Linking Iterated Belief Change Operations to Nonmonotonic Reasoning

Gabriele Kern-Isberner
Faculty of Computer Science
University of Technology Dortmund, Germany

Abstract

The study of the exact relationships between belief revision and belief update, on one side, and belief revision and nonmonotonic reasoning, on the other, has raised considerable interest, but still, the picture is far from being complete. In this paper, we add some new details to this line of research by making a foundational contribution to the discussions on the very nature of belief change operations, and by introducing universal inference operations as a proper counterpart in nonmonotonic reasoning to iterated belief change. Belief change is investigated within an abstract framework of epistemic states and (qualitative or quantitative) conditionals here. We show how belief revision and belief update can be realized by one and the same generic change operator as simultaneous and successive change operations. We propose general postulates for revision and update that also apply to iterated change. The distinction between background knowledge and evidential information turns out to be a crucial feature in our framework, in order to analyse belief change in more depth.

Introduction

Nonmonotonic reasoning and belief change theory are closely related in that they both deal with reasoning under uncertainty and try to reveal sensible lines of reasoning in response to incoming information. The crucial difference between both areas is the role of the current epistemic state which is only implicit in nonmonotonic reasoning, but explicit and in fact in the focus of interest in belief revision. So the correspondences between axioms of belief change and properties of nonmonotonic inference operations are usually elaborated only in the case that revisions are based on a fixed theory (cf. (Makinson and Gärdenfors 1991)), and very little work has been done to incorporate iterated belief revision in that framework.

However, belief revision theory is not homogeneous in itself. Genuine belief revision following the AGM-theory (Alchourrón, Gärdenfors, and Makinson 1985) is different from belief update, as defined by Katsuno and Mendelzon (Katsuno and Mendelzon 1991), although it seems to be difficult to draw a clear line between these two change operations. Often, belief revision is understood as the process of adjusting prior beliefs to new information in a static world, whereas belief update should be used to adjust prior beliefs to new information in a possibly changing world. This distinction has raised considerably numerous and deep discussions in the past, and it is not clear how nonmonotonic reasoning as the process of deriving plausible beliefs from given knowledge or beliefs, respectively, can be linked to either change operation.

In this paper, we propose a unifying framework that allows us to understand and explore these relationships more thoroughly. Basically, we will follow the traditional line to relate belief change and nonmonotonic reasoning via the Ramsey test (cf. e.g. (Gärdenfors 1988)) $A \mid\!\sim_{(\Psi)} B \quad \text{iff} \quad \Psi * A \models B$, with Ψ denoting some epistemic state. That is to say that B is plausibly derived from A, given the epistemic background Ψ, iff incorporating A into Ψ yields belief in B. This last statement is often seen to be equivalent to saying that the conditional belief $(B|A)$ is accepted in Ψ, $\Psi \models (B|A)$. As revision strategies, conditionals are of major importance when dealing with iterated belief change.

Therefore, we will explore the relationships between nonmonotonic reasoning and general belief change by considering epistemic states and sets of conditionals instead of theories and propositional beliefs. We will provide a more general framework that not only allows a more accurate representation of belief change via nonmonotonic formalisms, but also gives, vice versa, an important impetus to handle iterated changes. So we will generalize the notion of an inference operation and introduce *universal inference operations* in nonmonotonic logics as a suitable counterpart to (full) change operators. And instead of taking a (propositional) theory as a reference point for inferences and revisions, we will make use of the more comprehensive notion of an epistemic state to base inferences on. In a purely qualitative environment, a preorder might be enough to represent an epistemic state, but one might also choose more sophisticated representation frameworks, such as possibility distributions, ranking functions, and probability distributions. In this paper, Ψ will denote an abstract epistemic state. For the purpose of illustration, we will apply our ideas both in a probabilistic and an ordinal environment.

Leaving the classical framework also allows a more accurate view on iterated change operations by differentiating between *simultaneous* and *successive revision*. The former will be seen to handle genuine revisions appropriately, while the latter may also model *updating*. This distinction is based on clearly separating background or generic knowledge from evidential or contextual knowledge, a feature that is listed in (Dubois and Prade 1996) as one of three basic requirements a *plausible exception-tolerant inference system* has to meet.

The new information the agent is going to incorporate into her beliefs is assumed to provide information about the context the agent is focusing on. Since this context may be complex, the information can be complex as well, consisting of a simple fact, some uncertain evidence, or even conditional information. For instance, being situated in some region that has been devastated by a natural disaster, the agent might find that the rules of daily life have changed drastically, and adopting these new conditional beliefs is essential for him to survive. The new information may encompass different chunks of information all pertaining to the given context, so we assume it to be specified by a set $\mathcal{R}$ containing conditionals as the most general items in our framework that may express information. This includes also facts, since facts are considered as degenerate conditionals with tautological antecedents.

The problems we will be dealing with in this paper can be stated as follows:

- Technically, given some epistemic state Ψ and some set of conditional beliefs $\mathcal{R}$, how to compute $\Psi * \mathcal{R}$ as an adaption of Ψ to new conditional information $\mathcal{R}$?
- Conceptually, how to handle background knowledge as well as prior contextual and new contextual information?
- How to take epistemic background explicitly into account in plausible inference?

The outline of the paper is as follows: In the next section, we will introduce some formal notations for epistemic states and conditionals. For the examples presented in this paper, we recall basic facts on ordinal conditional functions and on the principle of minimum cross entropy in a probabilistic domain. The following section is dedicated to universal inference operations that allow us to deal with nonmonotonic inference operations based on different epistemic backgrounds. Afterwards, we first discuss update as a basic, imperative change operation which is linked to universal inference operations afterwards. Finally, belief revision is dealt with, and its connection to universal inference operations and belief update is elaborated. We conclude with a summary and a brief outlook on future work.

Epistemic states and conditionals

We will use propositions $A, B, \ldots$ as basic atoms from a finite signature Σ to build up a propositional logical language $\mathcal{L}$ with the junctors $\wedge$ and $\neg$. The $\wedge$-junctor will mostly be omitted, so that AB stands for $A \wedge B$, and $\neg$ will usually be indicated by overlining the corresponding proposition, i.e. $\overline{A}$ means $\neg A$. The set of all propositional interpretations over Σ is denoted by Ω_Σ. As the signature will be fixed throughout the paper, we will usually omit the subscript and simply write Ω.

$\mathcal{L}$ is extended to a conditional language $(\mathcal{L} \mid \mathcal{L})$ by introducing a conditional operator $|$: $(\mathcal{L} \mid \mathcal{L}) = \{(B|A) \mid A, B \in \mathcal{L}\}$. $(\mathcal{L} \mid \mathcal{L})$ is a flat conditional language, no nesting of conditionals is allowed.

Conditionals are usually considered within richer semantic structures such as *epistemic states*. Besides certain knowledge, epistemic states also allow the representation of preferences, beliefs, assumptions etc of an intelligent agent. In a purely qualitative setting, epistemic states can be represented by systems of spheres, or simply by a pre-ordering on $\mathcal{L}$ (which is mostly induced by a pre-ordering on worlds). In a (semi-)quantitative setting, also degrees of plausibility, probability, possibility, necessity etc can be expressed. We briefly describe two well-known representations of epistemic states, probability distributions and ordinal conditional functions; both will serve to illustrate the ideas presented in this paper.

Probability distributions in a logical environment can be identified with probability functions $P : \Omega \to [0, 1]$ with $\sum_{\omega \in \Omega} P(\omega) = 1$. The probability of a formula $A \in \mathcal{L}$ is given by $P(A) = \sum_{\omega \models A} P(\omega)$. Since $\mathcal{L}$ is finite, Ω is finite, too, and we only need additivity instead of σ-additivity. Moreover, we assume all subsets of Ω to be measurable. Conditionals are interpreted via conditional probabilities, so that $P(B|A) = \frac{P(AB)}{P(A)}$ for $P(A) > 0$, and $P \models (B|A)\,[x]$ iff $P(A) > 0$ and $P(B|A) = x$ ($x \in [0, 1]$).

Ordinal conditional functions (OCFs), (also called *ranking functions*) $\kappa : \Omega \to \mathbb{N} \cup \{\infty\}$ with $\kappa^{-1}(0) \neq \emptyset$, were introduced first by Spohn (Spohn 1988). They express degrees of plausibility of propositional formulas A by specifying degrees of disbeliefs of their negations $\overline{A}$. More formally, we have $\kappa(A) := \min\{\kappa(\omega) \mid \omega \models A\}$, so that $\kappa(A \vee B) = \min\{\kappa(A), \kappa(B)\}$. Hence, due to $\kappa^{-1}(0) \neq \emptyset$, at least one of $\kappa(A), \kappa(\overline{A})$ must be 0. A proposition A is believed if $\kappa(\overline{A}) > 0$ (which implies particularly $\kappa(A) = 0$). Degrees of plausibility can also be assigned to conditionals by setting $\kappa(B|A) = \kappa(AB) - \kappa(A)$. A conditional $(B|A)$ is accepted in the epistemic state represented by κ, or κ *satisfies* $(B|A)$, written as $\kappa \models (B|A)$, iff $\kappa(AB) < \kappa(A\overline{B})$, i.e. iff AB is more plausible than $A\overline{B}$. OCF's are the qualitative counterpart of probability distributions. Their plausibility degrees may be taken as order-of-magnitude abstractions of probabilities (cf. (Goldszmidt, Morris, and Pearl 1993; Goldszmidt and Pearl 1996)).

