Representing and learning conditional information in possibility theory

Gabriele Kern-Isberner

FernUniversität Hagen, Dept. of Computer Science,
P.O. Box 940, D-58084 Hagen, Germany
gabriele.kern-isberner@fernuni-hagen.de

1 Introduction

Conditionals (*if-then-rules*, *default rules*) are most important objects in knowledge representation and commonsense reasoning. They are closely related to epistemic notions such as plausibility, possibility, probability and the like, in that a conditional "*If A then B*" is generally considered to be accepted if its confirmation, $A \wedge B$, is judged to be more plausible, possible or probable, than its refutation, $A \wedge \neg B$. Conditionals can even be quantified by degrees of belief or by (statistical or subjective) probabilities, expressing to what extent they are taken to be established.

Due to their non-classical nature, however, conditionals are not easily dealt with. On the one hand, to validate conditionals, richer epistemic structures than plain propositional interpretations are needed, as e.g. *probability distributions* (see, e.g., [Par94]) and *possibility distributions* as fuzzy representations of epistemic attitudes (cf. [DP91c,DLP94,BDP99]). On the other hand, inference from conditional knowledge bases is much more difficult than in classically logical environments, since it is not a strictly deductive one. Rather it is based on nonmonotonic reasoning [DP91a,LM92,DP91b,BDP97] and involves belief revision theory [Rot91,FH94,KS91]. A rigid structure seems to be lacking which guides reasoning with conditionals and, the other way around, helps us to identify relevant conditional knowledge (*rules*) from statistical databases, making *knowledge discovery* possible.

The aim of this paper is to present an algebraic theory of conditional structures on possible worlds and to show how this theory can be applied in possibility theory to obtain adequate inductive representations of conditional knowledge bases by possibility distributions. These possibility distributions can then be used for inferences. The basic idea is to develop a new methodology of making interactions of conditionals transparent, being aware that any attempt to force conditionals into a classical framework will cut down their unique dynamic expressiveness. Rather, conditionals are considered here as agents having effects on worlds so as to establish conditional beliefs. With the notion of *conditional structures*, we can then define what it means that a possibility distribution is *indifferent with respect to a set of conditionals*, thereby respecting the structures imposed by the conditionals under consideration. Note that we will be dealing with (finite) *sets of conditionals*, not with single conditionals, at a time, taking the joint effects of several conditionals into account. Moreover, the same techniques can be applied to handle revisions of possibility distributions by incorporating new conditional beliefs. Last not least, the methodology provides an approach to discovering conditional knowledge in statistical databases (rule extraction) by considering the induced possibility distribution (see, e.g., [GK97]).

As we already said, we will largely avoid classically logical techniques. That does not mean, however, that this new approach to conditionals lacks a solid foundation. Quite to the contrary – the underlying conditional theory is a group theoretical one, e.g., conditional structures are defined via homomorphisms of groups. This embedding into a group theoretical framework, as awkward and strange as it appears at first sight, is the key to handle complex conditional interactions adequately, as it is necessary for inductive representations of conditionals, by representing these interactions as products while allowing cancellations. In search of appropriate structures, homomorphisms are excellent tools, since they are designated to transfer structures. In particular, the kernels of these homomorphisms (as usual) play a prominent role in identifying structures. So, knowledge discovery comes down to elaborating kernels and factorizing groups, inducing operations on conditionals so as to obtain concise ("short") rules compatible with the numerical structure.

This theory is not only devised for possibility theory but actually provides a powerful approach to conditionals which may also be applied in a probabilistic framework or when using *ordinal conditional functions* [Spo88], which are quite similar to possibility distributions, and it even yields essential insights in purely qualitative environments. It is presented in full in [KI99,KI01a], some results for ordinal conditional functions were published in [KI00a], and probabilistic rule extraction is sketched in [KI00b]. Now this paper will present new results for a possibilistic framework.

2 Conditionals in possibility theory

In the sequel, conditionals "*If A then B*" are formally represented as $(B|A)$, expressing a relationship between two (propositional) formulas A, the *antecedent* or *premise*, and B, the *consequent*, of a (finite) propositional language $\mathcal{L}$. *Single-elementary conditionals* are conditionals whose antecedents are conjunctions of literals, and whose consequents consist of one single literal.

Let $(\mathcal{L} \mid \mathcal{L})$ denote the set of all conditionals $(B|A)$ with $A, B \in \mathcal{L}$. To simplify notation, we will write $\overline{A}$ instead of $\neg A$, and AB instead of $A \wedge B$, for formulas $A, B \in \mathcal{L}$. Let Ω be the complete set of interpretations of $\mathcal{L}$, where each $\omega \in \Omega$ is taken to be a *possible world* for $\mathcal{L}$. We will only consider finite sets $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$ of conditionals.

A *possibility distribution* is a mapping $\pi : \Omega \to [0,1]$, where each $\pi(\omega)$ reflects the degree of possibility an intelligent agent assigns to ω. π induces a *possibility measure*

$$\Pi : \mathcal{L} \to [0,1], \Pi(A) = \max\{\pi(\omega) \mid \omega \models A\},$$

and a dual *necessity measure*

$$N = N_\pi : \mathcal{L} \to [0,1], N(A) = 1 - \Pi(\overline{A})$$

(cf., e.g., [DLP94]). To avoid confusion, we will not distinguish between a possibility distribution, π, on Ω and its induced possibility measure, Π, on $\mathcal{L}$, and will use the same symbol π for both. Furthermore, a possibility degree can also be assigned to a conditional $(B|A)$ by setting

$$\pi(B|A) = \frac{\pi(AB)}{\pi(A)},$$

in full analogy to Bayesian conditioning in probability theory. The corresponding possibility distribution is

$$\pi(\omega|A) = \begin{cases} \frac{\pi(\omega)}{\Pi(A)} & \text{if} \omega \models A \\ 0, & \text{otherwise} \end{cases}$$

Note that we also make use of the product operation in $[0,1]$. That means, that our approach is not only based upon comparing numbers, but also takes relations between numbers into account. These numerical relationships encode important information about the (relative) strength of conditionals which proves to be particularly crucial for representation and revision tasks. This amounts to carrying over Spohn's argumentation in [Spo88] to a possibilistic framework (see also [KI99] and [BDP97]).

A conditional $(B|A)$ *is accepted in* π, $\pi \models (B|A)$, iff $\pi(AB) > \pi(A\overline{B})$ (which is equivalent to $\pi(\overline{B}|A) < 1$ and $N_\pi(B|A) = 1 - \pi(\overline{B}|A) > 0$) (cf. [DP94]). So, in accordance with intuition, a conditional $(B|A)$ is accepted in the epistemic state modeled by a possibility distribution, if its confirmation (AB) is considered to be more possible (or plausible) than its refutation $(A\overline{B})$.

********** **quantified conditionals?** ***************

We assume all possibility distributions to be normalized, i.e. $\pi(\top) = 1$, and, only for conciseness of presentation, to be positive, i.e. $\pi(\omega) > 0$ for all $\omega \in \Omega$.

3 Conditional structures

A world $\omega \in \Omega$ is said to *verify* a conditional $(B|A)$, if $\omega \models AB$; it *falsifies* $(B|A)$, if $\omega \models A\overline{B}$; if $\omega \models \overline{A}$, then $(B|A)$ is not applicable to ω. By observing this behavior of worlds with respect to it, each conditional $(B|A)$ can be considered as a generalized (namely three-valued) indicator function on worlds:

$$(B|A)(\omega) = \begin{cases} 1 & : \quad \omega \models AB \\ 0 & : \quad \omega \models A\overline{B} \\ u & : \quad \omega \models \overline{A} \end{cases} \tag{1}$$

where u stands for *undefined* [Cal91,DeF74]. When we consider (finite) sets of conditionals $\mathcal{R} = \{(B_1|A_1), \ldots, (B_n|A_n)\} \subseteq (\mathcal{L} \mid \mathcal{L})$, we have to modify this representation appropriately to identify the effect of each conditional in $\mathcal{R}$ on worlds in Ω. To this end, we replace the numbers 0 and 1 by abstract symbols, $\mathbf{a}_i^+, \mathbf{a}_i^-$, that we associate to each conditional $(B_i|A_i)$ in $\mathcal{R}$. Moreover, we will make use of a group structure to represent the joint impact of conditionals on worlds. So let $\mathcal{F}_\mathcal{R} = \langle \mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^- \rangle$ be the free abelian group with generators $\mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^-$, i.e. $\mathcal{F}_\mathcal{R}$ consists of all elements of the form $(\mathbf{a}_1^+)^{r_1}(\mathbf{a}_1^-)^{s_1} \ldots (\mathbf{a}_n^+)^{r_n}(\mathbf{a}_n^-)^{s_n}$ with integers $r_i, s_i \in \mathbb{Z}$ (the ring of integers). Each element of $\mathcal{F}_\mathcal{R}$ can be identified by its exponents, so that $\mathcal{F}_\mathcal{R}$ is isomorphic to $\mathbb{Z}^{2n}$ (cf. [LS77]). For each $i, 1 \leqslant i \leqslant n$, we further define a function $\sigma_i = \sigma_{(B_i|A_i)} : \Omega \to \mathcal{F}_\mathcal{R}$ by setting

$$\sigma_i(\omega) = \sigma_{(B_i|A_i)}(\omega) := \begin{cases} \mathbf{a}_i^+ & \text{if} \quad (B_i|A_i)(\omega) = 1 \\ \mathbf{a}_i^- & \text{if} \quad (B_i|A_i)(\omega) = 0 \\ 1 & \text{if} \quad (B_i|A_i)(\omega) = u \end{cases}$$

$\sigma_i(\omega)$ represents the manner in which the conditional $(B_i|A_i)$ applies to the possible world ω. The neutral element 1 of $\mathcal{F}_\mathcal{R}$ corresponds to the non-applicability of $(B_i|A_i)$ in case that the antecedent A_i is not satisfied. The function $\sigma_\mathcal{R} : \Omega \to \mathcal{F}_\mathcal{R}$,

$$\sigma_\mathcal{R}(\omega) = \prod_{1 \leqslant i \leqslant n} \sigma_i(\omega) = \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \mathbf{a}_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \mathbf{a}_i^- \tag{2}$$

describes the all-over effect of $\mathcal{R}$ on ω. $\sigma_\mathcal{R}(\omega)$ is called the *conditional structure of* ω *with respect to* $\mathcal{R}$. Due to $\mathcal{F}_\mathcal{R}$ being commutative, we will be able to take the simultaneous impact of all conditionals in $\mathcal{R}$ into account, without assuming any order of applying the conditionals.

