

# Qualitative and Semi-Quantitative Inductive Reasoning with Conditionals

## Technical Project Report

Christian Eichhorn<sup>1</sup> · Gabriele Kern-Isberner<sup>1</sup>

Published online: 6 June 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Conditionals like “birds fly—if bird then fly” are crucial for commonsense reasoning. In this technical project report we show that conditional logics provide a powerful formal framework that helps understanding if-then sentences in a way that is much closer to human reasoning than classical logic and allows for high-quality reasoning methods. We describe methods that inductively generate models from conditional knowledge bases. For this, we use both qualitative (like preferential models) and semi-quantitative (like Spohn’s ranking functions) semantics. We show similarities and differences between the resulting inference relations with respect to formal properties. We further report on two graphical methods on top of the ranking approaches which allow to decompose the models into smaller, more feasible components and allow for local inferences.

**Keywords** Conditionals · Nonmonotonic reasoning · Induction · System P · Networks · Ordinal conditional function · Conditional structures

---

We thank the anonymous referees for their valuable hints that helped us improving the paper. This work was supported by grant KI 1413/5 – 1 to Gabriele Kern-Isberner from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program “New Frameworks of Rationality” (SPP 1516). Christian Eichhorn is supported by this Grant.

---

✉ Christian Eichhorn  
christian.eichhorn@tu-dortmund.de

Gabriele Kern-Isberner  
gabriele.kern-isberner@cs.tu-dortmund.de

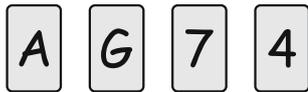
<sup>1</sup> Technische Universität Dortmund, Lehrstuhl Informatik 1, Dortmund, Germany

## 1 Introduction

The Wason selection task [34, 35] became famous for proving experimentally that people are poor deductive reasoners. In this task, people were presented with cards showing a letter on one side and a number on the other. Then, they were asked to validate the following (conditional) statement by turning as few cards as necessary: “If there is a vowel on the front side of the card, then there is an even number on its back side”. Figure 1 shows an exemplary set of cards for the Wason selection task.

The card showing “A” was turned by nearly all people (which is correct) and the card showing “G” was turned by only a small percentage of the people (of course, not everybody can be expected to be a good deductive reasoner). Problems with classical logic became obvious when considering the frequencies with which the cards showing “7” and “4”, resp., were turned by the test persons – far more persons turned “4” than “7”, which is incompatible with deductive reasoning. By turning “7”, the contrapositive form of the statement under consideration can be validated, so turning “7” is a must from the viewpoint of deduction. On the other hand, turning “4” is vacuous since this card would validate the statement in any case.

The results from such experiments have long been used to blame humans for being poor logicians, but other theories have been brought forward to explain human behaviour in the Wason task that found logic inadequate for evaluating humans, e.g., *relevance theory* [29], or a probabilistic theory that is based on information gain and claims to be even more general [25]. Quite recently, it has been pointed out that only *classical* logic is inadequate but other logics may help to model human behaviour in the Wason task: For instance in [4], logic programming under the weak completion semantics has been applied to the Wason



**Fig. 1** Exemplary card set for the Wason selection task

task and yields interesting results that distinguish between the abstract Wason task (as described above) and an analogical task having a social context. In addition, the authors also show relationships to the suppression task and abduction. Similarly, Furbach and Schon [8] succeeded in making a difference to previous findings by using deontic logic, interpreting the sentence “If there is a vowel on the front side of the card, then there is an even number on its back side” as a normative rule. Nevertheless, still the question remains what general *kinds of logic* are suitable to model human behaviour for the Wason selection task, even if no further assumptions on the context are made.

As also [4] and [8] noticed, the statement to be validated in the Wason selection task is a conditional one, so the most obvious candidate for looking for alternatives would be the framework of conditional logics. Although much work has been done in conditional logics both from symbolic and probabilistic points of views (cf., e.g., [1, 5, 21]) the understanding of an “if-then” statement as a material implication or as a “directed rule” is still the prevalent one. Conditionals as formal entities in the form  $(\psi|\phi)$  establish relationships between propositions  $\phi$  and  $\psi$  that need suitable semantics to be interpreted, and it depends on the specific kind of semantics which understanding of the conditional is appropriate. Popular semantics for conditionals in artificial intelligence are ranked models such as Spohn’s ranking functions [10, 30] (expressing plausibility, normality, and the like), modal logics [11] (expressing reachability), and—with quite different techniques and semantical annotations—probabilities [26]. But generally, the main difference between conditionals  $(\psi|\phi)$  and material implications  $\phi \Rightarrow \psi$  as formal representations of statements of the form “if  $\phi$  then  $\psi$ ” is that the acceptance of a conditional does not depend only on the truth values of its premise  $\phi$  and its conclusion  $\psi$  (so it is not *material*). Rather, it establishes a meaningful link between premise and conclusion that can be verified (in case of  $\phi \wedge \psi$ ) or falsified (in case of  $\phi \wedge \neg\psi$ ), but is vacuous if  $\neg\phi$  holds. This makes conditionals three-valued logical entities, far more expressive than the two-valued material implication. Conditionals are also different from strict or nonmonotonic rules used in [4] because they cannot be blocked, it is rather that the rule itself is defeasible.