Probability distributions will serve to illustrate the fully quantitative framework, while OCFs will only be used as a convenient way of expressing conditional beliefs qualitatively. However, since OCFs inherently use a discrete numerical scale with arithmetics, they provide a more powerful framework for belief change operations than what is possible in a purely qualitative framework.

Both probability distributions and ordinal conditional functions belong to the class of so-called *conditional valuation functions* which were introduced in (Kern-Isberner 2001) to abstract from numbers and reveal more clearly and uniformly the way in which (conditional) knowledge may be represented and treated within epistemic states. In short, conditional valuation functions assign abstract algebraic degrees of certainty, plausibility, possibility etc to propositional formulas and to conditionals. As a crucial feature, they make use of two different operations to distinguish between the handling of purely propositional information and conditionals, respectively. In this way, they provide a framework for considering and treating conditional knowledge as substantially different from propositional knowledge, a point that is stressed by various authors and that seems to be indispensable for representing epistemic states adequately (cf. (Darwiche and Pearl 1997)). However, the close relationship between propositions and conditionals by considering propositions as conditionals of a degenerate form via identifying A with $(A|\top)$ is respected. For more details on conditional valuation functions, please see (Kern-Isberner 2001; 2004).

In the following, let Ψ be any epistemic state, specified e.g. by a preorder, some kind of conditional valuation function, or some other structure that is found appropriate to express conditional beliefs via a suitable conditional language $(\mathcal{L} \mid \mathcal{L})^{\times}$, in which conditionals may be equipped with quantitative degrees of belief or not, according to the chosen framework. For instance, for probability functions, $(\mathcal{L} \mid \mathcal{L})^{\times} = (\mathcal{L} \mid \mathcal{L})^{prob} = \{(B|A)[x] \mid A, B \in \mathcal{L}, x \in [0,1]\}$, and for ordinal conditional functions, $(\mathcal{L} \mid \mathcal{L})^{\times} = (\mathcal{L} \mid \mathcal{L})$. Moreover, there is given an entailment relation $\models$ between epistemic states and conditionals; basically, $\Psi \models (B|A)^{\times}$ means that $(B|A)^{\times}$ is accepted in Ψ. Usually, we will assume that epistemic states are *complete* in that they are uniquely characterized by all conditional beliefs which they accept (up to equivalence).

Let $\mathcal{E} = \mathcal{E}_{\Sigma}$ denote the set of all such epistemic states using $(\mathcal{L} \mid \mathcal{L})^{\times}$ for representation of (conditional) beliefs. For an epistemic state $\Psi \in \mathcal{E}$, the set of its conditional consequences is denoted by

$$Th(\Psi) = \{\phi \in (\mathcal{L} \mid \mathcal{L})^{\times} \mid \Psi \models \phi\}.$$

Moreover, epistemic states are considered as (epistemic) models of sets of conditionals $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$:

$$Mod(\mathcal{R}) = \{\Psi \in \mathcal{E} \mid \Psi \models \mathcal{R}\}.$$

This allows us to extend semantical entailment to sets of conditionals by setting

$$\mathcal{R}_1 \models \mathcal{R}_2 \quad \text{iff} \quad Mod(\mathcal{R}_1) \subseteq Mod(\mathcal{R}_2),$$

and to define a (monotonic) consequence operation $Cn^{\times} : 2^{(\mathcal{L}|\mathcal{L})^{\times}} \to 2^{(\mathcal{L}|\mathcal{L})^{\times}}$ by $Cn^{\times}(\mathcal{R}) = \{\phi \in (\mathcal{L} \mid \mathcal{L})^{\times} \mid \mathcal{R} \models \phi\}$, in analogy to classical consequence. Two sets of conditionals $\mathcal{R}_1, \mathcal{R}_2 \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$ are called *(epistemically) equivalent* iff $Mod(\mathcal{R}_1) = Mod(\mathcal{R}_2)$. As usual, $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$ is *consistent* iff $Mod(\mathcal{R}) \neq \emptyset$, i.e. iff there is an epistemic state representing $\mathcal{R}$.

Belief change in probabilistic and ordinal environments

This section is dedicated to recalling two change operators, one for adapting probability distributions to sets of probabilistic conditionals, and one for changing ordinal conditional functions by new qualitative conditional beliefs. Both operators will serve to illustrate the ideas presented in this paper and are linked by the property that they both satisfy the *principle of conditional preservation*, as specified e.g. in (Kern-Isberner 2001; 2004). This principle makes use of the arithmetic structures underlying probabilities and rankings and allows a very accurate and precise handling of conditional information under belief change. Moreover, it has implications for purely qualitative conditional environments as well, ensuring that all postulates for conditional belief change as stated in (Kern-Isberner 2004) are satisfied. These postulates generalize the postulates proposed by Darwiche and Pearl (Darwiche and Pearl 1997). In (Kern-Isberner 2001; 2004), the term *c-revision* has been introduced for change operators satisfying this principle of conditional preservation. We prefer to use the term *c-change* here, where "c" refers to using conditionals as a major guideline for change. So, both change operators to be used in this paper are c-change operators. Although there is no room to give all details here, the structural similarity between the operators will become obvious (see equations (2) and (3) below).

The principle of minimum cross entropy

Given two probability distributions Q and P, the *cross entropy* between them is defined by

$$R(Q,P) = \sum_{\omega \in \Omega} Q(\omega) \log \frac{Q(\omega)}{P(\omega)}$$

(with $0 \log \frac{0}{0} = 0$ and $Q(\omega) \log \frac{Q(\omega)}{0} = \infty$ for $Q(\omega) \neq 0$) Cross entropy is a well-known information-theoretic measure of dissimilarity between two distributions and has been studied extensively (see, for instance, (Csiszár 1975; Hajek, Havranek, and Jirousek 1992; Jaynes 1983a; Kullback 1968); for a brief, but informative introduction and further references, cf. (Shore 1986); see also (Shore and Johnson 1981)). Cross entropy is also called *directed divergence* since it lacks symmetry, i.e. $R(Q,P)$ and $R(P,Q)$ differ in general, so it is not a metric. But cross entropy is *positive*, that means we have $R(Q,P) \geqslant 0$, and $R(Q,P) = 0$ iff $Q = P$ (cf. (Csiszár 1975; Hajek, Havranek, and Jirousek 1992; Shore 1986)).

Consider the probabilistic belief revision problem

($*_{\mathbf{prob}}$) Given a (prior) distribution P and some set of probabilistic conditionals $\mathcal{R} = \{(B_1|A_1)[x_1], \ldots, (B_n|A_n)[x_n]\} \subseteq (\mathcal{L} \mid \mathcal{L})^{prob}$, how should P be modified to yield a (posterior) distribution P^* with $P^* \models \mathcal{R}$?

When solving ($*_{prob}$), the paradigm of *informational economy*, i.e. of minimal loss of information (see (Gärdenfors 1988, p. 49)), is realized in an intuitive way by following the *principle of minimum cross entropy*

$$\min R(Q,P) = \sum_{\omega \in \Omega} Q(\omega) \log \frac{Q(\omega)}{P(\omega)} \quad (1)$$

$$\text{s.t. } Q \text{ is a probability distribution with } Q \models \mathcal{R}$$

For a distribution P and some set $\mathcal{R}$ of probabilistic conditionals compatible with P, there is a (unique) distribution

$P_{ME} = P_{ME}(P, \mathcal{R}) = P *_{ME} \mathcal{R}$ that satisfies $\mathcal{R}$ and has minimal cross entropy to the prior P (cf. (Csiszár 1975)), i.e. P_{ME} solves (1) and thereby ($*_{prob}$). Note that ($*_{prob}$) exceeds the framework of the classical AGM-theory with regard to several aspects: an *epistemic state* (P) is to be revised by a *set of conditionals* representing *uncertain knowledge*.