$\sigma_{\mathcal{R}}$ is defined on the set Ω, assigning group elements of $\mathcal{F}_{\mathcal{R}}$ to the possible worlds. Now, to comply with the group theoretical structure in $\mathcal{F}_{\mathcal{R}}$, we also take the worlds $\omega \in \Omega$ as formal generators of a free abelian group

$$\widehat{\Omega} := \langle \omega \mid \omega \in \Omega \rangle$$

$\widehat{\Omega}$ consists of all products $\widehat{\omega} = {\omega_1}^{r_1} \dots {\omega_m}^{r_m}$, with $\omega_1, \dots, \omega_m \in \Omega$, and integers $r_1, \dots r_m$. Now $\sigma_{\mathcal{R}}$ may be extended to $\widehat{\Omega}$ in a straightforward manner by setting

$$\sigma_{\mathcal{R}}({\omega_1}^{r_1} \dots {\omega_m}^{r_m}) = \sigma_{\mathcal{R}}(\omega_1)^{r_1} \dots \sigma_{\mathcal{R}}(\omega_m)^{r_m},$$

yielding a *homomorphism of groups* $\sigma_{\mathcal{R}} : \widehat{\Omega} \to \mathcal{F}_{\mathcal{R}}$.

Having the same conditional structure defines an equivalence relation $\equiv_{\mathcal{R}}$ on $\widehat{\Omega}$:

$$\widehat{\omega}_1 \equiv_{\mathcal{R}} \widehat{\omega}_2 \text{ iff } \sigma_{\mathcal{R}}(\widehat{\omega}_1) = \sigma_{\mathcal{R}}(\widehat{\omega}_2) \tag{3}$$

Those elements of $\widehat{\Omega}$ that are balanced with respect to the effects of conditionals in $\mathcal{R}$ are contained in the *kernel of* $\sigma_{\mathcal{R}}$,

$$ker\, \sigma_{\mathcal{R}} = \{\widehat{\omega} \in \widehat{\Omega} \mid \sigma_{\mathcal{R}}(\widehat{\omega}) = 1\}$$

To take the normalizing constraint $\Pi(\top) = 1$ into acccount, we also focus on equivalence with respect to $\sigma_{(\top|\top)}$. Since $\sigma_{(\top|\top)}$ simply counts the generators occurring in $\widehat{\omega}$, two elements $\widehat{\omega}_1 = \omega_1^{r_1} \dots \omega_m^{r_m}$, $\widehat{\omega}_2 = \nu_1^{s_1} \dots \nu_p^{s_p} \in \widehat{\Omega}$ are $\sigma_{(\top|\top)}$-equivalent, $\widehat{\omega}_1 \equiv_{\top} \widehat{\omega}_2$, iff $\sum_{1 \leqslant j \leqslant m} r_j = \sum_{1 \leqslant k \leqslant p} s_k$. This means, $\widehat{\omega}_1 \equiv_{\top} \widehat{\omega}_2$ iff they both are a (cancelled) product of the same number of generators, each generator being counted with its corresponding exponent. In particular, $\omega_1 \equiv_{\top} \omega_2$ for any possible worlds $\omega_1, \omega_2 \in \Omega$. Therefore, $ker\, \sigma_{(\top|\top)} = \{\widehat{\omega} = {\omega_1}^{r_1} \cdot \ldots \cdot {\omega_m}^{r_m} \in \widehat{\Omega} \mid \sum_{j=1}^{m} r_j = 0\}$. Set $ker_0\, \sigma_{\mathcal{R}} := ker\, \sigma_{\mathcal{R}} \cap ker\, \sigma_{(\top|\top)}$, that is,

$$ker_0\, \sigma_{\mathcal{R}} = \{\widehat{\omega} = {\omega_1}^{r_1} \cdot \ldots \cdot {\omega_m}^{r_m} \in \widehat{\Omega} \mid \sum_{j=1}^{m} r_j = 0 \text{ and } \sigma_{\mathcal{R}}(\widehat{\omega}) = 1\}$$

We will illustrate the benefits of these technical definitions in the next example.

Example 1. Consider the set $\mathcal{R} = \{(d|a), (d|b), (d|c)\}$ of conditionals using the atoms a, b, c, d. Let $\mathbf{a}_1^{\pm}, \mathbf{a}_2^{\pm}, \mathbf{a}_3^{\pm}$ be the group generators associated with $(d|a), (d|b), (d|c)$, respectively, so $\mathcal{F}_{\mathcal{R}} = \langle \mathbf{a}_1^+, \mathbf{a}_1^-, \mathbf{a}_2^+, \mathbf{a}_2^-, \mathbf{a}_3^+, \mathbf{a}_3^- \rangle$. The possible world $\omega = abcd$ verifies all three conditionals $(d|a), (d|b), (d|c)$, so its conditional structure is $\sigma_{\mathcal{R}}(abcd) = \mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^+$. For $\omega = ab\overline{c}d$, we calculate $\sigma_{\mathcal{R}}(ab\overline{c}d) = \mathbf{a}_1^+\mathbf{a}_2^+$, since $ab\overline{c}d$ verifies the first two conditionals, and the third conditional is not applicable to it.

Moreover, by conditional structures, it is possible to compare worlds, or sets of worlds, as to their behavior with respect to the conditionals in $\mathcal{R}$. For instance, we have

$$\begin{aligned}\sigma_{\mathcal{R}}(ab\overline{c}d)\sigma_{\mathcal{R}}(a\overline{b}cd)\sigma_{\mathcal{R}}(\overline{a}bcd) &= (\mathbf{a}_1^+\mathbf{a}_2^+)(\mathbf{a}_1^+\mathbf{a}_3^+)(\mathbf{a}_2^+\mathbf{a}_3^+)\\ &= (\mathbf{a}_1^+)^2(\mathbf{a}_2^+)^2(\mathbf{a}_3^+)^2 = (\mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^+)^2 = \sigma_{\mathcal{R}}(abcd)^2.\end{aligned}$$

That is to say, that the elements of the set $\{ab\overline{c}d, a\overline{b}cd, \overline{a}bcd\}$ show collectively the same conditional behavior as two copies of $abcd$. In the group theoretical framework using $\widehat{\Omega}$, this can be expressed as

$$\sigma_{\mathcal{R}}\left(ab\overline{c}d \cdot a\overline{b}cd \cdot \overline{a}bcd\right) = \sigma_{\mathcal{R}}\left(abcd^2\right)$$

or

$$\sigma_{\mathcal{R}}\left(\frac{ab\overline{c}d \cdot a\overline{b}cd \cdot \overline{a}bcd}{abcd^2}\right) = 1.$$

So $\frac{ab\overline{c}d \cdot a\overline{b}cd \cdot \overline{a}bcd}{abcd^2} \in ker\, \sigma_{\mathcal{R}}$, but $\frac{ab\overline{c}d \cdot a\overline{b}cd \cdot \overline{a}bcd}{abcd^2} \notin ker_0\, \sigma_{\mathcal{R}}$, since $ab\overline{c}d \cdot a\overline{b}cd \cdot \overline{a}bcd \not\equiv_{\top} abcd^2$ (three generators versus two generators).

4 Conditional indifference

Conditional structures, on the one hand, reflect clearly the (isolated) impact of each conditional in a set $\mathcal{R}$, and, on the other hand, make the interactions between different conditionals in $\mathcal{R}$ transparent and computable. In this section, we will link the algebraic notion of conditional structures to the numerical properties of possibility distributions $\Pi : \Omega \rightarrow [0,1]$. To this end, we now focus on the behavior of Π with respect to the elements in $\widehat{\Omega}$.

To fit into the group theoretical framework, we extend each such distribution to a homomorphism, $\pi : \widehat{\Omega} \rightarrow (\mathbb{R}^+, \cdot)$[1], by setting

$$\pi({\omega_1}^{r_1} \cdot \ldots \cdot {\omega_m}^{r_m}) = \pi(\omega_1)^{r_1} \cdot \ldots \cdot \pi(\omega_m)^{r_m}$$

This allows us to analyze numerical relationships holding between different $\pi(\omega)$, and to elaborate the conditionals whose structures π follows:

Definition 1. *Let $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$ be a set of conditonals, let π be a (positive) possibility distribution. π is said to be* indifferent with respect to $\mathcal{R}$ *iff*

$$\pi(\widehat{\omega}_1) = \pi(\widehat{\omega}_2) \tag{4}$$

whenever $\sigma_{\mathcal{R}}(\widehat{\omega}_1) = \sigma_{\mathcal{R}}(\widehat{\omega}_2)$ for $\widehat{\omega}_1 \equiv_{\top} \widehat{\omega}_2 \in \widehat{\Omega}$.

π being indifferent with respect to $\mathcal{R}$ means that it does not distinguish between different elements $\widehat{\omega}_1, \widehat{\omega}_2$ with the same conditional structure with respect to $\mathcal{R}$. Conversely, if π is indifferent with respect to $\mathcal{R}$, then any inequality $\sigma_{\mathcal{R}}(\widehat{\omega}) \neq 1$ can be explained by an unbalanced conditional structure $\sigma_{\mathcal{R}}(\widehat{\omega}) \neq 1$. Normalizing constraints are taken into account by observing $\equiv_{\top}$-equivalence. Note that we presupposed π to be positive only for technical reasons, in order to make this definition a very concise one. For the general case, cf. [KI01a].

The following proposition provides an obvious characterization of conditional indifference in terms of the kernels of π and $\sigma_{\mathcal{R}}$:

Proposition 1. *Let $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$ be a set of conditonals, let π be a (positive) possibility distribution. π is indifferent with respect to $\mathcal{R}$ iff $ker_0\ \sigma_{\mathcal{R}} \subseteq ker_0\ \pi$.*

The notion of conditional indifference captures interactions of conditionals of arbitrary depth by making use of the group homomorphism induced by π. It also handles interactions on the superficial level of the worlds $\omega \in \Omega$:

Proposition 2. *Suppose $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$ to be a set of conditonals, let π be a (positive) possibility distribution which is indifferent with respect to $\mathcal{R}$. Let $\omega_1, \omega_2 \in \Omega$ be worlds with identical conditional structures with respect to $\mathcal{R}$, i.e. $\sigma_{\mathcal{R}}(\omega_1) = \sigma_{\mathcal{R}}(\omega_2)$. Then $\pi(\omega_1) = \pi(\omega_2)$.*

This proposition is clear by observing that any two possible worlds are $\equiv_{\top}$-equivalent.