Coming back to the Wason selection task, let us see how a basic, but truly conditional-logical view may help to understand why people turned the cards showing “7” and “4” in this order: The “4” may verify the conditional

(*even\_number|vowel*), expressing that “If a vowel is observed, then an even number on the back side of the card is expected, normal, usual, etc.”; please note that we do not presuppose a specific semantics here, even a deontic one as in [8] can be assumed. So, “4” is *not* irrelevant for the conditional while “7” may falsify the conditional but conditionals may have exceptions, so even seeing a vowel on the back side of “7” would not completely invalidate the conditional. In any case, contraposition does not hold in general for conditional reasoning such that there is no conditional-logical obligation to turn the card. Summarizing these observations, conditionals in a very general sense may not completely *explain* human behaviour in the Wason selection task, but they provide a logic that is *compatible* with it.

The project “Rational reasoning with conditionals and probabilities” in the DFG Priority Programme 1516 “New Frameworks of Rationality” investigates rational reasoning with conditionals, using conditionals as a common interface between different approaches for inference since they may encode crucial guidelines for rational human reasoning, as described above. In particular, being able to use conditionals both in qualitative and quantitative ways provides a perspective that helps to overcome the limits of specific frameworks. Here we report on the computer science part of this project; in particular, we show how conditionals can serve as a base for qualitative and semi-quantitative<sup>1</sup> approaches to generate inference relations from conditional knowledge bases. In this way, we illustrate that conditionals are not only suitable representations of commonsense knowledge but that a variety of methods and implementations are available to draw inferences from knowledge bases consisting of conditionals.

First, we recall preferential models as a basic approach for nonmonotonic inference in Sect. 2, and then show in Sect. 3 how to generate preferential models inductively using conditional structures and ranking approaches. Finally, we show that network approaches can be used to decompose the presented inductive task and associated inference calculations into smaller, more feasible components in Sect. 4. We conclude in Sect. 5.

## 2 Conditionals and Inference Mechanisms

We consider conditionals  $(\psi|\phi)$  over a standard propositional language  $\mathcal{L}$  defined from a finite set of atoms  $\Sigma$ , i.e.,  $\psi, \phi \in \mathcal{L}$ . A conditional  $(\psi|\phi)$  can be used to encode a nonmonotonic, defeasible inference  $\phi \sim \psi$ —“from  $\phi$ ,  $\psi$  can be defeasibly concluded”. A finite set of conditionals is called a *conditional knowledge base*. A major idea in

<sup>1</sup> i.e., we use numbers in a qualitative way.

nonmonotonic reasoning is to rank propositional models according to preference relations and draw conclusions by considering only the most preferred (i.e., the most plausible) ones.

A *preferential model* [23] is a triplet of a set of states, a satisfaction relation and a preferential relation. Here we use the set  $\Omega$  of (propositional) interpretations (or *possible worlds*) as states and classical satisfaction  $\models$  as satisfaction relation, so we use classical logic as a base. A preferential model induces a nonmonotonic inference relation  $\sim$  between formulas such that  $\phi \sim \psi$  iff for all models of  $\phi \wedge \neg\psi$  there is a model of  $\phi \wedge \psi$  that is strictly preferred with respect to the preferential relation. In terms of conditionals ( $\psi|\phi$ ), this means that verification ( $\phi \wedge \psi$ ) is more plausible than falsification ( $\phi \wedge \neg\psi$ ). This is the framework we build upon for reasoning inductively from conditional knowledge bases. “Inductive” here means that we aim at completing the partial knowledge given by the knowledge base by computing a preferential model. The preferential relation will be induced by the conditional knowledge base in various ways.

Given that the satisfaction relation of the preferential model behaves like the classical satisfaction relation and the preference relation is transitive and does not allow endless descending chains, preferential models ensure inference relations of high quality, satisfying the well-appreciated axiom system P [1, 18, 23] consisting of *Reflexivity*, *Left Logical Equivalence*, *Right Weakening*, *Cautious Monotony*, *Cut* and *Or* (cf. Fig. 2). This system has not only proven to be of high formal quality, but also is most useful when describing human reasoning: Empirical studies show that human reasoners make use of each of the properties of System P [3, 28], so even with recent work challenging that human reasoners use System P as an inference calculus [19], each property is useful for the inference relations we consider in this paper. We here discuss approaches of inference mechanisms that allow us to set up inference relations which satisfy System P (and further properties) and hence are capable of realising properties of commonsense reasoning.