The principle of minimum cross entropy is a very powerful change methodology, though a bit opaque when considered only as a solution to an optimization problem. The conditional logical structure underlying the *ME*-revision is revealed better by the following representation that can be derived by the usual Lagrange techniques (see, for instance, (Jaynes 1983b)) in a straightforward way:

$$P *_{ME} \mathcal{R}(\omega) = \alpha_0 P(\omega) \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^{1-x_i} \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^{-x_i}, \quad (2)$$

with the α_i's being exponentials of the Lagrange multipliers, one for each conditional in $\mathcal{R}$, and have to be chosen properly to ensure that $P *_{ME} \mathcal{R}$ satisfies all conditionals in $\mathcal{R}$. α_0 is simply a normalizing factor.

Ordinal c-change

An *ordinal c-change operator* $*$ returns for each ranking function κ and each consistent set $\mathcal{R}$ of conditionals a ranking function $\kappa * \mathcal{R}$ that satisfies the principle of conditional preservation, as specified in (Kern-Isberner 2001; 2004). In these papers, also a handy characterization of ordinal c-change operators is provided:

A change operation $\kappa^* = \kappa * \mathcal{R}$ is an ordinal c-change operation iff there are numbers $\kappa_0, \kappa_i^+, \kappa_i^- \in \mathbb{Q}, 1 \leqslant i \leqslant n$, such that, for all $\omega \in \Omega$,

$$\kappa^*(\omega) = \kappa_0 + \kappa(\omega) + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \kappa_i^+ + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \kappa_i^-. \quad (3)$$

C-change operations can be defined for any finitely valued OCF κ and any consistent set $\mathcal{R}$ of conditionals; if κ also takes on infinite values, some basic demands for compatibility between κ and $\mathcal{R}$ have to be observed (cf. (Kern-Isberner 2001)). So, c-change operators can be obtained easily by using the schema provided by equation (3), and the requirement $\kappa^* \models \mathcal{R}$ yields the following conditions for κ_i^+, κ_i^- in a straightforward way:

$$\kappa_i^- - \kappa_i^+ > \min_{\omega \models A_i B_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j B_j}} \kappa_j^+ + \sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-) \quad (4)$$
$$- \min_{\omega \models A_i \overline{B}_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j B_j}} \kappa_j^+ + \sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-)$$

This can be simplified considerably by setting $\kappa_i^+ := 0$ and focussing only on the κ_i^-, preferably choosing them in a minimal way. However, as has been observed by several authors (cf. e.g. (Goldszmidt, Morris, and Pearl 1993; Bourne and Parsons 1999)), more than one minimal solution may exist. So, we have to constraint the choice of the κ_i^- in a reasonable way to ensure the uniqueness of the change result.

Supposing that $\mathcal{R}$ is consistent, a tolerance partitioning $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_k$ of $\mathcal{R}$ may be computed, as is done for system Z (Goldszmidt and Pearl 1996). For all conditionals $r_i \in \mathcal{R}$, $1 \leqslant i \leqslant n$, set $\kappa_i^+ := 0$, and set successively, for each partitioning set $\mathcal{R}_m$, $0 \leqslant m \leqslant k$, starting with $\mathcal{R}_0$, and for each conditional $r_i = (B_i|A_i) \in \mathcal{R}_m$

$$\kappa_i^- := 1 + \max \{ \min_{\substack{\omega \models A_i B_i \\ r(\omega) \neq 0 \forall r \in \cup_{l=m}^k \mathcal{R}_l}} (\kappa(\omega) + \sum_{\substack{r_j \in \cup_{l=0}^{m-1} \mathcal{R}_l \\ \omega \models A_j \overline{B}_j}} \kappa_j^-)$$
$$-\kappa(A_i \overline{B}_i), \ -1\} \quad (5)$$

Finally, choose κ_0 appropriately to make $\kappa^*(\omega) = \kappa_0 + \kappa(\omega) + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_i^-$ an ordinal conditional function. It is straightforward to check that indeed, $\kappa^* \models \mathcal{R}$, so κ^* is a c-change of κ by $\mathcal{R}$. We will call the so-defined ordinal change operator a *simple c-change operator*.

We will illustrate both methodologies later on, when discussing belief change operations in more detail. By applying the c-change approach to uniform epistemic states, i.e. the uniform distribution for probabilities and the uniform ranking $\kappa_u(\omega) = 0$ for all $\omega \in \Omega$, we obtain quite easily very well-behaved inductive inference operations of incomplete conditional knowledge bases. This connection between belief change and nonmonotonic inference will be elaborated in the following sections.

Before discussing belief revision and update in a more general epistemic and conditional environment, we will first develop a framework for nonmonotonic inference operations using epistemic states as background knowledge.

Universal inference operations

From quite a general, abstract point of view, inference operations C map sets of formulas to sets of formulas – given a set of formulas, C is to return which formulas can be derived from this set with respect to some classical or commonsense logic. In this paper, conditionals are assumed to be a basic, most general logical means to express knowledge or beliefs. Therefore, we will consider *(conditional) inference operations* $C : 2^{(\mathcal{L}|\mathcal{L})^\times} \to 2^{(\mathcal{L}|\mathcal{L})^\times}$ associating with each set of conditionals a set of inferred conditionals. This also covers the notion of propositional inference operations, taking propositional facts as degenerated conditionals. A straightforward example for a conditional inference operation is given by the classical operation $Cn^\times$ defined above. In order to link conditional inference to epistemic states, complete inference operations are of particular interest.

Definition 1. A conditional inference operation C is called *complete* iff it specifies for each consistent set $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$ a complete epistemic state $\Psi_\mathcal{R}$, i.e. iff there is an epistemic state $\Psi_\mathcal{R}$ such that $C(\mathcal{R}) = Th(\Psi_\mathcal{R})$.

Complete inference operations realize inductive, model based inference, i.e. they make the information given by some (consistent) set of conditionals complete. Well-known examples of complete conditional inference operations are induced by system Z and system Z^* (Goldszmidt and Pearl 1996) for ordinal epistemic states, and the principle of maximum entropy (Paris and Vencovska 1997) for probabilistic epistemic states.

As we are going to study inference operations which are based on *different* epistemic backgrounds, we make this background explicit and define a formal structure to cover all possible backgrounds, i.e. all epistemic states in $\mathcal{E}$.

Definition 2. A *universal inference operation* $\mathbf{C}$ assigns a complete conditional inference operation $C_\Psi : 2^{(\mathcal{L}|\mathcal{L})^\times} \to 2^{(\mathcal{L}|\mathcal{L})^\times}$ to each epistemic state $\Psi \in \mathcal{E}$: $\mathbf{C} : \Psi \mapsto C_\Psi$. $\mathbf{C}$ is said to be *reflexive (idempotent, cumulative)* iff all its involved inference operations have the corresponding property.

If $\mathbf{C} : \Psi \mapsto C_\Psi$ is a universal inference operation, C_Ψ is complete for each $\Psi \in \mathcal{E}$. That means, for each consistent set $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$, $C_\Psi(\mathcal{R})$ specifies completely (up to equivalence) an epistemic state $\Phi_{\Psi,\mathcal{R}}$:

$$C_\Psi(\mathcal{R}) = Th(\Phi_{\Psi,\mathcal{R}}) \qquad (6)$$

Define the set of all such epistemic states by $\mathcal{E}(C_\Psi) = \{\Phi \in \mathcal{E} \mid \exists \mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times : C_\Psi(\mathcal{R}) = Th(\Phi)\}$.

We are now going to study some interesting properties of universal inference operations.

Definition 3. A universal inference operation $\mathbf{C}$ is *founded* iff for each epistemic state Ψ and for any $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$, $\Psi \models \mathcal{R}$ implies $C_\Psi(\mathcal{R}) = Th(\Psi)$.

The property of foundedness establishes a close and intuitive relationship between an epistemic state Ψ and its associated inference operation C_Ψ, distinguishing Ψ as its stable starting point. In particular, if $\mathbf{C}$ is founded then $C_\Psi(\emptyset) = Th(\Psi)$. As to the universal inference operation $\mathbf{C}$, foundedness ensures injectivity, as can be proved easily.

Proposition 1. If $\mathbf{C}$ is founded, then it is injective.

In standard nonmonotonic reasoning, as it was developed in (Makinson 1994), cumulativity occupies a central and fundamental position, claiming the inferences from a set $\mathcal{S}$ that "lies in between" another set $\mathcal{R}$ and its nonmonotonic consequences $C(\mathcal{R})$ to coincide with $C(\mathcal{R})$.

To establish a similar well-behavedness of $\mathbf{C}$ with respect to *epistemic states*, we introduce suitable relations to compare epistemic states with one another.

Definition 4. Let $\mathbf{C} : \Psi \mapsto C_\Psi$ be a universal inference operation. For each epistemic state Ψ, define a relation $\sqsubseteq_\Psi$ on $\mathcal{E}(C_\Psi)$ by setting $\Phi_1 \sqsubseteq_\Psi \Phi_2$ iff there are sets $\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq (\mathcal{L} \mid \mathcal{L})^\times$ such that $Th(\Phi_1) = C_\Psi(\mathcal{R}_1)$ and $Th(\Phi_2) = C_\Psi(\mathcal{R}_2)$

So, $\Phi_1 \sqsubseteq_\Psi \Phi_2$ iff, based on the same background knowledge Ψ, Φ_2 makes use of a larger knowledge base than Φ_1.