The next theorem gives a simple criterion to check for conditional indifference:

Theorem 1. *Let $\mathcal{R} = \{(B_1|A_1), \ldots, (B_n|A_n)\}$ be a set of conditionals.*

A (positive) possibility distribution π is indifferent with respect to $\mathcal{R}$ iff there are positive real numbers $\alpha_0, \alpha_1^+, \alpha_1^-, \ldots, \alpha_n^+, \alpha_n^- \in \mathbb{R}^+$ such that for all $\omega \in \Omega$,

$$\pi(\omega) = \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^- \tag{5}$$

[1] This denotes the group of non-negative real numbers with respect to multiplication

Sketch of proof. According to Lemma ??, the equivalence relation (3) provides a rough classification of the worlds in Ω with respect to the conditionals in $\mathcal{R}$. Obtaining a representation of the form (5) then amounts to checking the solvability of a linear equational system. The proof of this theorem is very similar to the proof of the analogous theorem for probabilistic representation of knowledge given in [KI98]. □

5 Representing conditionals by possibility distributions

************ noch ausarbeiten ************

BenferhatDuboisPrade92 BenferhatSaffiottiSmets00 Representing conditionals in possibility theory

****

Definition 2. *Possibility distributions π which represent a set $\mathcal{R}$ of conditionals, $\pi \models \mathcal{R}$, and are indifferent with respect to it, are called* c-representations *of $\mathcal{R}$.*

Theorem 1 provides an intelligible schema to construct conditionally indifferent distributions. To obtain a c-representation π of $\mathcal{R}$, we simply set up π according to (5) and choose the $\alpha_1^+, \alpha_1^-, \ldots, \alpha_n^+, \alpha_n^-$ appropriately to ensure that π is a possibility distribution that represents $\mathcal{R}$:

Theorem 2. *A (positive) possibility distribution π is a c-representation of a set $\mathcal{R}$ of conditionals, iff there are positive real numbers $\alpha_0, \alpha_1^+, \alpha_1^-, \ldots, \alpha_n^+, \alpha_n^- \in \mathbb{R}^+$ such that for all $\omega \in \Omega$,*

$$\pi(\omega) = \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^-$$

and for all $j, 1 \leqslant j \leqslant n$,

$$\alpha_j^+ \left(\max_{\omega \models A_j B_j} \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^- \right) > \alpha_j^- \left(\max_{\omega \models A_j \overline{B}_j} \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^- \right) \quad (6)$$

Proof. Let π be a (positive) possibility distributions, and let $\mathcal{R} \subset (\mathcal{L} \mid \mathcal{L})$ be a set of conditionals. π is a c-representation of $\mathcal{R}$ iff π is indifferent with respect to $\mathcal{R}$ and $\pi \models \mathcal{R}$, i.e. iff there are positive real numbers $\alpha_0, \alpha_1^+, \alpha_1^-, \ldots, \alpha_n^+, \alpha_n^- \in \mathbb{R}^+$ such that

$$\pi(\omega) = \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^-$$

and for all $j, 1 \leqslant j \leqslant n$, $\pi(A_j B_j) > \pi(A_j \overline{B}_j)$, which is equivalent to

$$\begin{aligned}
& \max_{\omega \models A_j B_j} \pi(\omega) > \max_{\omega \models A_j \overline{B}_j} \pi(\omega) \\
\Longleftrightarrow\ & \max_{\omega \models A_j B_j} \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^- > \max_{\omega \models A_j \overline{B}_j} \alpha_0 \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^- \\
\Longleftrightarrow\ & \alpha_0 \max_{\omega \models A_j B_j} \alpha_j^+ \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^- > \alpha_0 \max_{\omega \models A_j \overline{B}_j} \alpha_j^- \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^- \\
\Longleftrightarrow\ & \alpha_j^+ \cdot \max_{\omega \models A_j B_j} \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^- > \alpha_j^- \cdot \max_{\omega \models A_j \overline{B}_j} \prod_{\substack{i \neq j \\ \omega \models A_i B_i}} \alpha_i^+ \prod_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \alpha_i^-
\end{aligned}$$

Note that the normalizing constant α_0 only has to be chosen appropriately to ensure that actually a normalized possibility distribution is obtained.

Therefore, c-representations of a set $\mathcal{R} = \{(B_1|A_1), \ldots, (B_n|A_n)\}$ of conditionals can be obtained by replacing the formal symbols $\mathbf{a}_i^+, \mathbf{a}_i^-$, associated to each conditional in R and used to reflect the conditional structures of worlds (cf. (2)), by appropriate numerical values. We will call a mapping $I : \{\mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^-\} \to \mathbb{R}^+$, $\mathbf{a}_i^+ \mapsto \alpha_i^+, \mathbf{a}_i^- \mapsto \alpha_i^-, 1 \leqslant i \leqslant n$, such that (6) holds, a *c-interpretation of* $\mathcal{R}$. From c-interpretations, c-representations arise in a straightforward way via the formal tool of conditional structures.

The structure imposed on possible worlds by a c-representation is as finely grained as can be justified by differences in conditional structures. c-representations observe conditional structures and thereby handle interactions of conditionals in a most adequate way. Therefore, they are most appropriate to represent conditional knowledge. We will illustrate this by examples in the following. But first, as a general observation, we will show that all c-representations satisfy the property of *irrelevance*

Lemma 1. *Suppose $\mathcal{R}$ is a set of conditionals, and let d be an atomic proposition not appearing in any of the conditionals in $\mathcal{R}$. Then, for any c-representation π_c of $\mathcal{R}$, and for any conditional $(B|A) \in (\mathcal{L} \mid \mathcal{L})$, $\pi_c \models (B|A)$ iff $\pi_c \models (B|Ad)$.*

This is easily proved by observing that for any propositional interpretation ω' over $\mathcal{V} - \{d\}$, we have $\sigma_{\mathcal{R}}(\omega' d) = \sigma_{\mathcal{R}}(\omega' \overline{d})$, and therefore, $\pi_c(A\dot{B}) = \pi_c(A\dot{B}d)$ for $\dot{B} \in \{B, \overline{B}\}$.

Let $\mid\!\sim^c$ be the (nonmonotonic) inference relation based on considering all c-representations:

$$\mathcal{R} \mid\!\sim^c (B|A) \quad \text{iff} \quad \pi_c \models (B|A) \text{ for all c-representations } \pi_c \text{ of } \mathcal{R}, \tag{7}$$

where $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L}), (B|A) \in (\mathcal{L} \mid \mathcal{L})$. Then Lemma 1 yields

Corollary 1. *Let $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$, $(B|A) \in (\mathcal{L} \mid \mathcal{L})$, let d be an atomic proposition not appearing in any of the conditionals in $\mathcal{R}$. Then $\mathcal{R} \mid\!\sim^c (B|A)$ iff $\mathcal{R} \mid\!\sim^c (B|Ad)$.*

******* **system P?** *******

Furthermore, the next example will show that $\mid\!\sim^c$ does not suffer from the so-called *drowning problem* which is a problem for general possibilistic inference (cf. [BDP97]).

Example 2. Let $\mathcal{R}$ consist of the following conditionals:

$r_1 : (w|p)$	*Penguins have wings.*
$r_2 : (f|w)$	*Winged animals fly.*
$r_3 : (\overline{f}|p)$	*Penguins do not fly.*
$r_4 : (b|p)$	*Penguins are birds.*
$r_5 : (e|b)$	*Birds lay eggs.*

In Table 1, we list the conditional structures of all possible worlds with respect to $\mathcal{R}$. Possibility values can be calculated directly from that table. For instance, $\pi(pw\overline{f}be) = \alpha_0 \alpha_1^+ \alpha_2^- \alpha_3^+ \alpha_4^+ \alpha_5^+$, where the α's are suitably chosen so as to fulfill (6) for that special set of conditionals. For α_4^+, α_4^- and α_5^+, α_5^-, e.g., we obtain (α_0 has already been cancelled)

$$\begin{aligned}
&\alpha_4^+ \cdot \max\{\alpha_1^+ \alpha_2^+ \alpha_3^- \alpha_5^+, \alpha_1^+ \alpha_2^+ \alpha_3^- \alpha_5^-, \alpha_1^+ \alpha_2^- \alpha_3^+ \alpha_5^+, \alpha_1^+ \alpha_2^- \alpha_3^+ \alpha_5^-,\\
&\qquad \alpha_1^- \alpha_3^- \alpha_5^+, \alpha_1^- \alpha_3^- \alpha_5^-, \alpha_1^- \alpha_3^+ \alpha_5^+, \alpha_1^- \alpha_3^+ \alpha_5^-\}\\
&\quad > \alpha_4^- \cdot \max\{\alpha_1^+ \alpha_2^+ \alpha_3^-, \alpha_1^+ \alpha_2^- \alpha_3^+, \alpha_1^- \alpha_3^-, \alpha_1^- \alpha_3^+\}
\end{aligned}$$

$$\begin{aligned}
&\alpha_5^+ \cdot \max\{\alpha_1^+ \alpha_2^+ \alpha_3^- \alpha_4^+, \alpha_1^+ \alpha_2^- \alpha_3^+ \alpha_4^+, \alpha_1^- \alpha_3^- \alpha_4^+, \alpha_1^- \alpha_3^+ \alpha_4^+, \alpha_2^+, \alpha_2^-, 1\}\\
&\quad > \alpha_5^- \cdot \max\{\alpha_1^+ \alpha_2^+ \alpha_3^- \alpha_4^+, \alpha_1^+ \alpha_2^- \alpha_3^+ \alpha_4^+, \alpha_1^- \alpha_3^- \alpha_4^+, \alpha_1^- \alpha_3^+ \alpha_4^+, \alpha_2^+, \alpha_2^-, 1\}
\end{aligned}$$

