, T.S., Kanal, L.N. & Lemmer, J.F. eds. *Uncertainty in Artificial Intelligence 4*. Amsterdam, North Holland, Elsevier Science Publishers, pp.149-158.

Subrahmanian, V.S. (1994) Amalgamating Knowledge Bases. *ACM Transactions on Database Systems*, 19(2), pp.291-331.

Thomas, S.R. (1999) A survey of agent-oriented programming. In: Wooldridge, M.J. & Rao, A.S. eds. *Foundations of Rational Agency*. Applied Logic Series, 14. Dordrecht, Kluwer Academic Publishers, pp.263-274.

Tversky, A. (1972) Elimination by aspects: A theory of choice, *Psychological Review*, 79, pp.281-299.

Tversky, A. & Kahneman, D. (1974) Judgement under Uncertainty: Heuristics and Biases, *Science*, 185, pp.1124-1131.

Reprinted as Chapter 1 in (1982) Kahneman, D., Slovic, P., & Tversky, A. (1982) *Judgement under uncertainty: Heuristics and biases*. Cambridge, NY, Cambridge

University Press.

Reprinted as Chapter 2 in: (2000) Connolly, T., Arkes, H.R., & Hammond, K.R. eds. *Judgement and Decision Making, An Interdisciplinary Reader,* 2nd ed. Cambridge University Press.

Tversky, A. & Kahneman, D. (1981) The framing of decisions and the psychology of choice, *Science*, 211, pp.453-458.

Tversky, A. & Kahneman, D. (1983) Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review*, 90, pp.293-315.

Tversky, A. & Kahneman, D. (1986) Rational choice and the framing of decisions, *Journal of Business*, 59, pp.261-278.

Tversky, A. & Shafir, E. (1992) Choice under conflict: The dynamics of deferred decision, *Psychological Science*, 3(6), pp.358-361.

van Benthem, J.F.A.K. (2001) Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4), pp.219-248.

van der Hoek, W. & Meyer, J-J. Ch. (1992) Graded Modalities in Epistemic Logic. In: Nerode, A. & Taitslin, M. eds. *Proceedings of the 2nd International Symposium on Logical Foundations of Computer Science (Tver'92)*, 20-24 July, Tver, Russia. Lecture Notes in Computer Science, 620. Springer-Verlag, pp.503-514.

van der Hoek, W. & Wooldridge, M.J. (2003) Towards a Logic of Rational Agency. *Logic Journal of the IGPL*, 11(2), pp.133-157.

van der Torre, L. (2003) Contextual Deontic Logic: Normative Agents, Violations and Independence. *Annals of Mathematics and Artificial Intelligence*, 37(1-2), pp.33-63.

van der Westhuizen, P.L., van der Poll, J.A., & Labuschagne, W.A. (2006) Templated Revision. In: Bishop, J. & Kourie, D. eds. *Service-oriented software systems, Proceed-*

ings of SAICSIT 2006 - Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists, 9-11 October, Somerset West, South Africa. A Volume in the ACM International Conference Proceedings Series. Pretoria, SAICSIT, pp.218-229.

von Wright, G.H. (1951) Deontic logic. *Mind*, 60, pp.1–15.

von Wright, G.H. (1963) *Norm and Action: A Logical Inquiry.* London, Routledge & Kegan Paul.

Williams, M-A. (1994) Transmutations of Knowledge Systems. In: Doyle, J., Sandewall, E., & Torasso, P. eds. *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, 24-27 May, Bonn, Germany. Morgan Kaufmann, pp.619-629.

Williams, M-A. (1996) Towards a Practical Approach to Belief Revision: Reason-Based Change. In: Aiello, L.C., Doyle, J., & Shapiro, S.C. eds. *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR'96)*, 5-8 November, Cambridge, Massachusetts, USA. Morgan Kaufmann, pp.412-420.

Winslett, M. (1988) Reasoning about actions using a possible models approach. In: *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence (AAAI'88)*, 21-26 August, St. Paul, Minnesota, USA. AAAI Press/The MIT Press, pp.89-93.

Winslett, M. (1990) *Updating Logical Databases.* Cambridge, UK, Cambridge University Press.

Wooldridge, M.J. (2000) *Reasoning about Rational Agents.* Cambridge, MA, MIT Press.

Wooldridge, M.J. & Jennings, N.R. (1995) Agent Theories, Architectures, and Languages: A Survey. In: Wooldridge, M.J. & Jennings, N.R. eds. *Intelligent Agents, Proceed-*

ings of the ECAI-94 Workshop on Agent Theories, Architectures, and Languages, 8-9 August, Amsterdam, The Netherlands. Lecture Notes in Artificial Intelligence, 890. Springer-Verlag, pp.1-39.

Wooldridge, M.J. & Rao, A. (1999) *Foundations of Rational Agency.* Applied Logic Series, 14. Dordrecht, Kluwer Academic Publishers.

Zadeh, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, pp.3-28.

Zsambok C.E. & Klein, G. (1997) *Naturalistic Decision Making.* Mahwah, New Jersey, Lawrence Erlbaum Associates, Publishers.

Index