For founded universal inference operations, we have in particular $C_\Psi(\emptyset) = Th(\Psi)$ for all $\Psi \in \mathcal{E}$, so Ψ is a minimal element of $\mathcal{E}(C_\Psi)$ with respect to $\sqsubseteq_\Psi$:

Proposition 2. If $\mathbf{C}$ is founded, then for all $\Psi \in \mathcal{E}$ and for all $\Phi \in \mathcal{E}(C_\Psi)$, it holds that $\Psi \sqsubseteq_\Psi \Phi$.

We will now generalize the notion of cumulativity for universal inference relations:

Definition 5. A universal inference operation $\mathbf{C}$ is called *strongly cumulative* iff for each $\Psi \in \mathcal{E}$ and for any epistemic states $\Phi_1, \Phi_2 \in \mathcal{E}(C_\Psi)$, $\Psi \sqsubseteq_\Psi \Phi_1 \sqsubseteq_\Psi \Phi_2$ implies: whenever $\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq (\mathcal{L} \mid \mathcal{L})^\times$ such that $Th(\Phi_1) = C_\Psi(\mathcal{R}_1)$ and $Th(\Phi_2) = C_\Psi(\mathcal{R}_2)$, then $Th(\Phi_2) = C_\Psi(\mathcal{R}_2) = C_{\Phi_1}(\mathcal{R}_2)$.

Strong cumulativity describes a relationship between inference operations based on different epistemic states, thus linking up the inference operations of $\mathbf{C}$. In the definition above, Φ_1 is an epistemic state intermediate between Ψ and Φ_2, with respect to the relation $\sqsubseteq_\Psi$, and strong cumulativity claims that the inferences based on Φ_1 coincide with the inferences based on Ψ within the scope of Φ_2.

The next proposition is immediate:

Proposition 3. Let $\mathbf{C}$ be a universal inference operation which is strongly cumulative. Suppose $\Psi \in \mathcal{E}$, $\Phi \in \mathcal{E}(C_\Psi)$ such that $Th(\Phi) = C_\Psi(\mathcal{R}), \mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$. Then $C_\Psi(\mathcal{S}) = C_\Phi(\mathcal{S})$ for any $\mathcal{S} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$ with $\mathcal{R} \subseteq \mathcal{S}$.

The following theorem justifies the name "strong cumulativity": It states that strong cumulativity actually generalizes cumulativity for an important class of universal inference operations:

Theorem 1. If $\mathbf{C}$ is founded, then strong cumulativity implies cumulativity.

Universal inference operations will prove to be an adequate formal counterpart of iterated change operations on the nonmonotonic reasoning side. Before elaborating this in more detail, we will first develop a comparable formal machinery for belief change.

Updating epistemic states by conditional beliefs

Basically, from our point of view, the difference between revision and updating is mainly due to different belief change scenarios, using different information. This implies that the change process should be distinguished from scenario assumptions that are often mixed with the pure change process. For instance, Lang (Lang 2007) stresses the importance of recognizing the nature of the new information, whether it is an observation or an action effect, for telling update and revision apart. Hunter and Delgrande (Hunter and Delgrande 2005) distinguish between ontic and epistemic actions to make differences. From the point of view we take in this paper, such important details on the new information should be represented explicitly, and belief change operators should be able to handle such annotations. Furthermore, world changes are not only brought about by actions, but also, in a more general way, by events.

Moreover, belief update is often thought of as an operator that should be able to bring about more rigid changes. For instance, in (Lang, Marquis, and Williams 2001), updating is realized by making use of a forgetting operation that cancels any dependencies between involved variables. This is a reasonable approach as far as contingent beliefs are concerned, e.g. if either the book or the magazine is on the floor accidentally. However, update is not expected to change everything, coherent prior beliefs should be respected and not given up without reason. The focus of this paper is more on these coherence aspects of the world that may be modified under evolution but are seldom completely lost. Hence, the ideas presented here aim at clarifying more foundational points on the nature of change processes, with a motivation similar to the work by Dubois (Dubois 2006).

So, first of all, we think of $*$ to be a basic imperative change operator that solves the technical belief change problem successfully:

$$\Psi * \mathcal{R} \in \mathcal{E} \text{ such that } \Psi * \mathcal{R} \models \mathcal{R}. \qquad (7)$$

It is crucial to point out that $*$ is a full change operator taking *two* entries, namely an epistemic state Ψ on its left and a (compatible and consistent) set of conditionals on its right. Beyond success, what can be expected from $*$? Since the epistemic state Ψ is assumed to be completely characterized by its set of accepted conditional beliefs, $\mathcal{R}$ can be consistent with Ψ only in a trivial way, i.e., in case of $\Psi \models \mathcal{R}$.

Due to this completeness of knowledge we demand for representing epistemic states, there is few or no room, respectively, to model ignorance. To check whether new information is consistent with an epistemic state thus generally comes down to check whether this information is already represented. On the other hand, this implies that incorporating $\mathcal{R}$ might change lots of (factual) beliefs. In a probabilistic framework, for instance, changing a distribution so as to assimilate new information will usually change every single atomic probability (even if the change operation is as simple as conditioning). The propositional AGM view of obeying minimal change principles which are based on set inclusion does not make much sense here. This has been recognized already when dealing with conditional beliefs under revision (cf. e.g. (Darwiche and Pearl 1997)). We would rather need some kind of distance measure between epistemic states that helps us finding posterior epistemic states closest to prior epistemic states. In the probabilistic framework, cross entropy (cf. e.g. (Jaynes 1983a)) is a proper candidate for that, and similar ideas could be realized in other frameworks, too.

However, there still should be some qualitative close connection between prior and posterior epistemic state. The key idea here is to postulate that the reasoning structures underlying Ψ should also be effective in $\Psi * \mathcal{R}$ to a largest possible degree, that is, if no new information in $\mathcal{R}$ force them to change. Since in our framework, such reasoning structures are implemented by conditionals, we should focus on conserving conditional beliefs, hence following a line of thought similar to that of Darwiche and Pearl in their work on iterated belief revision (Darwiche and Pearl 1997). In several papers, we have made precise how a *principle of conditional preservation* can be realized in different frameworks (Kern-Isberner 1999; 2004).

Here in this paper, we generalize these ideas to the meta level of change itself, proposing reasonable postulates for an *ideal operator* $*$ changing epistemic states by sets of conditionals. The attribute *ideal* is to be taken in several meanings: First, as we do not stick to one representation framework, the postulates are quite general. Second, we are willing to accept that not all of them can be fulfilled in any framework, nevertheless suggesting them as postulates a "good" change operator should strive for to satisfy. Third, even in frameworks where these postulates do not fully apply, we might think of some ideal change operator that the operator under investigation can be based upon (e.g. as a projection), hence deriving from the postulates reasonable properties of belief change in simpler frameworks.

Postulates for updating epistemic states by sets of conditionals:
Let Ψ be an epistemic state, and let $\mathcal{R} \subset (\mathcal{L} \mid \mathcal{L})^{\times}$; let $\Psi * \mathcal{R}$ denote the result of updating Ψ by $\mathcal{R}$:

(CSR1) Success: $\Psi * \mathcal{R} \models \mathcal{R}$.

(CSR2) Stability: If $\Psi \models \mathcal{R}$ then $\Psi * \mathcal{R} = \Psi$.

(CSR3) Semantical equivalence: If $\mathcal{R}_1$ and $\mathcal{R}_2$ are (semantically) equivalent, then $\Psi * \mathcal{R}_1 = \Psi * \mathcal{R}_2$.

(CSR4) Reciprocity: If $\Psi * \mathcal{R}_1 \models \mathcal{R}_2$ and $\Psi * \mathcal{R}_2 \models \mathcal{R}_1$ then $\Psi * \mathcal{R}_1 = \Psi * \mathcal{R}_2$.

(CSR5) Logical coherence: $\Psi * (\mathcal{R}_1 \cup \mathcal{R}_2) = (\Psi * \mathcal{R}_1) * (\mathcal{R}_1 \cup \mathcal{R}_2)$.