ω	$\sigma_{\mathcal{R}}(\omega)$	ω	$\sigma_{\mathcal{R}}(\omega)$	ω	$\sigma_{\mathcal{R}}(\omega)$	ω	$\sigma_{\mathcal{R}}(\omega)$
$pwfbe$	$\mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^-\mathbf{a}_4^+\mathbf{a}_5^+$	$p\overline{w}fbe$	$\mathbf{a}_1^-\mathbf{a}_3^-\mathbf{a}_4^+\mathbf{a}_5^+$	$\overline{p}wfbe$	$\mathbf{a}_2^+\mathbf{a}_5^+$	$\overline{p}\,\overline{w}fbe$	$\mathbf{a}_5^+$
$pwfb\overline{e}$	$\mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^-\mathbf{a}_4^+\mathbf{a}_5^-$	$p\overline{w}fb\overline{e}$	$\mathbf{a}_1^-\mathbf{a}_3^-\mathbf{a}_4^+\mathbf{a}_5^-$	$\overline{p}wfb\overline{e}$	$\mathbf{a}_2^+\mathbf{a}_5^-$	$\overline{p}\,\overline{w}fb\overline{e}$	$\mathbf{a}_5^-$
$pwf\overline{b}e$	$\mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^-\mathbf{a}_4^-$	$p\overline{w}f\overline{b}e$	$\mathbf{a}_1^-\mathbf{a}_3^-\mathbf{a}_4^-$	$\overline{p}wf\overline{b}e$	$\mathbf{a}_2^+$	$\overline{p}\,\overline{w}f\overline{b}e$	1
$pwf\overline{b}\overline{e}$	$\mathbf{a}_1^+\mathbf{a}_2^+\mathbf{a}_3^-\mathbf{a}_4^-$	$p\overline{w}f\overline{b}\overline{e}$	$\mathbf{a}_1^-\mathbf{a}_3^-\mathbf{a}_4^-$	$\overline{p}wf\overline{b}\overline{e}$	$\mathbf{a}_2^+$	$\overline{p}\,\overline{w}f\overline{b}\overline{e}$	1
$pw\overline{f}be$	$\mathbf{a}_1^+\mathbf{a}_2^-\mathbf{a}_3^+\mathbf{a}_4^+\mathbf{a}_5^+$	$p\overline{w}\overline{f}be$	$\mathbf{a}_1^-\mathbf{a}_3^+\mathbf{a}_4^+\mathbf{a}_5^+$	$\overline{p}w\overline{f}be$	$\mathbf{a}_2^-\mathbf{a}_5^+$	$\overline{p}\,\overline{w}\overline{f}be$	$\mathbf{a}_5^+$
$pw\overline{f}b\overline{e}$	$\mathbf{a}_1^+\mathbf{a}_2^-\mathbf{a}_3^+\mathbf{a}_4^+\mathbf{a}_5^-$	$p\overline{w}\overline{f}b\overline{e}$	$\mathbf{a}_1^-\mathbf{a}_3^+\mathbf{a}_4^+\mathbf{a}_5^-$	$\overline{p}w\overline{f}b\overline{e}$	$\mathbf{a}_2^-\mathbf{a}_5^-$	$\overline{p}\,\overline{w}\overline{f}b\overline{e}$	$\mathbf{a}_5^-$
$pw\overline{f}\,\overline{b}e$	$\mathbf{a}_1^+\mathbf{a}_2^-\mathbf{a}_3^+\mathbf{a}_4^-$	$p\overline{w}\overline{f}\,\overline{b}e$	$\mathbf{a}_1^-\mathbf{a}_3^+\mathbf{a}_4^-$	$\overline{p}w\overline{f}\,\overline{b}e$	$\mathbf{a}_2^-$	$\overline{p}\,\overline{w}\overline{f}\,\overline{b}e$	1
$pw\overline{f}\,\overline{b}\overline{e}$	$\mathbf{a}_1^+\mathbf{a}_2^-\mathbf{a}_3^+\mathbf{a}_4^-$	$p\overline{w}\overline{f}\,\overline{b}\overline{e}$	$\mathbf{a}_1^-\mathbf{a}_3^+\mathbf{a}_4^-$	$\overline{p}w\overline{f}\,\overline{b}\overline{e}$	$\mathbf{a}_2^-$	$\overline{p}\,\overline{w}\overline{f}\,\overline{b}\overline{e}$	1

Table 1. Conditional structures for Example 2

So, $\alpha_5^+ > \alpha_5^-$, and therefore,

$$\begin{aligned}&\alpha_4^+\alpha_5^+ \cdot \max\{\alpha_1^+\alpha_2^+\alpha_3^-, \alpha_1^+\alpha_2^-\alpha_3^+, \alpha_1^-\alpha_3^-, \alpha_1^-\alpha_3^+\}\\ &\quad > \alpha_4^- \cdot \max\{\alpha_1^+\alpha_2^+\alpha_3^-, \alpha_1^+\alpha_2^-\alpha_3^+, \alpha_1^-\alpha_3^-, \alpha_1^-\alpha_3^+\}\end{aligned}$$

which implies $\alpha_4^+\alpha_5^+ > \alpha_4^-$.

Although penguins are exceptional birds in that they are not able to fly, they should still inhibit the property of laying eggs from their superclass *birds*. This is not guaranteed in general possibilistic inference – the conditional $(e|b)$ is "drowned" by the conditionals $(w|p), (\overline{f}|p)$ which obtain a higher degree of possibility (see [BDP97, p. 273]). By observing conditional structures, however, the impact of $(e|b)$ can be preserved to establish the validity of $(e|p)$: For any c-representation π of $\mathcal{R}$, we compute

$$\begin{aligned}\pi(ae) = \alpha_0 \max\{&\alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^+\alpha_5^+, \alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^-, \alpha_1^+\alpha_2^-\alpha_3^+\alpha_4^+\alpha_5^+, \alpha_1^+\alpha_2^-\alpha_3^+\alpha_4^-,\\ &\alpha_1^-\alpha_3^-\alpha_4^+\alpha_5^+, \alpha_1^-\alpha_3^-\alpha_4^-, \alpha_1^-\alpha_3^+\alpha_4^+\alpha_5^+, \alpha_1^-\alpha_3^+\alpha_4^-\}\\ \pi(a\overline{e}) = \alpha_0 \max\{&\alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^+\alpha_5^-, \alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^-, \alpha_1^+\alpha_2^-\alpha_3^+\alpha_4^+\alpha_5^-, \alpha_1^+\alpha_2^-\alpha_3^+\alpha_4^-,\\ &\alpha_1^-\alpha_3^-\alpha_4^+\alpha_5^-, \alpha_1^-\alpha_3^-\alpha_4^-, \alpha_1^-\alpha_3^+\alpha_4^+\alpha_5^-, \alpha_1^-\alpha_3^+\alpha_4^-\}\end{aligned}$$

By comparing the sets which contribute to the maxima element by element, we see that $\pi(pe) > \pi(p\overline{e})$, due to $\alpha_5^+ > \alpha_5^-$ and $\alpha_4^+\alpha_5^+ > \alpha_4^-$ (e.g., $\alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^+\alpha_5^+ > \alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^-, \alpha_1^+\alpha_2^+\alpha_3^-\alpha_4^+\alpha_5^-$). This means, that $\pi \models (e|p)$, for all c-representations π of $\mathcal{R}$. Hence, $\mathcal{R} \mid\!\sim^c (e|p)$.

C-representations handle conditional interactions in a numerical way, but also allow us to use structural arguments when considering conditional entailment.

Our general approach based on conditional structures and c-representations may be improved when combined with the *principle of least commitment* and with the *principle of minimum specificity*, that is, *choose the α's in a least committed and least specific way.* The consequences of these principles are twofold:

1. To allow worlds verifying conditionals to be as normal and least specific as possible, set $\alpha_i^+ = 1$ for $1 \leqslant i \leqslant n$. Since $\mathcal{R}$ is consistent iff there are worlds not falsifying any of the conditionals in $\mathcal{R}$, this implies $\alpha_0 = 1$. Furthermore, to comply with intuition (falsifying conditionals should decrease the possibility of worlds), we choose $\alpha_i^- < 1$ for $1 \leqslant i \leqslant n$.
2. In order not to introduce any relations which are not explicitly sanctioned by the representation postulate $\pi \models \mathcal{R}$, choose $\alpha_1^-, \ldots, \alpha_n^-$ according to (6), but

in a least committed way. This amounts to partition $\mathcal{R}$ into sets $\mathcal{R}_1, \ldots, \mathcal{R}_m$ of conditionals, where conditionals in subsets with lower indices can be established more easily than those in subsets with higher indices. We may then choose α_i^- such that for all conditionals in one set $\mathcal{R}_j$, all α_i^- are the same, and the α_i^- are then chosen following some discrete scale, e.g., $\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \ldots$, i.e. $\alpha_i^- = \frac{1}{2^j}$ for all conditionals $(B_i|A_i) \in \mathcal{R}_j$.

A c^*-representation of $\mathcal{R}$ is a c-representation $\pi^*(\omega) = \alpha_0 \prod\limits_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^+ \prod\limits_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^-$ of $\mathcal{R}$ whose factors α_i^+, α_i^- satisfy both (1) and (2) above. In particular, $\alpha_0 = 1$, $\alpha_i^+ = 1$, $\alpha_i^- < 1$, and

$$\left(\max_{\omega \models A_i B_i} \prod_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \alpha_j^- \right) > \alpha_i^- \left(\max_{\omega \models A_i \overline{B}_i} \prod_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \alpha_j^- \right) \tag{8}$$

for $1 \leqslant i \leqslant n$. Let $\mid\!\sim^*$ be the inference relation based on considering all c^*-representations:

$$\mathcal{R} \mid\!\sim^* (B|A) \quad \text{iff} \quad \pi^* \models (B|A) \text{ for all } c^*\text{-representations } \pi^* \text{ of } \mathcal{R}.$$

The approach we thereby obtain is very similar to the *lcd-approach* of Benferhat, Saffiotti and Smets [BSS00]. Instead of *proto-infinitesimals* and *infinitesimals* which are used in that paper, we use formal symbols $\mathbf{a}_1^-, \ldots, \mathbf{a}_n^-$ and positive real numbers $\alpha_1^-, \ldots, \alpha_n^-$. So our c-representations are nearly the same as their interpretations of proto-infinitesimals. Their *auto-deduction principle* is the same as our (quite implicit) representation postulate $\pi \models \mathcal{R}$. They base their considerations upon belief functions, we use possibility distributions (which correspond to plausibility functions). The basic idea to construct infinitesimal lcd-belief functions is to interpret proto-infinitesimals according to a *least committed ϵ-stratification*, where each ϵ-stratification corresponds to a partition of $\mathcal{R}$, and least committedness comes down to minimizing the number of classes, and to assign as many proto-infinitesimals (i.e. α_1^-) as possible to classes with low index. This amounts exactly to choose $\alpha_1^-, \ldots, \alpha_n^-$ according to (6) in a least committed way. So $\mid\!\sim^*$ and $\mid\!\sim_{lcd}$ of [BSS00] coincide.