We use the car start problem [10] as a running example for setting up a conditional knowledge base: A car usually will start ( $s$ ) if the battery is charged ( $b$ ), otherwise it usually will not start ( $\bar{s}$ ). We also know that a car usually will start if the fuel tank is sufficiently filled ( $f$ ), otherwise

it usually will not start. Additionally, we know that if the battery is charged and the fuel tank is empty, the car usually will not start, and if the battery is discharged and the fuel tank is filled, the car usually will not start. We usually switch off the headlights overnight ( $\bar{h}$ ) and the fuel-tank usually is filled. If the headlights have been left switched on overnight ( $h$ ), the battery will usually be empty; if we have switched the headlights off ( $\bar{h}$ ), the battery usually will be charged. This is formalised in the knowledge base

$$\mathcal{R} = \left\{ \begin{array}{llll} \delta_1 : (s|b), & \delta_2 : (\bar{s}|\bar{b}), & \delta_3 : (s|f), & \delta_4 : (\bar{s}|\bar{f}), \\ \delta_5 : (\bar{s}|b\bar{f}), & \delta_6 : (\bar{s}|\bar{b}f), & \delta_7 : (\bar{h}|\top), & \delta_8 : (f|\top), \\ \delta_9 : (\bar{b}|h), & \delta_{10} : (b|\bar{h}) \end{array} \right\}.$$

### 3 Inductive Conditional Reasoning

The information given by a knowledge base usually is incomplete, i.e., not every situation is specified directly by the conditionals. In the car start example, for instance, it is not stated whether the car will start if both battery and fuel-tank are empty. Here inductive methods are used to fill up missing information, creating the preference relations needed for the preferential model.

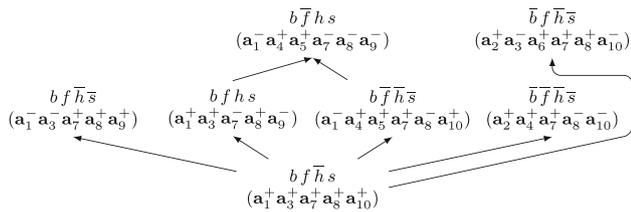
Every inductive method presented in the following generates a preferential model of some sort based on the knowledge base. This means the preferential model and hence the resulting preferential inference relation are based mainly on the information encoded in the conditionals of the knowledge base.

#### 3.1 Qualitative Inference by Conditional Structures

We start with a rather simple approach based on conditional structures to make core ideas of conditional reasoning understandable. Given a knowledge base  $\mathcal{R} = \{(\psi_1|\phi_1), \dots, (\psi_n|\phi_n)\}$  we assign to each  $(\psi_i|\phi_i) \in \mathcal{R}$  a pair of abstract symbols  $\mathbf{a}_i^+$  and  $\mathbf{a}_i^-$  to illustrate the impact of conditionals on worlds. To connect a world and the impact of a conditional to this world we define the function  $\sigma_i$  by  $\sigma_i(\omega) := \mathbf{a}_i^+$  iff  $\omega \models \phi \wedge \psi$ ,  $\sigma_i(\omega) := \mathbf{a}_i^-$  iff  $\omega \models \phi \wedge \neg\psi$  and  $\sigma_i(\omega) := 1$  iff  $\omega \models \neg\phi$ , for each  $1 \leq i \leq n$ . So,  $\mathbf{a}_i^+$  ( $\mathbf{a}_i^-$ ) indicates that  $\omega$  verifies (falsifies)

Fig. 2 Axioms of system P

(Reflexivity)	$\phi$			implies	$\phi \sim \phi$
(Cut)	$\phi \sim \psi$	and	$\phi \wedge \psi \sim \chi$	imply	$\phi \sim \chi$
(Cautious Monotony)	$\phi \sim \psi$	and	$\phi \sim \chi$	imply	$\phi \wedge \psi \sim \chi$
(Right Weakening)	$\phi \sim \psi$	and	$\psi \models \chi$	imply	$\phi \sim \chi$
(Left Logical Equivalence)	$\phi \equiv \psi$	and	$\psi \sim \chi$	imply	$\phi \sim \chi$
(Or)	$\phi \sim \chi$	and	$\psi \sim \chi$	imply	$(\phi \vee \psi) \sim \chi$



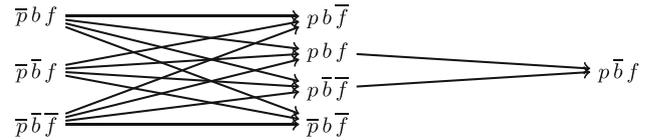
**Fig. 3** Excerpt from the conditional structures and preferences for the car start example (sources are preferred to targets)

$(\psi_i|\phi_i)$ , and 1 corresponds to non-applicability of the conditional. For a possible world  $\omega$ , the conditional structure  $\sigma(\omega)$  is the combination of all these symbols,  $\sigma(\omega) = \prod_{i=1}^n \sigma_i(\omega)$  (where product simply means concatenation) [12, 13]. In the car start example  $b\bar{f}hs$  has the conditional structure  $\sigma(b\bar{f}hs) = a_1^+ a_4^- a_5^- a_7^- a_8^- a_9^-$  because it verifies  $\delta_1$ , falsifies  $\delta_4, \delta_5, \delta_7, \delta_8$  and  $\delta_9$ , while  $\delta_2, \delta_3, \delta_6$  and  $\delta_{10}$  are not applicable.