Postulates (CSR1) implements *success*, a crucial property of an imperative change operator. (CSR2) guarantees *stability* if the information represented by $\mathcal{R}$ is not new but already believed. As *success*, also *stability* might be debatable in cases where we would like new confirming information to have a strengthening effect. Again, this is not a matter of pure change processes, but involves different processes that might better be discussed in a merging framework. (CSR3) means that for the basic change operation, only the semantics of new pieces of information matter. This should, however, not be confused with the idea of *syntax independence* which is sometimes used to indicate refraining from using explicit beliefs in belief bases. (CSR4) states that two changing procedures with respect to sets $\mathcal{R}_1$ and $\mathcal{R}_2$ should result in the same epistemic state if each revision represents the new information of the other. This property is called *reciprocity* in the framework of nonmonotonic logics (cf. (Makinson 1994)) and appears as axiom (U6) in the work of Katsuno and Mendelzon (Katsuno and Mendelzon 1991). (CSR5) is the only seemingly extraordinary and new axiom here. It demands that adjusting any intermediate epistemic state $\Psi * \mathcal{R}_1$ to the full information $\mathcal{R}_1 \cup \mathcal{R}_2$ should result in the same epistemic state as adjusting Ψ by $\mathcal{R}_1 \cup \mathcal{R}_2$ in one step. The rationale behind this axiom is that if the information about the new world drops in in parts, changing any intermediate state of belief by the full information should result unambigously in a final belief state. So, it guarantees the change process to be *logically coherent*. (CSR5) goes back to (Shore and Johnson 1981) where it was used to describe the change process induced by the principle of minimum cross entropy. In (Kern-Isberner 1998), it plays a crucial role for characterizing this probabilistic principle. But so far, to the best of our knowledge, it has not been recognized as an important postulate for iterated belief change in general, although it explicitly links iterated belief change to one-step belief change. Interestingly, by reducing the complex conditional inputs to plain propositions, we obtain axiom (C1) in (Darwiche and Pearl 1997). When presupposing success and stability, (CSR5) implies (CSR4). However, we prefer to list both postulates here, as (CSR5) is very hard to satisfy in full generality.

Note that (CSR5) does not claim $(\Psi * \mathcal{R}_1) * \mathcal{R}_2$ and $(\Psi * \mathcal{R}_1) * (\mathcal{R}_1 \cup \mathcal{R}_2)$ to be the same, so it does not boil down iterated belief change to simple belief change. Just to the contrary – these epistemic states will be expected to differ in general, because the first is not supposed to maintain prior contextual information, $\mathcal{R}_1$, whereas the second should do so, according to axiom (CSR1).

We call such a basic imperative change operator an update operator, because it should be able to bring about more thorough changes, which is necessary in an evolving world. No easy comparison with the update operator by Katsuno and Mendelzon (Katsuno and Mendelzon 1991) seems to be possible, except for saying that changes induced by our update operator may affect each (classical) model, or its appertaining degree of belief, respectively, just as each model is taken into regard by update according to Katsuno and Mendelzon.

In (Kern-Isberner 2001) it is shown, that the probabilistic update operation via the principle of minimum cross entropy satisfy all postulates (CSR1) to (CSR5). The simple ordinal c-change operator defined above only satisfies (CSR1) to (CSR3). We will give examples for the application of both change operators later on, when discussing update and revision operations in one framework. But first, we will link belief update to universal inference operations.

Universal inference operations and belief update

A straightforward relationship between universal inference operations $\mathbf{C} : \Psi \mapsto C_\Psi$ and binary update operators $*$ can now be established by setting $\Psi * \mathcal{R} \equiv \Phi_{\Psi,\mathcal{R}}$ (cf. (6) above) for $\Psi \in \mathcal{E}, \mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$, that is

$$C_\Psi(\mathcal{R}) = Th(\Psi * \mathcal{R}). \quad (8)$$

We will use this relationship as a formal vehicle to make correspondences between the postulates for updates and properties of universal inference operations explicit.

Proposition 4. Suppose update is being realized via a universal inference operation as in (8). Then the following coimplications hold:

(i) $*$ satisfies (CSR1) iff $\mathbf{C}$ is reflexive.

(ii) $*$ satisfies (CSR2) iff $\mathbf{C}$ is founded.

(iii) $*$ satisfies (CSR3) iff $\mathbf{C}$ satisfies left logical equivalence.

(iv) Assuming reflexivity resp. the validity of (CSR1), $*$ satisfies (CSR4) iff $\mathbf{C}$ is cumulative.

(v) Assuming foundedness resp. the validity of (CSR2), $*$ satisfies (CSR5) iff $\mathbf{C}$ is strongly cumulative.

The proofs are immediate. From this proposition, a representation result follows in a straightforward manner:

Theorem 2. If $*$ is defined by (8), it satisfies all of the postulates (CSR1)-(CSR5) iff the universal inference operation $\mathbf{C}$ is reflexive, founded, strongly cumulative and satisfies left logical equivalence.

So, in particular, the properties of foundedness and strong cumulativity turn out to be crucial to control iterated change operations.

Belief bases and belief revision

Finally, we will turn our attention to belief revision. The usual view on belief revision is that currently held beliefs are revised by new information, where conflicts are solved in favor of the new information. If we presuppose that the revision agent is neither confused nor deceived or stupid, the observations he makes are correct, and conflicts may only arise between new information and plausibly derived beliefs.

For instance, if we hear in the news that a plane crashed, we are shocked and believe that many people are dead or badly hurt. However, we might not react in a paranoid way, but believe that the crash was an accident and no terrorist attack. Fortunately, like a miracle, the plane could land on a field and did not catch fire, so we now hope that most of the passengers could be saved. But later on, we read in the newspaper that a mysterious letter has been found referring to the crash, so it seems plausible now that the crash was caused indeed by some attack. It is obvious that revision processes took place in this example, dismissing prior beliefs and adopting new beliefs, but the pieces of information that we obtained concerning the plane crash were consistent with one another. However, each new information triggered plausible inferences which might have to be revised later on.

So, we are going to distinguish between background knowledge and evidential knowledge, and all evidential information the agent collects about the "static world" under consideration – which defines the relevant *context* – must be consistent.

Revision will be considered in epistemic states Ψ' that reveal some history by telling background knowledge Ψ and contextual information $\mathcal{R}$ apart: $\Psi' = \Psi * \mathcal{R}$, where the imperative change operator $*$ is used to adopt background knowledge to context. In order to handle both parts of the epistemic state appropriately, we first introduce belief bases that can be used to build up the epistemic state.

Definition 6. A *belief base* is a pair $(\Psi, \mathcal{R})$, where Ψ is an epistemic state (*background knowledge*), and $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})^\times$ is a set of conditionals representing *contextual* (or *evidential*) *knowledge*.

The transition from belief bases to epistemic states is then achieved by a binary belief update operator $*$:

$$*(\Psi, \mathcal{R}) := \Psi * \mathcal{R}. \quad (9)$$

So, flux of knowledge is modelled quite naturally: Prior knowledge serves as a base for obtaining an adequate full description of the present context which may be used again as background knowledge for further change operations.

In the following, we will develop postulates for revising belief bases $(\Psi, \mathcal{R})$ by new conditional information $\mathcal{S} \subseteq$

$(\mathcal{L} \mid \mathcal{L})^{\times}$, yielding a new belief base $(\Psi, \mathcal{R}) \circ \mathcal{S}$, in the sense of the AGM-postulates.

Due to distinguishing background knowledge from context information, we are able to compare the knowledge stored in different belief bases:

Definition 7. A pre-ordering $\sqsubseteq$ on belief bases is defined by $(\Psi_1, \mathcal{R}_1) \sqsubseteq (\Psi_2, \mathcal{R}_2)$ iff $\Psi_1 = \Psi_2$ and $\mathcal{R}_2 \models \mathcal{R}_1$ $(\Psi_1, \mathcal{R}_1)$ and $(\Psi_2, \mathcal{R}_2)$ are $\sqsubseteq$-*equivalent*, $(\Psi_1, \mathcal{R}_1) \equiv_{\sqsubseteq} (\Psi_2, \mathcal{R}_2)$, iff $(\Psi_1, \mathcal{R}_1) \sqsubseteq (\Psi_2, \mathcal{R}_2)$ and $(\Psi_2, \mathcal{R}_2) \sqsubseteq (\Psi_1, \mathcal{R}_1)$.

Therefore $(\Psi_1, \mathcal{R}_1) \equiv_{\sqsubseteq} (\Psi_2, \mathcal{R}_2)$ iff $\Psi_1 = \Psi_2$ and $\mathcal{R}_1$ and $\mathcal{R}_2$ are semantically equivalent, i.e. iff both belief bases reflect the same epistemic (background and contextual) knowledge.

The following postulates do not make use of the basic change operation $*$, but are to characterize pure belief base revision by the revision operator $\circ$:

Postulates for conditional base revision:

Let Ψ be an epistemic state, and let $\mathcal{R}, \mathcal{R}_1, \mathcal{S} \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$ be sets of conditionals.

(CBR1) $(\Psi, \mathcal{R}) \circ \mathcal{S}$ is a belief *base*.

(CBR2) If $(\Psi, \mathcal{R}) \circ \mathcal{S} = (\Psi, \mathcal{R}_1)$ then $\mathcal{R}_1 \models \mathcal{S}$.