The crucial difference between our work and that of Benferhat, Saffiotti and Smets [BSS00] is how the special product form of possibility distributions (see (5)), or belief functions, respectively, is motivated: Benferhat, Saffiotti and Smets derive the product form from Dempster's combination rule, whereas in our framework, (5) arises from an underlying conditional theory, which is much more general and applies not only in a possibilistic environment, but also in other quantitative (probabilistic) and semi-quantitative (ordinal conditional functions) environments and has also important consequences for handling conditionals adequately in a purely qualitative manner (cf. [KI01a]). Actually, in the framework of ordinal conditional functions, the application of our conditional approach turns out to be a generalization of system-Z^* (cf. [GMP93]; see [KI01b]). Note that the ideas of considering conditional structures and using conditional indifference were born when we investigated maximum entropy inference (cf. [KI98]; conditional indifference was termed *conditional preservation* there, as we actually considered probabilistic revision). So the same crucial principle, namely that of conditional indifference, underlies our possibilistic representation approach as well as the probabilistic maximum entropy approach, the generalized system-Z^* and the lcd-approach of Benferhat, Saffiotti and Smets. It is this principle that guarantees a most appropriate handling of conditionals in uncertain reasoning.

In Example 2, we showed that our approach solves the drowning problem; in particular, it also solves the *blocking of property inheritance problem* (cf. [BDP97]). Lemma 1 proves it to observe irrelevance. In the next example, we will show that even the inference relation $\mid\!\sim^*$ based on c^*-representations does not introduce conditional dependencies where there are none – it respects *ambiguity* (also cf. [BSS00]):

Example 3. Consider the following set of conditionals $\mathcal{R}$:

$(f\|b)$	*Birds fly.*
$(b\|p)$	*Penguins are birds.*
$(\overline{f}\|p)$	*Penguins do not fly.*
$(f\|m)$	*Objects with metal wings fly.*

The factors $\alpha_1^-, \ldots, \alpha_4^-$ of any c^*-representation π^* of $\mathcal{R}$ must be chosen in a least committed way so as to satisfy (8), which means $\alpha_1^- < 1, \alpha_2^- < \max\{\alpha_1^-, \alpha_3^-\}, \alpha_3^- < \max\{\alpha_1^-, \alpha_2^-\}, \alpha_4^- < 1$. Therefore, c^*-representations are obtained by choosing $\alpha_2^- = \alpha_3^- < \alpha_1^- = \alpha_4^- < 1$.

If we now see Super-Tweety, the penguin with metal wings, waddling around, can we decide whether it flies or not? In our rule base $\mathcal{R}$, there is evidence in favor of and against Super-Tweety's flying, but we are not able to make up our mind. And indeed, this ambiguity is reflected by c^*-representations: $\pi^*(pmf) = \max\{\alpha_3^-, \alpha_2^-\alpha_3^-\} = \alpha_3^-$ and $\pi^*(pm\overline{f}) = \max\{\alpha_1^-\alpha_4^-, \alpha_2^-\alpha_4^-\} = \alpha_4^- \cdot \max\{\alpha_1^-, \alpha_2^-\} = \alpha_1^-\alpha_4^-$, and α_3^- and $\alpha_1^-\alpha_4^-$ are incomparable. So neither $\mathcal{R} \mid\!\sim^* (f|pm)$ nor $\mathcal{R} \mid\!\sim^* (\overline{f}|pm)$. The same holds if we also take into account that Super-Tweety is a bird – neither $(f|pbm)$ nor $(\overline{f}|pbm)$ are accepted.

6 Revising possibility distributions by conditional beliefs

Suppose a prior possibility distribution π is given, and some new conditionals $\mathcal{R} \subseteq (\mathcal{L} \mid \mathcal{L})$ have to be incorporated, so that a revised possibility distribution π^* results. The crucial point is to maintain the conditional structures in π as far as possible under revision, which amounts to postulate that the function π^*/π should be indifferent with respect to $\mathcal{R}$ (for further details, also cf. [KI99]).

******** noch anpassen! ******* The theory of conditional structures also provides a solid foundation to deal with conditionals appropriately when revising epistemic states by (propositional and conditional) beliefs. Minimality of change is a crucial paradigm for belief revision, and a "principle of conditional preservation" is to realize this idea of minimality when conditionals are involved in change. Minimizing absolutely the changes in conditional beliefs, as in [BG93], is an important proposal to this aim, but it does not always lead to intuitive results (see [DP97]). The idea we will develop here rather aims at *preserving the conditional structure of knowledge* within an epistemic state which, in this section, we assume to be represented by an OCF κ. Similar results can be stated for probability functions, too.

We just explained what it means for an OCF, κ, to follow the structure imposed by a set of conditionals $\mathcal{R}$ on the possible worlds by introducing the notion of conditional indifference. Pursuing this approach further in the framework of belief revision, a revision of κ by simultaneously incorporating the conditionals in $\mathcal{R}$, $\kappa^* = \kappa * \mathcal{R}$, can be said to preserve the conditional structure of κ with respect to $\mathcal{R}$ if the *relative change function* $\kappa^* - \kappa$ is indifferent with respect to $\mathcal{R}$.

Definition 3. *Let κ be an OCF, and let $\mathcal{R}$ be a (finite) set of conditionals. Let $\kappa^* = \kappa * \mathcal{R}$ denote the result of revising κ by $\mathcal{R}$. κ^* is* indifferent with respect to $\mathcal{R}$ and κ *iff $(\kappa^* - \kappa)(\widehat{\omega}_1) = (\kappa^* - \kappa)(\widehat{\omega}_2)$ whenever $\sigma_{\mathcal{R}}(\widehat{\omega}_1) = \sigma_{\mathcal{R}}(\widehat{\omega}_2)$ and $\widehat{\omega}_1 \equiv_{\top} \widehat{\omega}_2$, for $\widehat{\omega}_1, \widehat{\omega}_2 \in \widehat{\Omega}$.*

The principle of conditional preservation is now realized as an indifference property:

Definition 4. *Let κ be an OCF, and let $\mathcal{R}$ be a finite set of conditionals. A revision $\kappa^* = \kappa * \mathcal{R}$ satisfies the* principle of conditional preservation *iff κ^* is indifferent with respect to $\mathcal{R}$ and κ.*

So $\kappa * \mathcal{R}$ satisfies the principle of conditional preservation if any change in plausibility is clearly and unambigously induced by $\mathcal{R}$. The next theorem characterizes revisions of ordinal conditional functions that satisfy the principle of conditional preservation. The theorem is obvious by observing Theorem ??.

Theorem 3. *Let κ, κ^* be OCF's, and let $\mathcal{R} = \{(B_1|A_1), \ldots, (B_n|A_n)\}$ be a (finite) set of conditionals in $(\mathcal{L} \mid \mathcal{L})$. A revision $\kappa^* = \kappa * \mathcal{R}$ satisfies the principle of conditional preservation iff there are numbers $\kappa_0, \kappa_i^+, \kappa_i^- \in \mathbb{Q}, 1 \leqslant i \leqslant n$, such that for all $\omega \in \Omega$,*

$$\kappa^*(\omega) = \kappa(\omega) + \kappa_0 + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \kappa_i^+ + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \kappa_i^- \tag{9}$$

An analogous theorem for probabilistic revisions shows that also the minimum cross-entropy method satisfies the principle of conditional preservation (cf. [KI98,KI01a]).

Comparing Theorems ?? and 3 with one another, we see that an OCF κ is indifferent with respect to a finite set of conditionals $\mathcal{R}$ iff it can be taken as a revision $\kappa_0 * \mathcal{R}$ satisfying the principle of conditional preservation, where $\kappa_0(\omega) = 0$ for all $\omega \in \Omega$ is the *uniform ordinal conditional function*. So knowledge representation and belief revision can be dealt with in a common framework, based on conditionals as crucial pieces of knowledge. At least for this reason, the topic of revising epistemic states by conditional beliefs should not be looked upon as a superfluous artifact, but should be appreciated as providing a most general and enhanced framework for representing and revising knowledge. Due to the fact, that propositions, A, can be considered as degenerated conditionals, $(A|T)$, the theory may also be applied to propositional belief revision. In fact, the principle of conditional preservation, as defined here, can be shown to cover the postulates for iterative belief revision of [DP97] (cf. [KI01a]).

7 Knowledge discovery/Rule extraction

The concept of conditional structures and the principle of conditional indifference provide a powerful framework for inductive knowledge representation and belief revision in possibility theory. Embedding the techniques into a group theoretical framework makes calculations more convenient and allows an elegant presentation of the basic ideas. In this section, we will show how this link between group theory and conditionals may be used to extract conditional relationships from data.

Let a possibility distribution π be given. The aim is to find a most concise set $\mathcal{R}$ of conditionals of which π is a c-representation. This amounts to find $\mathcal{R} \subset (\mathcal{L} \mid \mathcal{L})$ such that $\pi \models \mathcal{R}$ and $ker_0\ \sigma_{\mathcal{R}} \subseteq ker_0\ \pi$ (where π is the induced homomorphism, cf. 1). So numerical relationships within π – represented by elements of $ker_0\ \pi$ – are taken as manifestations of conditional interactions. Then these numerical relationships are translated into equations between group elements. Solving these equations means disentangling conditional interactions.

Our method is based on the following prerequisites:

- We will concentrate on the discovery of single-elementary conditionals, that is, we assume $\mathcal{R}$ to be a set of single-elementary conditionals. This restriction should not be considered as a heavy drawback, bearing in mind the expressibility of single-elementary conditionals.

– To exploit numerical relationships successfully, we must presuppose, that no equation $\pi(\widehat{\omega}) = 1$ holds accidentally, but that any of these equations is clearly induced by $\mathcal{R}$. Therefore, we assume

$$ker_0\ \pi \subseteq ker_0\ \sigma_{\mathcal{R}} \tag{10}$$

This last prerequisite (10) is a bit problematic in a possibilistic framework, because possibility values are often leveled for the sake of minimum specificity, and this leveling will usually blur conditional structures. In particular, c^*-representations which minimize conditional dependencies to the debit of conditional clearness, are bad candidates for conditional knowledge discovery. Nevertheless, in spite of these objections, our method is interesting in that it provides a completely new and theoretically profound approach to discover rules from data.