With conditional structures we can set up a preferential relation with respect to verification or falsification of conditionals, or both. Here we use a penalising approach, i.e., a world  $\omega$  is *structurally preferred* to a world  $\omega'$  (written  $\omega \prec_\sigma \omega'$ ), iff  $\omega'$  falsifies strictly more (in terms of set inclusion) conditionals than  $\omega$ . Figure 3 shows an excerpt of the structural preferences for the car start example, where we have, e.g.,  $b\bar{f}h\bar{s} \prec_\sigma b\bar{f}hs$  since  $b\bar{f}h\bar{s}$  falsifies  $\delta_1$  and  $\delta_8$ , whereas  $b\bar{f}hs$  falsifies  $\delta_1, \delta_8, \delta_7$  and  $\delta_9$ . From this relation we obtain a preferential model and thereby an inference relation  $\vdash_{\mathcal{R}}^\sigma$ , called *structural inference*, as stated above, such that  $\phi \vdash_{\mathcal{R}}^\sigma \psi$  iff for every world that falsifies the conditional  $(\psi|\phi)$  we find a world that verifies the conditional and the latter is preferred to the former [12, 17]. In the car start example, e.g., we have  $\bar{h} \vdash_{\mathcal{R}}^\sigma s$  because for every model of  $\bar{h}\bar{s}$ , i.e., the worlds  $b\bar{f}h\bar{s}, b\bar{f}h\bar{s}, \bar{b}\bar{f}h\bar{s}$  and  $\bar{b}\bar{f}h\bar{s}$ , we find a structurally preferred model of  $\bar{h}s$ , namely  $b\bar{f}hs$  (cf. Fig. 3). An implementation of inference by conditional structures can be found in the TWEETY-library [32].

Being set up by a preferential model that meets the criteria sketched in Sect. 2, i.e., the satisfaction is classical, the preferential relation is irreflexive and transitive and the set of models is finite, the inference relation  $\vdash_{\mathcal{R}}^\sigma$  satisfies System P. Structural inference compares the complete conditional structures of the respective worlds. Therefore, for the generation of the preference relation as well as for inference tasks, it is ensured that every conditional in the knowledge base is taken into account.

However, the preferential relation  $\prec_\sigma$  is not total, so we have incomparable worlds, e.g., in the car start example the worlds  $b\bar{f}hs$  with the conditional structure  $\sigma(b\bar{f}hs) = a_1^+ a_3^+ a_7^- a_8^- a_9^-$  and  $b\bar{f}h\bar{s}$  with  $\sigma(b\bar{f}h\bar{s}) = a_1^- a_3^- a_7^+ a_8^+ a_9^+$  cannot be compared using  $\prec_\sigma$ . Therefore, it may happen that we cannot infer all conditionals in the knowledge base



**Fig. 4** Structural preferences in the Tweety example

(thus violating the property of being able to validate all conditionals in the knowledge base, known as *Direct Inference* [22]). To illustrate this we use the simple Tweety example with a knowledge base  $\mathcal{R}^{Tweety} = \{(f|b), (\bar{f}|p), (b|p)\}$  formalising the rules that birds usually can fly, penguins usually cannot fly and penguins usually are birds (with structural preferences given in Fig. 4). Here for the world  $pbf$  there is no  $\prec_\sigma$ -preferred world  $\omega \models p\bar{f}$  and thus we have  $p \not\vdash_{\mathcal{R}}^\sigma \bar{f}$ , though  $(\bar{f}|p) \in \mathcal{R}^{Tweety}$  [14, 17]. So we obtain that comparing conditional structures is not enough and, in case of incomparable worlds, we need to weigh the severity of falsification of rules. For this, we broaden our semantic scope to the more expressive framework of ordinal conditional functions (OCF).

**3.2 Semi-Quantitative Inference by OCF**

An *ordinal conditional function* (OCF), also known as *ranking function*, is a function that assigns a rank of *disbelief* or *implausibility* to each world, i.e., the higher the rank of a world, the less plausible this world is.

**Definition 1** (*Ranking function (OCF, [30, 31])*) An *Ordinal Conditional Function* (OCF) is a function  $\kappa : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$  s.t. the set  $\{\omega | \kappa(\omega) = 0\}$  is not empty, i.e., there have to be maximally plausible worlds.