(CBR3) $(\Psi, \mathcal{R}) \sqsubseteq (\Psi, \mathcal{R}) \circ \mathcal{S}$.

(CBR4) $(\Psi, \mathcal{R}) \circ \mathcal{S}$ is a minimal belief base (with respect to $\sqsubseteq$) among all belief bases satisfying (PR1)-(PR3).

(CBR1) is the most fundamental axiom and coincides with the demand for *categorical matching*. (CBR2) is the *success postulate* here: the new context information is now represented (up to epistemic equivalence). (CBR3) states that revision should preserve prior *contextual* knowledge. Thus it is crucial for revision in contrast to updating. Finally, (CBR4) is in the sense of *informational economy*: No unnecessary changes should occur. The following characterization may be proved easily:

Theorem 3. The revision operator $\circ$ satisfies the axioms (CBR1) – (CBR4) iff

$$(\Psi, \mathcal{R}) \circ \mathcal{S} \equiv_{\sqsubseteq} (\Psi, \mathcal{R} \cup \mathcal{S}). \tag{10}$$

So, from (CBR1)-(CBR4), other properties of the revision operator also follow in a straightforward manner which are usually found among characterizing postulates:

Proposition 5. Suppose the revision operator $\circ$ satisfies (10). Then it fulfills the following properties:

(i) If $\mathcal{R} \models \mathcal{S}$, then $(\Psi, \mathcal{R}) \circ \mathcal{S} \equiv_{\sqsubseteq} (\Psi, \mathcal{R})$;

(ii) If $(\Psi_1, \mathcal{R}_1) \sqsubseteq (\Psi_2, \mathcal{R}_2)$ then $(\Psi_1, \mathcal{R}_1) \circ \mathcal{S} \sqsubseteq (\Psi_2, \mathcal{R}_2) \circ \mathcal{S}$;

(iii) $((\Psi, \mathcal{R}) \circ \mathcal{S}_1) \circ \mathcal{S}_2 \equiv_{\sqsubseteq} (\Psi, \mathcal{R}) \circ (\mathcal{S}_1 \cup \mathcal{S}_2)$,

where $(\Psi, \mathcal{R}), (\Psi_1, \mathcal{R}_1), (\Psi_2, \mathcal{R}_2)$ are belief bases and $\mathcal{S}, \mathcal{S}_1, \mathcal{S}_2 \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$.

(i) shows a minimality of change, while (ii) is stated in (Gärdenfors 1988) as a monotonicity postulate. (iii) deals with the handling of non-conflicting iterated revisions.

Note that revising a belief base $(\Psi, \mathcal{R})$ by $\mathcal{S} \subseteq (\mathcal{L} \mid \mathcal{L})^{\times}$ also induces a change of the corresponding belief state $\Psi^* = \Psi * \mathcal{R}$ to $(\Psi^*)' = *((\Psi, \mathcal{R}) \circ \mathcal{S})$. So, revision of epistemic states is realized here by making use of base revision and update, or universal inference operations, respectively. According to Theorem 3, if the involved update operation satisfies (CSR3), i.e. if the (underlying) universal inference operation **C** satisfies left logical equivalence, then the only reasonable revision operation (as specified by (CBR1)-(CBR4)) is given on the level of epistemic states by

$$*((\Psi, \mathcal{R}) \circ \mathcal{S}) = \Psi * (\mathcal{R} \cup \mathcal{S}), \tag{11}$$

and therefore, $Th(*((\Psi, \mathcal{R}) \circ \mathcal{S})) = C_{\Psi}(\mathcal{R} \cup \mathcal{S})$. This means that the revision of $\Psi * \mathcal{R}$ by $\mathcal{S}$ is given by $\Psi * (\mathcal{R} \cup \mathcal{S})$. Hence, the epistemic status that $\Psi * \mathcal{R}$ assigns to (conditional) beliefs occurring in $\mathcal{S}$ will differ from those in Ψ as well as from those in $\Psi * (\mathcal{R} \cup \mathcal{S})$ with expanded contextual knowledge. So, the degree of belief of the conditionals in $\mathcal{S}$ is indeed *revised*.

Revision and update

Investigating belief change in the generalized framework of epistemic states and conditionals provides us with deeper insights into the mechanisms underlying the belief change process. As a crucial difference to propositional belief change, it is possible to distinguish between *simultaneous* and *successive* change of epistemic states: In general, we have

$$\Psi * (\mathcal{R} \cup \mathcal{S}) \neq (\Psi * \mathcal{R}) * \mathcal{S}, \tag{12}$$

instead, we may only postulate strong cumulativity or logical coherence, respectively, $\Psi * (\mathcal{R} \cup \mathcal{S}) = (\Psi * \mathcal{R}) * (\mathcal{R} \cup \mathcal{S})$, which is essentially weaker.

The distinction between revision and update can now be made clear on a conceptual level. Revision is to process pieces of contextual information $\mathcal{R}_1, \ldots, \mathcal{R}_n$ pertaining to the same background knowledge Ψ, e.g. the epistemic state of the agent at a given time. Hence, revision should be performed by simultaneous belief change. A proper example for this is the process of making up a consistent mental picture on some event which marks a clear point on the time line. Any reliable information an agent obtains on that event clearly has to be handled on the same level. So, if he receives two pieces of information, $\mathcal{R}_1$ and $\mathcal{R}_2$, on that event, one after the other, then he should revise $\Psi * \mathcal{R}_1$ by $\mathcal{R}_2$ to obtain $\Psi * (\mathcal{R}_1 \cup \mathcal{R}_2)$. On the other hand, update is a kind of successive belief change which is able to override any previously held beliefs, just using the current epistemic state as background knowledge. In the scenario with Ψ as a starting point and two pieces of information, $\mathcal{R}_1$ and $\mathcal{R}_2$, coming in one after the other and each referring to the respective current context, the agent should update $\Psi * \mathcal{R}_1$ by $\mathcal{R}_2$ to obtain $(\Psi * \mathcal{R}_1) * \mathcal{R}_2$, which is, as we pointed out, in general different from $\Psi * (\mathcal{R}_1 \cup \mathcal{R}_2)$. The distinction between update and revision thus becomes primarily a question of making a proper decision how to handle the available information, not of techniques.

This makes clear that in our conceptual framework of belief change, incorporating several pieces of new information into our stock of belief can be achieved in different ways by using the same change operator. This allows a more accurate view on iterated belief change by differentiating between simultaneous and successive revision. This means, having to deal with different pieces of information, the crucial question is not whether one information is more recent than others, but which pieces of information should be considered to be on the same level (which may, but is not restricted to be, of temporal type), or to be more precise, which pieces of information refer to the same context. Basically, informations on the same level are assumed to be compatible with one another, so simple set union will return a consistent set of formulas. Otherwise, more complex merging operations have to be considered. This could be handled in our framework as well by first merging pieces of information equipped with different reliabilities into one consistent rule base which then will serve as input to the change process. Informations on different levels do not have to be consistent, here later ones may override those on previous levels.

Finally, we will try to get a clearer view on formal parallels and differences, respectively, between revision and updating. For an adequate comparison, we have to observe the changes of epistemic *states* that are induced by revision of belief *bases*. Observing (9), (CBR2) and (CBR3) translate into

(CBR2') $*((\Psi, \mathcal{R}) \circ \mathcal{S}) \models \mathcal{S}$.

(CBR3') $*((\Psi, \mathcal{R}) \circ \mathcal{S}) \models \mathcal{R}$.

While (CBR2') parallels (CSR1), (CBR3') establishes a crucial difference between revision and updating: revision is supposed to preserve prior contextual knowledge while updating does not, neither in a classical nor in a generalized framework.

The intended effects of revision and updating on an epistemic state $\Psi * \mathcal{R}$ that is generated by a belief base $(\Psi, \mathcal{R})$ are made obvious by – informally! – writing

$$(\Psi * \mathcal{R}) \circ \mathcal{S} = \Psi * (\mathcal{R} \cup \mathcal{S}) \neq (\Psi * \mathcal{R}) * \mathcal{S} \qquad (13)$$

(cf. (11)). This reveals clearly the difference, but also the relationship between revision and updating: Revising $\Psi * \mathcal{R}$ by $\mathcal{S}$ results in the same epistemic state as updating Ψ by (the full contextual information) $\mathcal{R} \cup \mathcal{S}$.