So assume $\mathcal{R} = \{(b_1|A_1), \ldots, (b_n|A_n)\}$ is an existing, but hidden set of single-elementary conditionals, such that (10) holds. Let us further suppose that $ker_0\ P$ (or parts of it) is known from exploiting numerical relationships. Let $\sigma_{\mathcal{R}} : \widehat{\Omega} \to \mathcal{F}_{\mathcal{R}} = \langle \mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^- \rangle$ denote a conditional structure homomorphism with respect to $\mathcal{R}$.

Our method is a *bottom-up approach* generalizing conditionals in accordance with the conditional structure revealed by $ker_0\ P$. We start with considering basic single-elementary conditionals, which are single-elementary conditionals with antecedents of maximal length. For each atom $v \in \mathcal{L}$, choose an arbitrary, but fixed numbering of the remaining atoms $w \neq v$, $(w_0, w_1, \ldots, w_{\#(atoms)-1})$. Then *basic single-elementary conditionals* are conditionals of the form

$$\psi_{v,l} = (v \mid \bigwedge_j w_j^{\epsilon_j}) \tag{11}$$

with $\epsilon_j \in \{0,1\}, w_j^1 := w_j, w_j^0 := \overline{w_j}, 0 \leqslant j \leqslant \#(atoms) - 1$ and $l = \sum_j \epsilon_j 2^j$. We will abbreviate the antecedent of $\psi_{v,l}$ by $C_{v,l}$. Let

$$\mathcal{B} = \{\psi_{v,l} \mid v \text{ atom in } \mathcal{L}, 0 \leqslant l \leqslant 2^{\#(atoms)-1} - 1\}$$

denote the set of all basic single-elementary conditionals in $(\mathcal{L} \mid \mathcal{L})$, and let $\mathcal{F}_{\mathcal{B}} = \langle \mathbf{b}_{v,l}^+, \mathbf{b}_{v,l}^- \mid v \text{ atom in } \mathcal{L}, 0 \leqslant l \leqslant 2^{\#(atoms)-1} - 1 \rangle$ be the free abelian group corresponding to $\mathcal{B}$ with conditional structure homomorphism $\sigma_{\mathcal{B}} : \widehat{\Omega} \to \mathcal{F}_{\mathcal{B}}$,

$$\sigma_{\mathcal{B}} = \prod_{v,l} \sigma_{v,l},\ \sigma_{v,l}(\omega) = \begin{cases} \mathbf{b}_{v,l}^+, & \text{if } \omega = C_{v,l} v \\ \mathbf{b}_{v,l}^-, & \text{if } \omega = C_{v,l} \overline{v} \\ 1, & \text{else} \end{cases}$$

Note that $\sigma_{\mathcal{B}}$ and $\mathcal{F}_{\mathcal{B}}$ are known, whereas $\sigma_{\mathcal{R}}$ and $\mathcal{F}_{\mathcal{R}}$ are not. We only know the kernel, $ker_0\ \sigma_{\mathcal{R}}$, of $\sigma_{\mathcal{R}}$, which is, by assuming faithfulness, the same as the kernel, $ker_0\ \pi$, of π. Now, to establish a connection between what is obvious ($\mathcal{B}$) and what is searched for ($\mathcal{R}$), we define a homomorphism $g : \mathcal{F}_{\mathcal{B}} \to \mathcal{F}_{\mathcal{R}}$ via

$$g(\mathbf{b}_{v,l}^{\pm}) = \prod_{\substack{1 \leqslant i \leqslant n \\ b_i = v, C_{v,l} \models A_i}} \mathbf{a}_i^{\pm}, \tag{12}$$

Actually, g is a "phantom" which is not explicitly given, but only known to exist. Its crucial meaning for the knowledge discovery task is revealed by the following theorem and its corollary:

Theorem 4. *Let $g : \mathcal{F}_{\mathcal{B}} \to \mathcal{F}_{\mathcal{R}}$ be as in (12). Then $\sigma_{\mathcal{R}} = g \circ \sigma_{\mathcal{B}}$.*

Corollary 2. *$\widehat{\omega} \in ker_0\ \sigma_{\mathcal{R}} = ker_0\ \pi$ iff $\widehat{\omega} \in \widehat{\Omega}_0$ and $\sigma_{\mathcal{B}}(\widehat{\omega}) \in ker\, g$.*

This means, that numerical relationships observed in π (and represented by elements of $ker_0\ \pi$) translate into group theoretical equations modulo the kernel of g:

Proposition 3. *Let* $\widehat{\omega} = \omega_1^{r_1} \cdot \ldots \cdot \omega_m^{r_m} \in \widehat{\Omega}_0$. *Then* $\sigma_{\mathcal{B}}(\omega_1^{r_1} \cdot \ldots \cdot \omega_m^{r_m}) \in ker\, g$ *iff for all atoms* v *in* $\mathcal{L}$,

$$\prod_{\substack{1 \leqslant k \leqslant m \\ \omega_k = C_{v,l} v}} (\mathbf{b}_{v,l}^+)^{r_k}, \quad \prod_{\substack{1 \leqslant k \leqslant m \\ \omega_k = C_{v,l} \overline{v}}} (\mathbf{b}_{v,l}^-)^{r_k} \in ker\, g. \tag{13}$$

So each (generating) element of $ker_0\ \sigma_{\mathcal{R}}$ gives rise to an equation modulo $ker\, g$ for the generators $\mathbf{b}_{v,l}^+, \mathbf{b}_{v,l}^-$ of $\mathcal{F}_{\mathcal{B}}$. Moreover, Proposition 3 allows us to split up equations modulo $ker_0\ g$ to handle each atom separately as a consequent of conditionals.

The idea of our data mining procedure is to exploit the relations mod $ker\ g$ holding between the group elements $\mathbf{b}_{v,l} \in \mathcal{F}_{\mathcal{B}}$ with the aim to define a finite sequence of sets $\mathcal{S}^{(0)}, \mathcal{S}^{(1)}, \ldots$ of conditionals approximating $\mathcal{R}$:

$$ker_0\ \sigma_{\mathcal{S}^{(0)}} \subseteq ker_0\ \sigma_{\mathcal{S}^{(1)}} \subseteq \ldots \subseteq ker_0\ \sigma_{\mathcal{R}} \tag{14}$$

The set $\mathcal{B}$ of basic single elementary conditionals proves to be an ideal starting point:

Lemma 2. $\sigma_{\mathcal{B}}$ *is injective, i.e.* $ker_0\ \sigma_{\mathcal{B}} = \{1\}$.

So $\sigma_{\mathcal{B}}$ provides the most finely grained conditional structure on $\widehat{\Omega}$: No different elements $\widehat{\omega}_1 \neq \widehat{\omega}_2$ are equivalent with respect to $\mathcal{B}$. So set $\mathcal{S}^{(0)} := \mathcal{B}$ and $g^{(0)} := g$.

By solving the equations modulo $ker\ g^{(t)}$, $t \geqslant 0$, we obtain a finite sequence of sets $\mathcal{S}^{(t)}$ of single-elementary conditionals $\varphi_{v,j}^{(t)} = (v|D_{v,j}^{(t)})$ with corresponding group generators ${\mathbf{s}_{v,j}^{(t)}}^+, {\mathbf{s}_{v,j}^{(t)}}^-$ such that $ker_0\ \sigma_{\mathcal{S}^{(0)}} \subseteq \ldots ker_0\ \sigma_{\mathcal{S}^{(t)}} \subseteq ker_0\ \sigma_{\mathcal{S}^{(t+1)}} \subseteq \ldots \subseteq ker_0\ \sigma_{\mathcal{R}}$, as desired, together with homomorphisms $g^{(t)}$ describing their relationship to $\mathcal{R}$. Each $\mathcal{S}^{(t+1)}$ arises from $\mathcal{S}^{(t)}$ by eliminating conditionals, or by aggregating conditionals with identical consequents in conjoining their antecedents. These operations are induced by group equations modulo $ker_0\ g^{(t)}$ (equivalence modulo $ker_0\ g^{(t)}$ is denoted by $\equiv_{g^{(t)}}$ in the following, and the $+, -$-signs may be omitted): For instance, if $\mathbf{s}_{v,j}^{(t)} \equiv_{g^{(t)}} 1$, then simply eliminate $\varphi_{v,j}^{(t)}$, and if

$$\mathbf{s}_{v,j_0}^{(t)} \equiv_{g^{(t)}} \mathbf{s}_{v,j_1}^{(t)} \ldots \mathbf{s}_{v,j_m}^{(t)}, \tag{15}$$

then

1. eliminate $\varphi_{v,j_0}^{(t)}$ from $\mathcal{S}^{(t)}$;
2. replace each $\varphi_{v,j_k}^{(t)}$ by $\varphi_{v,j_k}^{(t+1)} = \varphi_{v,j_0}^{(t)} \sqcup \varphi_{v,j_k}^{(t)} = (v \mid D_{v,j_0}^{(t)} \vee D_{v,j_k}^{(t)})$, for $1 \leqslant k \leqslant m$; set $D_{v,j_k}^{(t+1)} = D_{v,j_0}^{(t)} \vee D_{v,j_k}^{(t)}$, $1 \leqslant k \leqslant m$;
3. retain all other $\varphi_{w,l}^{(t)}$.

For the formal details and theoretical justifications, cf. [KI00b]. We will briefly illustrate the method by an example.