The rank of a formula  $\phi \in \mathcal{L}$  is the minimal rank of all worlds that satisfy  $\phi$ ,  $\kappa(\phi) = \min\{\kappa(\omega) | \omega \models \phi\}$ . The rank of a conditional  $(\psi|\phi) \in (\mathcal{L}|\mathcal{L})$  is defined by  $\kappa(\psi|\phi) = \kappa(\phi \wedge \psi) - \kappa(\phi)$ . An OCF *accepts* a conditional  $(\psi|\phi)$  iff its verification is strictly more plausible than its falsification, formally  $\kappa \models (\psi|\phi)$  iff  $\kappa(\phi \wedge \psi) < \kappa(\phi \wedge \neg\psi)$ . An OCF is *admissible* with respect to a knowledge base  $\mathcal{R}$  iff  $\kappa \models (\psi|\phi)$  for all  $(\psi|\phi) \in \mathcal{R}$ . The inference relation of OCFs is constructed by preferential models with a total preference relation  $<_\kappa \subseteq \Omega \times \Omega$  s.t.  $\omega <_\kappa \omega'$  iff  $\kappa(\omega) < \kappa(\omega')$  and so a formula  $\phi$  infers a formula  $\psi$  via  $\kappa$ ,  $\phi \vdash_\kappa \psi$ , iff  $\kappa(\phi \wedge \psi) < \kappa(\phi \wedge \neg\psi)$  [12, 14, 17, 31]. For OCFs, acceptance is closely connected to inference, since we have  $\phi \vdash_\kappa \psi$  iff  $\kappa \models (\psi|\phi)$  which gives us that for an  $\mathcal{R}$ -admissible  $\kappa$ ,  $\vdash_\kappa$  satisfies Direct Inference, in contrast to  $\vdash_{\mathcal{R}}^\sigma$ . Being set up from a preferential model, every  $\kappa$ -inference satisfies System P. It has also been shown that  $\vdash_\kappa$  satisfies *Rational Monotony* (RM), i.e.,  $\vdash_\kappa$  behaves

monotonously with respect to formulas  $\chi$  of which the contrary is not inferable [17, 20], formally

$$\phi \vdash \chi \text{ and } \phi \not\vdash \neg\psi \text{ imply } \phi \wedge \psi \vdash \chi. \tag{RM}$$

OCF reasoning is semi-quantitative in that numbers are used for degrees of implausibility, but employed in a qualitative way for comparisons. We recall two established approaches for generating an OCF inductively.

As defined in Sect. 1 we use only “purely qualitative” conditionals, that is, knowledge bases that encode defeasible rules of the type of “if  $\phi$  then usually  $\psi$ ” without any further, quantitative addendum to inductively generate the semi-quantitative OCF representation. This means that, a priori, each conditional is equally important and the approaches generate a ranked model of this equally important conditionals which, a posteriori, provides a plausibility ranking of the possible worlds. For ranking approaches that also allow for semi-quantitative knowledge bases, i.e., knowledge bases which implement how strongly or firmly a conditional rule is believed, see [9, 16, 31].

**System Z** [27] is an inductive approach to generate the unique minimal admissible ranking function  $\kappa_{\mathcal{R}}^Z$  to a knowledge base  $\mathcal{R}$ . This system partitions the knowledge base algorithmically into subsets ordered by the degree of exceptionality of the rule. The ranking function assigns to each world the rank of the most exceptional falsified rule, and the exceptionality of a world is used as its implausibility rank (confer [27] for a complete description of the algorithm). Table 2 (upper row) gives the System Z ranking function for the car start example. For System Z we obtain an inference relation  $\vdash_{\mathcal{R}}^Z$  as defined for OCFs in general which satisfies System P and (RM) as described above.

Being an OCF, System Z satisfies Direct Inference and therefore resolves the unintuitivity we found for structural inference, but just taking the most exceptional rules into account results in a somewhat flat inference, as we can see in the car start example: Here we have  $b f h s \prec_{\sigma} b f h \bar{s}$  because the latter falsifies  $\delta_1$  and  $\delta_3$  additionally to  $b f h s$ , but since System Z just uses the maximum we have  $\kappa(b f h s) = \kappa(b f h \bar{s}) = 2$ , so we obtain that System Z and structural inference are different in general [17].

Also, System Z does not allow for subclass inheritance of exceptional subclasses, i.e., if a subclass is exceptional with respect to a single property of the superclass, it does not inherit any property of the superclass. This is called the *Drowning Problem* [2, 27]. It can be illustrated by extending the Tweety example with the rule that birds usually have wings, so our knowledge base for this example is  $\{(f|b), (\bar{f}|p), (b|p), (w|b)\}$ . For System Z we get the rankings given in Table 3 and see that we cannot infer that penguins have wings since  $\kappa(pw) = 1 = \kappa(p\bar{w})$  and therefore  $p \not\vdash_{\mathcal{R}}^Z w$ .

Overall, we obtain that System Z allows for different inferences than inference by conditional structures. However, the Drowning Problem still results in unintuitive inferences since the somewhat flat inference bears the risk that some important rules may be neglected when constructing the ranking function  $\kappa_{\mathcal{R}}^Z$ .