The representation of an epistemic state by a belief base, however, is not unique, different belief bases may generate the same epistemic state (the same holds for classical belief bases). So we could not define genuine epistemic revision on belief states, but had to consider belief bases in order to separate background and context knowledge unambigously. It is interesting to observe, however, that the logical coherence property (CSR5) of the basic change operator $*$ ensures at least a convenient independence of revision from background knowledge: If two belief bases $(\Psi_1, \mathcal{R}), (\Psi_2, \mathcal{R})$ with different prior knowledge but the same contextual knowledge give rise to the same belief state $\Psi_1 * \mathcal{R} = \Psi_2 * \mathcal{R}$, then, assuming logical coherence to hold, $\Psi_1 * (\mathcal{R} \cup \mathcal{S}) = (\Psi_1 * \mathcal{R}) * (\mathcal{R} \cup \mathcal{S}) = (\Psi_2 * \mathcal{R}) * (\mathcal{R} \cup \mathcal{S}) = \Psi_2 * (\mathcal{R} \cup \mathcal{S})$. So, in spite of using different background knowledge, revising $\Psi_1 * \mathcal{R}$ and $\Psi_2 * \mathcal{R}$ by $\mathcal{S}$ both result in the same epistemic state. Therefore, logical coherence guarantees a particular well-behavedness with respect not only to updating, but also to revision.

In the following, two examples to illustrate revision and update in different environments are given. We begin with belief change via the probabilistic principle of minimum cross entropy.

Example 1. A psychologist has been working with addicted people for a couple of years. His experiences concerning the propositions

$\mathcal{V}$: a : addicted to alcohol
d : addicted to drugs
y : being young

may be summarized by the following distribution P that expresses his epistemic state probabilistically:

ω	$P(\omega)$	ω	$P(\omega)$	ω	$P(\omega)$	ω	$P(\omega)$
ady	0.050	$\overline{a}dy$	0.333	$ad\overline{y}$	0.053	$\overline{a}d\overline{y}$	0.053
$a\overline{d}y$	0.093	$\overline{a}\overline{d}y$	0.102	$a\overline{d}\overline{y}$	0.225	$\overline{a}\overline{d}\overline{y}$	0.091

The following probabilistic conditionals may be entailed from P:

$(d|a)[0.242]$ (i.e. $P(d|a) = 0.242$)
$(d|\overline{a})[0.666]$ (i.e. $P(d|\overline{a}) = 0.666$)
$(a|y)[0.246]$ $(a|\overline{y})[0.660]$
$(d|y)[0.662]$ $(d|\overline{y})[0.251]$

Now the psychologist is going to change his job: He will be working in a clinic for people addicted only to alcohol and/or drugs. He is told that the percentage of persons addicted to alcohol, but also addicted to drugs, is higher than usual and may be estimated by 40 %.

So the information the psychologist has about the "new world" is represented by

$$\mathcal{R} = \{a \vee d[1], (d|a)[0.4]\}.$$

The distribution P from above is now supposed to represent background or prior knowledge, respectively. So the psychologist updates P by $\mathcal{R}$ using ME-change and obtains $P^* = P *_{ME} \mathcal{R}$ as new belief state:

ω	$P^*(\omega)$	ω	$P^*(\omega)$	ω	$P^*(\omega)$	ω	$P^*(\omega)$
ady	0.099	$\overline{a}dy$	0.425	$ad\overline{y}$	0.105	$\overline{a}d\overline{y}$	0.066
$a\overline{d}y$	0.089	$\overline{a}\overline{d}y$	0.0	$a\overline{d}\overline{y}$	0.216	$\overline{a}\overline{d}\overline{y}$	0.0

After having spent a couple of days in the new clinic, the psychologist realized that this clinic was for young people only. So he had to revise his knowledge about his new sphere of activity and arrived at the revised belief state $*_{ME}((P, \mathcal{R}) \circ y[1]) = P *_{ME} (\mathcal{R} \cup y[1]) =: P_1^*$ shown in the following table:

ω	$P_1^*(\omega)$	ω	$P_1^*(\omega)$	ω	$P_1^*(\omega)$	ω	$P_1^*(\omega)$
ady	0.120	$\overline{a}dy$	0.700	$ad\overline{y}$	0.0	$\overline{a}d\overline{y}$	0.0
$a\overline{d}y$	0.180	$\overline{a}\overline{d}y$	0.0	$a\overline{d}\overline{y}$	0.0	$\overline{a}\overline{d}\overline{y}$	0.0

This distribution obtained by *revision* is different from that one the psychologist would have obtained by *focusing* his knowledge represented by $P^* = P *_{ME} \mathcal{R}$ on a

ω	$\kappa(\omega)$	ω	$\kappa(\omega)$	ω	$\kappa(\omega)$	ω	$\kappa(\omega)$
$p\overline{k}\,\overline{d}bfw$	2	$\overline{p}k\overline{d}bfw$	0	$\overline{p}\overline{k}dbfw$	0	$\overline{p}\overline{k}\,\overline{d}bfw$	0
$p\overline{k}\,\overline{d}bf\overline{w}$	3	$\overline{p}k\overline{d}bf\overline{w}$	1	$\overline{p}\overline{k}dbf\overline{w}$	1	$\overline{p}\overline{k}\,\overline{d}bf\overline{w}$	1
$p\overline{k}\,\overline{d}b\overline{f}w$	1	$\overline{p}k\overline{d}b\overline{f}w$	1	$\overline{p}\overline{k}db\overline{f}w$	1	$\overline{p}\overline{k}\,\overline{d}b\overline{f}w$	1
$p\overline{k}\,\overline{d}b\overline{f}\overline{w}$	2	$\overline{p}k\overline{d}b\overline{f}\overline{w}$	2	$\overline{p}\overline{k}db\overline{f}\overline{w}$	2	$\overline{p}\overline{k}\,\overline{d}b\overline{f}\overline{w}$	2
$p\overline{k}\,\overline{d}\,\overline{b}fw$	4	$\overline{p}k\overline{d}\,\overline{b}fw$	1	$\overline{p}\overline{k}d\overline{b}fw$	1	$\overline{p}\overline{k}\,\overline{d}\,\overline{b}fw$	0
$p\overline{k}\,\overline{d}\,\overline{b}f\overline{w}$	4	$\overline{p}k\overline{d}\,\overline{b}f\overline{w}$	1	$\overline{p}\overline{k}d\overline{b}f\overline{w}$	1	$\overline{p}\overline{k}\,\overline{d}\,\overline{b}f\overline{w}$	0
$p\overline{k}\,\overline{d}\,\overline{b}\,\overline{f}w$	2	$\overline{p}k\overline{d}\,\overline{b}\,\overline{f}w$	1	$\overline{p}\overline{k}d\overline{b}\,\overline{f}w$	1	$\overline{p}\overline{k}\,\overline{d}\,\overline{b}\,\overline{f}w$	0
$p\overline{k}\,\overline{d}\,\overline{b}\,\overline{f}\overline{w}$	2	$\overline{p}k\overline{d}\,\overline{b}\,\overline{f}\overline{w}$	1	$\overline{p}\overline{k}d\overline{b}\,\overline{f}\overline{w}$	1	$\overline{p}\overline{k}\,\overline{d}\,\overline{b}\,\overline{f}\overline{w}$	0

Figure 1: Epistemic state κ after initialization given in Example 2

young patient, which can be computed via ME update as $P^* *_{ME} \{y[1]\} = P^*(\cdot|y) =: P^*_y$ (please note that the context has changed!):

ω	$P^*_y(\omega)$	ω	$P^*_y(\omega)$	ω	$P^*_y(\omega)$	ω	$P^*_y(\omega)$
ady	0.162	$\overline{a}dy$	0.693	$ad\overline{y}$	0.0	$\overline{a}d\overline{y}$	0.0
$a\overline{d}y$	0.145	$\overline{a}\,\overline{d}y$	0.0	$a\overline{d}\overline{y}$	0.0	$\overline{a}\overline{d}\overline{y}$	0.0

The following example makes use of the simple ordinal change operator defined for ranking functions and qualitative conditionals.

Example 2. Suppose we have the propositional atoms f - *flying*, b - *birds*, p - *penguins*, w - *winged* animals, k - *kiwis*, d - *doves*, . Let the set $\mathcal{R}$ consist of the following conditionals:

r_1 :	$(f\|b)$	*birds fly*
r_2 :	$(b\|p)$	*penguins are birds*
r_3 :	$(\overline{f}\|p)$	*penguins do not fly*
r_4 :	$(w\|b)$	*birds have wings*
r_5 :	$(b\|k)$	*kiwis are birds*
r_6 :	$(b\|d)$	*doves are birds*

Moreover, we assume the strict knowledge *penguins, kiwis, and doves are pairwise exclusive* to hold, which amounts in considering only those worlds as possible that make at most one of $\{p, k, d\}$ true.

The conditionals r_1, r_4, r_5, and r_6 are tolerated by $\mathcal{R}$, whereas r_2 and r_3 are not; but both r_2 and r_3 are tolerated by the set $\{r_2, r_3\}$. This yields the partitioning $\mathcal{R}_0 = \{r_1, r_4, r_5, r_6\}$, $\mathcal{R}_1 = \{r_2, r_3\}$ showing the consistency of $\mathcal{R}$.