Example 4. We consider formulas involving the three atomic propositions a - *being a student*, b - *being young*, and c - *being single (i.e. unmarried)*. The distribution π over a, b, c is given as follows:

ω	$\pi(\omega)$	ω	$\pi(\omega)$	ω	$\pi(\omega)$	ω	$\pi(\omega)$
abc	0.720	$ab\overline{c}$	0.900	$a\overline{b}c$	0.12	$a\overline{b}\overline{c}$	0.300
$\overline{a}bc$	0.800	$\overline{a}b\overline{c}$	1.000	$\overline{a}\overline{b}c$	0.400	$\overline{a}\overline{b}\overline{c}$	1.000

Important relationships between possibility values are $\pi(\overline{a}b\overline{c}) = \pi(\overline{a}\,\overline{b}\overline{c})$, $\frac{\pi(abc)}{\pi(ab\overline{c})} = \frac{\pi(\overline{a}bc)}{\pi(\overline{a}b\overline{c})}$, $\frac{\pi(a\overline{b}c)}{\pi(a\overline{b}\overline{c})} = \frac{\pi(\overline{a}\,\overline{b}c)}{\pi(\overline{a}\,\overline{b}\overline{c})}$. We list the twelve basic single-elementary conditionals of $\mathcal{B}$ which we start with:

$$
\begin{array}{lll}
\psi_{a,0}{=}(a \mid \overline{b}\overline{c}) & \psi_{b,0}{=}(b \mid \overline{a}\,\overline{c}) & \psi_{c,0}{=}(c \mid \overline{a}\overline{b}) \\
\psi_{a,1}{=}(a \mid \overline{b}c) & \psi_{b,1}{=}(b \mid \overline{a}c) & \psi_{c,1}{=}(c \mid \overline{a}b) \\
\psi_{a,2}{=}(a \mid b\overline{c}) & \psi_{b,2}{=}(b \mid a\overline{c}) & \psi_{c,2}{=}(c \mid a\overline{b}) \\
\psi_{a,3}{=}(a \mid bc) & \psi_{b,3}{=}(b \mid ac) & \psi_{c,3}{=}(c \mid ab)
\end{array}
$$

with corresponding generators $\mathbf{b}^{+}_{v,l}, \mathbf{b}^{-}_{v,l}$ of $\mathcal{F}_{\mathcal{B}}$. The numerical relationships yield the following group equations :

$$
\begin{array}{llll}
\mathbf{b}^{-}_{a,2}\mathbf{b}^{+}_{b,0}\mathbf{b}^{-}_{c,1} & = \sigma_{\mathcal{B}}(\overline{a}b\overline{c}) & = \sigma_{\mathcal{B}}(\overline{a}\,\overline{b}\overline{c}) & = \mathbf{b}^{-}_{a,0}\mathbf{b}^{-}_{b,0}\mathbf{b}^{-}_{c,0} \\
\mathbf{b}^{+}_{a,3}\mathbf{b}^{+}_{b,3}\mathbf{b}^{+}_{c,3} \cdot \mathbf{b}^{-}_{a,2}\mathbf{b}^{+}_{b,0}\mathbf{b}^{-}_{c,1} & = \sigma_{\mathcal{B}}(abc)\sigma_{\mathcal{B}}(\overline{a}b\overline{c}) & = \sigma_{\mathcal{B}}(ab\overline{c})\sigma_{\mathcal{B}}(\overline{a}bc) & = \mathbf{b}^{+}_{a,2}\mathbf{b}^{+}_{b,2}\mathbf{b}^{-}_{c,3} \cdot \mathbf{b}^{-}_{a,3}\mathbf{b}^{+}_{b,1}\mathbf{b}^{+}_{c,1} \\
\mathbf{b}^{+}_{a,1}\mathbf{b}^{-}_{b,3}\mathbf{b}^{+}_{c,2} \cdot \mathbf{b}^{-}_{a,0}\mathbf{b}^{-}_{b,0}\mathbf{b}^{-}_{c,0} & = \sigma_{\mathcal{B}}(a\overline{b}c)\sigma_{\mathcal{B}}(\overline{a}\,\overline{b}\overline{c}) & = \sigma_{\mathcal{B}}(a\overline{b}\overline{c})\sigma_{\mathcal{B}}(\overline{a}\,\overline{b}c) & = \mathbf{b}^{+}_{a,0}\mathbf{b}^{-}_{b,2}\mathbf{b}^{-}_{c,2} \cdot \mathbf{b}^{-}_{a,1}\mathbf{b}^{-}_{b,1}\mathbf{b}^{+}_{c,0}
\end{array}
$$

Considering these equations for each atom a, b, c separately and omitting the $\{+,-\}$-signs, we obtain the following equivalences modulo an appropriate subgroup of $\mathcal{F}_{\mathcal{B}}$: $\mathbf{b}_{a,0} \equiv \mathbf{b}_{a,1} \equiv \mathbf{b}_{a,2} \equiv \mathbf{b}_{a,3}$, $\mathbf{b}_{c,0} \equiv \mathbf{b}_{c,1} \equiv \mathbf{b}_{c,2} \equiv \mathbf{b}_{c,3}$, $\mathbf{b}_{b,0} \equiv 1$, $\mathbf{b}_{b,3} \equiv \mathbf{b}_{b,1}\mathbf{b}_{b,2}$. Eliminating $\psi_{b,0}$ and joining conditionals according to these equations results in the set $\mathcal{S} = \{(\overline{a}|\top), (\overline{c}|\top), (b|a), (b|c)\}$ of conditionals of which the above possibility distribution π is a c-representation. Note that π is *not* a c^*-representation.

8 Conditional structures and conditional events

We announced our approach to conditionals as a substantially new one, capturing the nature of conditionals peculiarly well. The results and examples presented above illustrate this, but what is actually new and different, compared with other, established conditional theories?

First formal work on conditionals dates back to the days of Boole. *Conditional events* were investigated particularly intensively in the nineties of the past century, when their crucial meaning for knowledge representation in AI had become apparent (for a historical review, cf. [NW94]; for a survey of conditional logics, cf. [Nut80]). Lewis [Lew76] made clear that conditionals cannot be handled as boolean entities, so new ways of representing and dealing with conditionals had to be explored. A very popular approach, e.g., is the following one (for an overview of important approaches to conditionals, see [DGC94]):

It is generally agreed that two conditionals $(B|A), (D|C)$ are equivalent, $(B|A) \equiv (D|C)$, iff $A \equiv C$ and $AB \equiv CD$. So it seems adequate to define a conditional $(B|A)$ by a set of boolean propositions X all fulfilling $XB = AB$: $(B|A) := \{X \in \mathcal{L} \mid XB = AB\} = \{X \in \mathcal{L} \mid AB \leqslant X \leqslant \overline{B} \vee A\}$, where $\leqslant$ is the usual entailment relation between boolean propositions (see [GN88]). That is, a conditional is identified with an interval of boolean propositions. This gives rise to a partial ordering $\leqslant$ between conditionals which is also frequently understood as an entailment relation between conditionals (cf. [DP91b,BDP97]):

$$(D|C) \leqslant (B|A) \quad \text{iff} \quad CD \leqslant AB \text{ and } A\overline{B} \leqslant C\overline{D} \tag{16}$$

Walker lists in [Wal94] several possible representations of conditional events and investigates different ways of defining logical connectives between conditionals. The representation that seems most adequate to compare our approach with others is the one that identifies a conditional $(B|A)$ with an ordered pair of its crucial components $AB, A\overline{B}$:

$$(B|A) \sim (AB, A\overline{B}), \tag{17}$$

so that $(\mathcal{L} \mid \mathcal{L}) \sim \{(X, Y) \mid X, Y \in \mathcal{L}, XY \equiv \perp\}$. Moreover, the appropriateness of the three-valued semantics of conditionals (1) which we base our theory upon is accepted by most other researchers in the field of conditionals. So the essential ingredients to our theory are established and nearly undebatable. The difference lies in the way we make use of them. Our focusing not on logical properties but on the effects conditionals exert on possible worlds suggests another way of relating conditionals to one another. First of all, both positive and negative effects are considered to be equally important, no priority is given to the confirming component AB (as is done in most approaches). So, a conditional $(D|C)$ is called a *subconditional* of $(B|A)$, iff its effect on worlds is in line with $(B|A)$, but possibly $(D|C)$ applies to fewer worlds:

$$(D|C) \sqsubseteq (B|A) \quad \text{iff for all } \omega \in \Omega, \begin{cases} \sigma_{(D|C)}(\omega) = 1 \Rightarrow \sigma_{(B|A)}(\omega) = 1 \\ \sigma_{(D|C)}(\omega) = 0 \Rightarrow \sigma_{(B|A)}(\omega) = 0 \end{cases} \tag{18}$$

which amounts to saying that $(D|C) \sqsubseteq (B|A)$ iff $CD \leqslant AB$ and $C\overline{D} \leqslant A\overline{B}$. Note the essential difference with (16), but we still have $(B|A) \equiv (D|C)$ iff $(B|A) \sqsubseteq (D|C)$ and $(D|C) \sqsubseteq (B|A)$, due to $(B|A) \equiv (D|C)$ iff $AB \equiv CD$ and $A\overline{B} \equiv C\overline{D}$.

Further, we can define connectives $\sqcap, \sqcup$ by setting $(B|A) \sqcap (D|C) := \inf\{(B|A), (D|C)\}$ and $(B|A) \sqcup (D|C) := \sup\{(B|A), (D|C)\}$, where the supremum is defined only if $ABC\overline{D} \equiv A\overline{B}CD \equiv \perp$. It is easy to check that

$$(B|A) \sqcap (D|C) = (BD \mid AC(BD \vee \overline{B}\,\overline{D})) \tag{19}$$

$$(B|A) \sqcup (D|C) = (AB \vee CD \mid A \vee C) \tag{20}$$

Although $(B|A) \sqcup (D|C)$ in its original sense as a supremum is defined only for $ABC\overline{D} \equiv A\overline{B}CD \equiv \perp$, (20) may be used for defining a connective for arbitrary $(B|A), (D|C)$. If the condition $ABC\overline{D} \equiv A\overline{B}CD \equiv \perp$ is not fulfilled, however, then $(B|A) \sqcup (D|C)$ is not the supremum of $(B|A)$ and $(D|C)$ (as can easily be seen by considering $(B|A)$ and $(\overline{B}|A)$). $\sqcap$ and $\sqcup$ are associative and distributive operations on $(\mathcal{L} \mid \mathcal{L})$, if all required suprema exist (cf. [KI01a]). Their connections to the boolean operations $\wedge$ and $\vee$ can be seen more directly by using the representation (17)

$$\begin{aligned}
(X, Y) \sqcap (W, Z) &= (X \wedge W, Y \wedge Z), \\
(X, Y) \sqcup (W, Z) &= (X \vee W, Y\overline{W} \vee Z\overline{X}), \\
&= (X \vee W, Y \vee Z) \quad \text{if } XZ \equiv YW \equiv \perp,
\end{aligned}$$

that is, the $\wedge$- and $\vee$-operator, respectively, are applied to the confirming and the refuting components separately, respectively. Both $\sqcap$ and $\sqcup$ are truth-functional, with $((B|A) \sqcap (D|C))(\omega) = 1$ iff $(B|A)(\omega) = 1$ and $(D|C)(\omega) = 1$, and $((B|A) \sqcap (D|C))(\omega) = 0$ iff $(B|A)(\omega) = 0$ and $(D|C)(\omega) = 0$, whereas $((B|A) \sqcup (D|C))(\omega) = 1$ iff $(B|A)(\omega) = 1$ or $(D|C)(\omega) = 1$, and $((B|A) \sqcup (D|C))(\omega) = 0$ iff $(B|A)(\omega) = 0$ and $(D|C)(\omega) \neq 1$, or $(D|C)(\omega) = 0$ and $(B|A)(\omega) \neq 1$. Therefore, we obtain the following truth tables for combining conditionals by $\sqcap$ and $\sqcup$, respectively (note that for the supremum, $0 \sqcup 1$ and $1 \sqcup 0$ do not exist):

$\sqcap$	0	1	u
0	0	u	u
1	u	1	u
u	u	u	u

$\sqcup$	0	1	u
0	0	1/*undef.*	0
1	1/*undef.*	1	1
u	0	1	u

These truth-tables reveal some incompatibilities with boolean logic: We have $0 \sqcap 1 = 1 \sqcap 0 = u$, and $0 \sqcup 1$ and $1 \sqcup 0$ are not defined, if $\sqcup$ is taken as a supremum. If one conditional is confirmed and the other refuted, nothing is said how they combine.