**c-representations:** Comparing worlds by their conditional structure only did not lead to an entirely satisfactory inference relation. The same holds for weighing the falsification of conditionals by the most exceptional conditional. With the approach of c-representations [12, 13], we develop structural inference further and assign an individual impact  $\kappa_i^- \in \mathbb{N}_0$  as abstract weight to each conditional in the knowledge base  $\mathcal{R} = \{(\psi_1|\phi_1), \dots, (\psi_n|\phi_n)\}$ . The rank of a world is the combined impact of all falsified conditionals, so a c-representation  $\kappa_{\mathcal{R}}^c$  is an OCF defined by

$$\kappa_{\mathcal{R}}^c(\omega) = \sum_{\omega \models \phi_i \wedge \neg \psi_i} \kappa_i^-, \tag{1}$$

where the impacts  $\kappa_i^- \in \mathbb{N}_0$  are chosen such that  $\kappa_{\mathcal{R}}^c$  is admissible with respect to  $\mathcal{R}$ . This is obtained if the impacts satisfy the following system of in Eqs. [12, 13]:

$$\kappa_i^- > \min_{\substack{\omega \in \Omega \\ \omega \models \phi_i \wedge \psi_i}} \left\{ \sum_{\substack{1 \leq j \leq n, j \neq i \\ \omega \models \phi_j \wedge \neg \psi_j}} \kappa_j^- \right\} - \min_{\substack{\omega \in \Omega \\ \omega \models \phi_i \wedge \neg \psi_i}} \left\{ \sum_{\substack{1 \leq j \leq n, j \neq i \\ \omega \models \phi_j \wedge \neg \psi_j}} \kappa_j^- \right\}.$$

So we generate a c-representation by solving the above system of in equations and by this setting the impacts to values such that they satisfy the system. We illustrate the calculation of a c-representation using the knowledge base from the Tweety example where we refer to  $(f|b)$  as conditional 1,  $(\bar{f}|p)$  as conditional 2, etc. The verification resp. falsification of these conditionals in the possible worlds of this example are given in Table 1. We exemplarily set up the first minimum in the first row of the system. Here, the set of worlds that verify the conditional, i.e.,  $\omega \models bf$ , is the set  $\{pbfw, pbf\bar{w}, \bar{p}bfw, \bar{p}bf\bar{w}\}$ . From these, the world  $pbfw$  falsifies the conditional  $(\bar{f}|p)$ , only, hence (according to the above introduced enumeration) the first sum of the minimum is  $\kappa_2^-$ . The world  $\bar{p}bfw$  falsifies no conditional, therefore the respective sum is empty, that is, 0. Applying these deliberations to all of the four worlds we get the term  $\min\{\kappa_2^-, \kappa_2^- + \kappa_4^-, 0, 0\}$  and since each value  $\kappa_i^-$  is non-negative we obtain directly that the minimal value is 0. Setting up the complete system in this way gives us

$$\begin{aligned} \kappa_1^- > 0 - 0 & \quad \kappa_2^- > \min\{\kappa_1^-, \kappa_3^-\} - 0 \\ \kappa_4^- > 0 - 0 & \quad \kappa_3^- > \min\{\kappa_1^-, \kappa_2^-\} - 0 \end{aligned}$$

We now can set  $\kappa_1^-$  and  $\kappa_4^-$  to any integer greater than 0. For this example, we settle for the minimal possible value,

**Table 1** Interpretation of the conditionals in the Tweety example, where “+” indicates falsification and an empty cell indicates non-applicability

	$pbfw$	$pbf\bar{w}$	$p\bar{b}fw$	$p\bar{b}f\bar{w}$	$p\bar{b}\bar{f}w$	$p\bar{b}\bar{f}\bar{w}$	$\bar{p}bfw$	$\bar{p}bf\bar{w}$	$\bar{p}\bar{b}fw$	$\bar{p}\bar{b}f\bar{w}$	$\bar{p}\bar{b}\bar{f}w$	$\bar{p}\bar{b}\bar{f}\bar{w}$
$(fb)$	+											
$(\bar{f} p)$		-										
$(b p)$			+									
$(w b)$												

hence  $\kappa_1^- = \kappa_2^- = 1$ . In the next step, we plug in these values into the system and obtain

$$\begin{aligned} \kappa_1^- &= 1 & \kappa_2^- &> \min\{1, \kappa_3^-\} \\ \kappa_4^- &= 1 & \kappa_3^- &> \min\{1, \kappa_2^-\}. \end{aligned}$$

Here we see that both  $\kappa_2^-$  and  $\kappa_3^-$  cannot be smaller than 1 and hence the minima for  $\kappa_2^-$  and  $\kappa_3^-$  can be solved to 1 which gives us

$$\begin{aligned} \kappa_1^- &= 1 & \kappa_2^- &> 1 \\ \kappa_4^- &= 1 & \kappa_3^- &> 1. \end{aligned}$$