As initial epistemic state, we represent inductively $\mathcal{R}$ by applying a simple ordinal c-change to the uniform ranking function κ_u, obtaining $\kappa = \kappa_u * \mathcal{R}$ as current epistemic state (cf. Figure 1).

It can easily be checked that $\kappa_u * \mathcal{R}$ yields the conditional beliefs that penguin-birds do not fly, but both kiwis and doves inherit the property of having wings from their superclass *birds*.

Suppose now that the agent gets to know that this is false for kiwis - kiwis do *not* possess wings - and we want the agent to adopt this new information which has escaped her knowledge before. So, the agent wants to change her beliefs about the world, but the world itself has not changed. Hence (genuine) revision is the proper belief change operation and amounts to computing a new inductive representation for the set $\mathcal{R}' = \{(f|b), (b|p), (\overline{f}|p), (w|b), (b|k), (b|d)\} \cup \{(\overline{w}|k)\}$, i.e. a simple ordinal c-change of κ_u by $\mathcal{R}'$ has to be computed: $\kappa^*_1 = \kappa_u * \mathcal{R}'$. Note that the new information $(\overline{w}|k)$ is not consistent with the prior epistemic state κ but with the context information $\mathcal{R}$ which is refined by $(\overline{w}|k)$.

On the other hand, let us now suppose that the agent learned from the news, that, due to some mysterious illness that has occurred recently among doves, the wings of new-born doves are nearly completely mutilated. She wants to adopt her beliefs to the new information $(\overline{w}|d)$. Obviously, the proper change operation in this case in an update operation as the world under consideration has changed by some event (the occurrence of the mysterious illness).

The updated epistemic state $\kappa^*_2 = \kappa * \{(\overline{w}|d)\}$ is a simple c-change of κ by $\{(\overline{w}|d)\}$ and can be obtained from κ by setting $\kappa^*_2(\omega) = \kappa(\omega) + 2$ for any ω with $\omega \models dw$ and setting $\kappa^*_2(\omega) = \kappa(\omega)$ otherwise.

While the revised state κ^*_1, by construction, still represents the six conditionals that have been known before (and, of course, the new conditional), it can be verified easily that the updated state κ^*_2 only represents the five conditionals $(f|b)$, $(b|p)$, $(\overline{f}|p)$, and $(w|b)$, $(b|k)$, but it no longer satisfies $(b|d)$ because $\kappa^*_2(bd) = \kappa^*_2(\overline{b}d) = 1$ - since *birds* and *wings* have been plausibly related by the conditional $(w|b)$, the property of not having wings casts (reasonably) doubt on doves being birds. Moreover, the agent is now uncertain about the ability of doves to fly, as also $\kappa^*_2(fd) = \kappa^*_2(\overline{f}d) = 1$. This illustrates that priorly stated, explicit knowledge is kept under revision, but might be given up under update. It can be easily checked that these effects would have been the same when updating κ^*_1 instead of κ, since the agent's beliefs on kiwis and doves do not interfere. So, revision and update can be realized within the same scenario, but yielding different results according to the different epistemic roles that have been assigned to the corresponding pieces of information in the change processes.

Conclusion and outlook

In this paper, we presented a formal, unifying framework for iterated belief change and nonmonotonic inference operations with explicit epistemic background knowledge. We proposed a clear conceptual distinction between belief revision and belief update and showed how this distinction can be realized technically. Both revision and update operations could be iterated, with different outcomes. General postulates served as cornerstones to classify change operations, in the tradition of the AGM-theory. Most of these postulates were related to similar statements in revision and update theories. The postulate (CSR5) of Logical Coherence, however, is new and refers explicitly to iterated belief change. It proved to be very strong, and is shown to correspond to strong cumulativity of universal inference operations.

We illustrated our ideas by the probabilistic change operation that is induced by the principle of minimum cross entropy, and by c-changes of ordinal conditional functions. Up to date, the minimum cross entropy operator is the only change operation which is known to satisfy all requirements listed in this paper. It is part of our ongoing work to explore which other frameworks are rich enough to comply with the postulates and properties presented in this paper. Furthermore, we will investigate, how our ideas can be combined with work on explicitly taking actions and its effects into account (see, e.g., (Hunter and Delgrande 2005; Lang 2007)). Another focus of research is to generalize the presented approach to handle inconsistent contextual information and to compare our work to (Delgrande, Dubois, and Lang 2006).

References

Alchourrón, C.; Gärdenfors, P.; and Makinson, P. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* 50(2):510–530.

Bourne, R., and Parsons, S. 1999. Maximum entropy and variable strength defaults. In *Proceedings Sixteenth International Joint Conference on Artificial Intelligence, IJCAI'99*, 50–55.

Csiszár, I. 1975. I-divergence geometry of probability distributions and minimization problems. *Ann. Prob.* 3:146–158.

Darwiche, A., and Pearl, J. 1997. On the logic of iterated belief revision. *Artificial Intelligence* 89:1–29.

Delgrande, J.; Dubois, D.; and Lang, J. 2006. Iterated revision and prioritized merging. In *Proc. KR'06*, 210–220. AAAI Press.

Dubois, D., and Prade, H. 1996. Non-standard theories of uncertainty in plausible reasoning. In Brewka, G., ed., *Principles of Knowledge Representation*. CSLI Publications.

Dubois, D. 2006. Three scenarios for the revision of epistemic states. In *Proceedings NMR'2006*.

Gärdenfors, P. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge, Mass.: MIT Press.

Goldszmidt, M., and Pearl, J. 1996. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence* 84:57–112.

Goldszmidt, M.; Morris, P.; and Pearl, J. 1993. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(3):220–232.

Hajek, P.; Havranek, T.; and Jirousek, R. 1992. *Uncertain Information Processing in Expert Systems*. Boca Raton, Florida: CRC-Press.

Hunter, A., and Delgrande, J. 2005. Iterated belief change: a transition system apporach. In *Proceedings IJCAI'2005*.

Jaynes, E. 1983a. *Papers on Probability, Statistics and Statistical Physics*. Dordrecht, Holland: D. Reidel Publishing Company.

Jaynes, E. 1983b. Where do we stand on maximum entropy? In *Papers on Probability, Statistics and Statistical Physics*. Dordrecht, Holland: D. Reidel Publishing Company. 210–314.

Katsuno, H., and Mendelzon, A. 1991. On the difference between updating a knowledge base and revising it. In *Proceedings Second International Conference on Principles of Knowledge Representation and Reasoning, KR'91*, 387–394. San Mateo, Ca.: Morgan Kaufmann.

Kern-Isberner, G. 1998. Characterizing the principle of minimum cross-entropy within a conditional-logical framework. *Artificial Intelligence* 98:169–208.

Kern-Isberner, G. 1999. Postulates for conditional belief revision. In *Proceedings Sixteenth International Joint Conference on Artificial Intelligence, IJCAI-99*, 186–191. Morgan Kaufmann.

Kern-Isberner, G. 2001. *Conditionals in nonmonotonic reasoning and belief revision*. Springer, Lecture Notes in Artificial Intelligence LNAI 2087.

Kern-Isberner, G. 2004. A thorough axiomatization of a principle of conditional preservation in belief revision. *Annals of Mathematics and Artificial Intelligence* 40(1-2):127–164.

Kullback, S. 1968. *Information Theory and Statistics*. New York: Dover.

Lang, J.; Marquis, P.; and Williams, M.-A. 2001. Updating epistemic states. In *AI 2007*. Springer, LNAI 2256. 297–308.

Lang, J. 2007. Belief update revisited. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence, IJCAI-07*, 2517–2522.

Makinson, D., and Gärdenfors, P. 1991. Relations between the logic of theory change and nonmonotonic logic. In *Proceedings Workshop The Logic of Theory Change, Konstanz, Germany, 1989*, 185–205. Berlin Heidelberg New York: Springer.

Makinson, D. 1994. General patterns in nonmonotonic reasoning. In Gabbay, D.; Hogger, C.; and Robinson, J., eds., *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. Oxford University Press. 35–110.

Paris, J., and Vencovska, A. 1997. In defence of the maximum entropy inference process. *International Journal of Approximate Reasoning* 17(1):77–103.

Shore, J., and Johnson, R. 1981. Properties of cross-entropy minimization. *IEEE Transactions on Information Theory* IT-27:472–482.

Shore, J. 1986. Relative entropy, probabilistic inference and AI. In Kanal, L., and Lemmer, J., eds., *Uncertainty in Artificial Intelligence*. Amsterdam: North-Holland. 211–215.

Spohn, W. 1988. Ordinal conditional functions: a dynamic theory of epistemic states. In Harper, W., and Skyrms, B., eds., *Causation in Decision, Belief Change, and Statistics, II*. Kluwer Academic Publishers. 105–134.