So no boolean rules are forced upon the interactions of conditionals. Therefore, the difference between our theory and most other approaches is that we do not expect the connectives $\sqcup$ and $\sqcap$ to observe boolean laws; in particular, though $\sqcup$ and $\sqcap$ are based upon $\vee$ and $\wedge$, respectively, they cannot be taken as extensions of those boolean operators (note, however, that the generalized version of $\sqcup$ does extend $\vee$ and is identical to Calabrese's conditional $\vee$-operator; see [Cal91]). Moreover, $\sqcap$ and $\sqcup$ are not linked by DeMorgan's laws; rather, the negation $(B|A)' := (\overline{B}|A)$ acts as a homomorphism with respect to both connectives: $((B|A) \sqcap (D|C))' = (B|A)' \sqcap (D|C)'$ and $((B|A) \sqcup (D|C))' = (B|A)' \sqcup (D|C)'$.

This breaking with boolean conditional traditions also has consequences for considering facts (i.e., boolean propositions) as conditionals: Usually, a proposition A is identified with the degenerate (two-valued) conditional $(A|\top)$. This appears to be justified by considering (1). Note, however, that actually, stating A means "*A is true*", whereas $(A|\top)$ means "*A is plausible*". This distinction is respected in our framework, in that $A \wedge B$ does not correspond to $(A|\top) \sqcap (B|\top)$ $(= (AB|AB \vee \overline{AB}) \neq (AB|\top)$; the conjunction of two plausible propositions need not be plausible), nor does $(A|\top) \sqsubseteq (B|\top)$ mean $A \subseteq B$; rather, we have $(A|\top) \sqsubseteq (B|\top)$ if and only if $A \equiv B$.

In summary, it can be said that, although the roots of our approach are firmly established in the foundations of conditional theories, the thorough application of the new concept of conditional structures brings forth quite a new theory of conditionals which is non-boolean to a very large extent. We think, however, that the loss of boolean conveniences is more than compensated by gaining a rich methodological framework for handling the dynamic power of conditionals more adequately, in particular in the areas of inductive knowledge representation and belief revision.

******** noch anpassen! *********

To exemplify what it means that c-representations impose conditional structures upon worlds, we state the following proposition which relates $\sqsubseteq$ to $\mid\!\sim^c$-inference (cf. (7)) and extends Corollary 1:

Proposition 4. *$(B|A) \mid\!\sim^c (D|C)$ iff $(D|C) \sqsubseteq (B|A)$, that is, if $(D|C)$ is a subconditional of $(B|A)$.*

Proof. Any (OCF-)c-representation of $(B|A)$ has the form

$$\kappa(\omega) = \begin{cases} \kappa_0 & \text{if} \quad \omega \models \overline{A}, \\ \kappa_0 + \kappa^+ & \text{if} \quad \omega \models AB, \\ \kappa_0 - \kappa^- & \text{if} \quad \omega \models A\overline{B} \end{cases}$$

with rational numbers $\kappa_0, \kappa^+, \kappa^-$ such that $\kappa^- - \kappa^+ > 0$ (cf. (9)). If $(D|C) \sqsubseteq (B|A)$, then also $\kappa \models (D|C)$, and conversely, if $(D|C)$ is not a subconditional of $(B|A)$, then $\kappa_0, \kappa^+, \kappa^-$ can be chosen appropriately so that $\kappa \not\models (D|C)$.

So, in the simple case where only one conditional is considered, inferring by observing conditional structures means accepting all subconditionals, too. If the inference is to be based on a set of conditionals, interactions between these conditionals have to be taken into account, and c-entailment should not be expected to be this simple. A thorough investigation of $\mid\!\sim^c$ is a topic of our current research; results will soon be published in a further paper.

9 Conclusion

****** to be written ******** All results can be generalized to a framework of quantified conditionals in a straightforward way.

References

[BDP97] S. Benferhat, D. Dubois, and H. Prade. Nonmonotonic reasoning, conditional objects and possibility theory. *Artificial Intelligence*, 92:259–276, 1997.

[BDP99] S. Benferhat, D. Dubois, and H. Prade. Possibilistic and standard probabilistic semantics of conditional knowledge bases. *Journal of Logic and Computation*, 9(6):873–895, 1999.

[BG93] C. Boutilier and M. Goldszmidt. Revision by conditional beliefs. In *Proceedings 11th National Conference on Artificial Intelligence (AAAI'93)*, pages 649–654, Washington, DC., 1993.

[BSS00] S. Benferhat, A. Saffiotti, and P. Smets. Belief functions and default reasoning. *Artificial Intelligence*, 122:1–69, 2000.

[Cal91] P.G. Calabrese. Deduction and inference using conditional logic and probability. In I.R. Goodman, M.M. Gupta, H.T. Nguyen, and G.S. Rogers, editors, *Conditional Logic in Expert Systems*, pages 71–100. Elsevier, North Holland, 1991.

[DeF74] B. DeFinetti. *Theory of Probability*, volume 1,2. John Wiley and Sons, New York, 1974.

[DGC94] D. Dubois, I.R. Goodman, and P.G. Calabrese. Special issue on the conditional event algebra. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(12), 1994.

[DLP94] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D.M. Gabbay, C.H. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 439–513. Oxford University Press, 1994.

[DP91a] D. Dubois and H. Prade. Conditional objects and non-monotonic reasoning. In *Proceedings 2nd Int. Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 175–185. Morgan Kaufmann, 1991.

[DP91b] D. Dubois and H. Prade. Conditioning, non-monotonic logic and non-standard uncertainty models. In I.R. Goodman, M.M. Gupta, H.T. Nguyen, and G.S. Rogers, editors, *Conditional Logic in Expert Systems*, pages 115–158. Elsevier, North Holland, 1991.

[DP91c] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. *Artificial Intelligence*, 50:223–239, 1991.

[DP94] D. Dubois and H. Prade. A survey of belief revision and updating rules in various uncertainty models. *Intern. Journal of Intelligent Systems*, 9:61–100, 1994.

[DP97] A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.

[FH94] N. Friedman and J.Y. Halpern. Conditional logics of belief change. In *Proceedings 12th National Conference on Artificial Intelligence, AAAI-94*, 1994.

[GK97] J. Gebhardt and R. Kruse. Background and perspectives of possibilistic graphical models. In *Proceedings First International Joint Conference on Qualitative and Quantitative Practical Reasoning, ECSQARU-FAPR'97,*, pages 108–121. Springer, 1997.

[GMP93] M. Goldszmidt, P. Morris, and J. Pearl. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):220–232, 1993.

[GN88] I.R. Goodman and H.T. Nguyen. Conditional objects and the modeling of uncertainties. In M.M. Gupta and T. Yamakawa, editors, *Fuzzy Computing - Theory, Hardware and Applications*, pages 119–138. North-Holland, 1988.

[KI98] G. Kern-Isberner. Characterizing the principle of minimum cross-entropy within a conditional-logical framework. *Artificial Intelligence*, 98:169–208, 1998.

[KI99] G. Kern-Isberner. *A unifying framework for symbolic and numerical approaches to nonmonotonic reasoning and belief revision.* Department of Computer Science, FernUniversität Hagen, 1999. Habilitation thesis.

[KI00a] G. Kern-Isberner. Conditional preservation and conditional indifference (preliminary version). In *Proceedings 8th International Workshop on Nonmonotonic Reasoning (NMR'2000)*, Breckenridge, Colorado, 2000.

[KI00b] G. Kern-Isberner. Solving the inverse representation problem. In *Proceedings 14th European Conference on Artificial Intelligence, ECAI'2000*, pages 581–585, Berlin, 2000. IOS Press.

[KI01a] G. Kern-Isberner. *Conditionals in nonmonotonic reasoning and belief revision.* Springer, Lecture Notes in Artificial Intelligence, 2001. (to appear).

[KI01b] G. Kern-Isberner. Handling conditionals adequately in uncertain reasoning. In *Proceedings ECSQARU'01*, 2001. (to appear).

[KS91] H. Katsuno and K. Satoh. A unified view of consequence relation, belief revision and conditional logic. In *Proceedings Twelfth International Joint Conference on Artificial Intelligence, IJCAI-91*, pages 406–412, 1991.

[Lew76] D. Lewis. Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, 85:297–315, 1976.

[LM92] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[LS77] R.C. Lyndon and P.E. Schupp. *Combinatorial group theory.* Springer, Berlin Heidelberg New York, 1977.

[Nut80] D. Nute. *Topics in Conditional Logic.* D. Reidel Publishing Company, Dordrecht, Holland, 1980.

[NW94] H.T. Nguyen and E.A. Walker. A history and introduction to the algebra of conditional events and probability logic. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(12):1671–1675, 1994.

[Par94] J.B. Paris. *The uncertain reasoner's companion – A mathematical perspective.* Cambridge University Press, 1994.

[Rot91] H. Rott. A nonmonotonic conditional logic for belief revision, part I: Semantics and logic of simple conditionals. In A. Fuhrmann and M. Morreau, editors, *The logic of theory change*, pages 135–181. Springer, Berlin, Heidelberg, New York, 1991.

[Spo88] W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W.L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics, II*, pages 105–134. Kluwer Academic Publishers, 1988.

[Wal94] E.A. Walker. Stone algebras, conditional events, and three valued logic. *IEEE Transactions on Systems, Man, and Cybernetics (Special Issue on The Conditional Event Algebra)*, 24(12):1699–1707, 1994.