Again, we use the smallest possible value and set  $\kappa_2^- = \kappa_3^- = 2$  and hence obtain a solution for the system of equations. With these values we set up  $\kappa_{\mathcal{R}}^c(\omega)$  as given in Table 3 (lower row) for the Tweety example according to (1), so, for example,  $\kappa_{\mathcal{R}}^c(pbfw) = \kappa_2^- = 2$  because  $pbfw$  falsifies  $(\bar{f}|p)$  and no other conditional, and  $\kappa_{\mathcal{R}}^c(p\bar{b}fw) = \kappa_2^- + \kappa_3^- = 4$  because  $p\bar{b}fw$  falsifies  $(\bar{f}|p)$  and  $(b|p)$ , only.

Each c-representation  $\kappa_{\mathcal{R}}^c$  gives rise to an inference relation  $\sim_{\mathcal{R}}^c$  which, as ranking inference, satisfies System P and (RM), and additionally Direct inference by construction, but also ensures that every conditional from  $\mathcal{R}$  is taken into account, hence ensuring that no conditional “drowns” in the effect of the others: For the car start example (lower row of Table 2) we have  $\kappa_{\mathcal{R}}^c(bfhs) = 2 < 3 = \kappa_{\mathcal{R}}^c(bfh\bar{s})$ , a preference also found in the conditional structures. The approach also allows for an exceptional subclass to inherit properties of its superclass; in the extended Tweety example (lower row of Table 3) we have  $\kappa_{\mathcal{R}}^c(pw) = 1 < 2 = \kappa_{\mathcal{R}}^c(p\bar{w})$  and hence we can infer that penguins have wings,  $p \sim_{\mathcal{R}}^c w$ .

Inference by c-representations is different in general from inference by System Z, which can be seen formally [17] as well as experimentally [33].

Until now we presented three different approaches for inductive conditional reasoning which all allow for high quality inferences. All approaches have exponential space complexity. Having additional properties (Direct Inference, no Drowning Problem) comes with the price of increasing time complexity, which is exponential in the alphabet and at least polynomial in the size of the knowledge base for each approach. In the following we introduce network approaches which promise to reduce the space complexity by decomposing the ranking function into smaller local components as well as reducing time complexity by dividing the knowledge base into smaller sets.

### 4 Network Based Nonmonotonic Inference

In this section, we recall network approaches for storing OCFs in local components and show how these can be set up using inductive approaches to combine both the

**Table 2** Ranking functions  $\kappa_{\mathcal{R}}^Z$  (upper row) and  $\kappa_{\mathcal{R}}^c$  (lower row) of the car start example

$\omega$	$b f h s$	$b f h \bar{s}$	$b \bar{f} h s$	$b \bar{f} h \bar{s}$	$\bar{b} f h s$	$\bar{b} f h \bar{s}$	$\bar{b} \bar{f} h s$	$\bar{b} \bar{f} h \bar{s}$	$\bar{b} f \bar{h} s$	$\bar{b} f \bar{h} \bar{s}$	$\bar{b} \bar{f} \bar{h} s$	$\bar{b} \bar{f} \bar{h} \bar{s}$	$\bar{b} \bar{f} h s$	$\bar{b} \bar{f} h \bar{s}$	$\bar{b} \bar{f} \bar{h} s$	$\bar{b} \bar{f} \bar{h} \bar{s}$
$\kappa_{\mathcal{R}}^Z(\omega)$	2	2	0	1	2	2	2	1	2	1	2	1	2	1	2	1
$\kappa_{\mathcal{R}}^c(\omega)$	2	3	0	1	4	3	2	1	2	1	2	1	2	1	2	1

inductive strength of the methods and the efficient properties of the network approaches. For this, we make use of ideas from probabilistics, starting with Bayesian networks [26]. Bayesian networks are used to (compactly) store the joint probability distribution of a set of variables under assumptions of conditional independence thus yielding the possibility to save only local probability information and calculate a global distribution by means of decomposition. We recall Bayesian-style networks with local ranking tables instead of local tables of probabilities [7, 10, 15, 16], so called *OCF-networks*.

OCF-networks resemble Bayesian networks in that they are directed acyclic graphs over the alphabet, with each variable  $V \in \Sigma$  being annotated with a conditional ranking table  $\kappa_V(V|pa(V))$  of this variable given its parents  $pa(V)$ . Like Bayesian networks, these local OCF tables build up a global OCF  $\kappa$ , such that [16, Proposition 1]

$$\kappa(\omega) = \sum_{V \in \Sigma} \kappa_V(V(\omega)|pa(V)(\omega)), \tag{2}$$

where  $V(\omega)$  resp.  $pa(V)(\omega)$  indicates the outcome  $\dot{v}$  of  $V$  with  $\omega \models \dot{v}$  resp. the configuration  $\dot{p}$  of the variables in  $pa(V)$  with  $\omega \models \dot{p}$ . Figure 5 is an OCF-network for the car start example with respective local ranking tables and the global ranking function that coincides with the OCF given in Table 2 (lower row). The global OCF of an OCF-network coincides with the local ranking tables on the respective marginals, i.e., we have  $\kappa(V(\omega)|pa(V)(\omega)) = \kappa_V(V(\omega)|pa(V)(\omega))$  for all  $\omega \in \Omega$  and all  $V \in \Sigma$  [16, Theorem 1], which allows us to use the lightweight local representation instead of the global OCF, locally.

In an OCF-network, every vertex is conditionally ( $\kappa$ )-independent from its non-descendants given its parents [16, Theorem 2] where conditional independence with respect to a ranking function  $\kappa$  is defined in full analogy to conditional independence with respect to a probability distribution [16, 26, 31]. This property, called the *local directed Markov Property*, also is an important property of Bayesian networks and allows for local calculations and local inferences in the network.

OCF-networks can be combined with inductive reasoning such that we combine the benefits of the inductive approach with the benefits of the network approach but only if the knowledge base is restricted to contain *single elementary conditionals*, i.e., conditionals whose conclusion is a literal and whose premise is a conjunction of

literals. Based on this subclass of conditional knowledge bases we can set up the network part using the algorithm of [10], where the set of variables defines the set of vertices and the set of edges is defined by the conditionals such that there is an edge from  $V'$  to  $V$  iff  $V'$  appears in the premise of a conditional with conclusion  $v$  or  $\neg v$ . For example, in the car start knowledge base we have the conditional  $(\bar{s}|\bar{b}f)$  which gives us the edges  $B \rightarrow S$  and  $F \rightarrow S$  in Fig. 5. The knowledge base  $\mathcal{R}$  is split up into partitions  $\mathcal{R}_V$  for all  $V \in \Sigma$  such that  $\mathcal{R}_V$  contains all conditionals with conclusion  $v$  or  $\bar{v}$  with the respective alphabet  $\Sigma_V = \{V\} \cup pa(V)$ . For these partitions we generate local ranking functions with an inductive approach. In the car start example we get  $\mathcal{R}_B = \{(\bar{b}|h), (b|\bar{h})\}$  as local knowledge base for  $B$  and each of System Z or c-representations gives us the ranking table  $\kappa_B(B|H)$  in Fig. 5 by calculating  $\kappa_B(\dot{b}|\dot{h}) = \kappa_{\mathcal{R}_B}(\dot{b}\dot{h}) - \kappa_{\mathcal{R}_B}(\dot{h})$  for every outcome  $\dot{b}, \dot{h}$  of  $B$  and  $H$ , respectively.

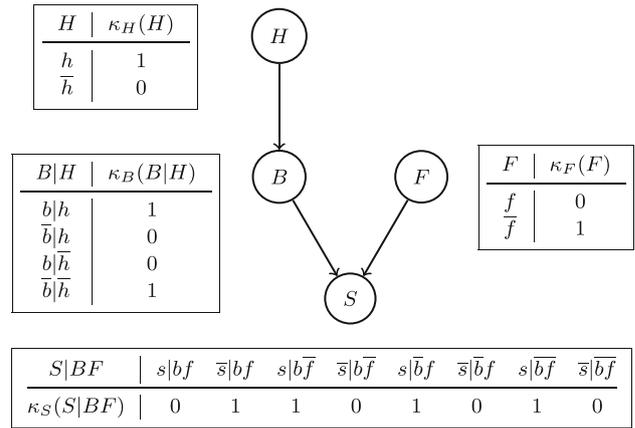
Overall, we see that we can use the approach of OCF-networks in combination with approaches of inductive reasoning (which is hereafter referred to as *inductive OCF-network*) with ranking functions and conditional knowledge bases. This combination allows us to use smaller knowledge bases and smaller alphabets for the inductive approaches, which gives us a significantly lower computational complexity. Also, the approach allows us to store the global ranking functions in the local tables, which reduces the space complexity of the inductive approaches, too. But, despite their formal and computational advantages, OCF-networks have two major drawbacks:

On the technical side, the local information stored in the ranking table is not a ranking function but a table of conditional ranking values. The global OCF coincides with these values, but these values are conditional values only. This means that to obtain the rank of a world or a formula, we have to calculate the ranks of this world or formula with respect to decomposition (2) and can *not* use the values in the table directly. In the car start example,  $\kappa(fs)$  is hence calculated as  $\kappa(fs) = \min\{\kappa_H(h) + \kappa_B(b|h) + \kappa_F(f) + \kappa_S(s|bf), \kappa_H(h) + \kappa_B(\bar{b}|h) + \kappa_F(f) + \kappa_S(s|\bar{b}f), \kappa_H(\bar{h}) + \kappa_B(b|\bar{h}) + \kappa_F(f) + \kappa_S(s|b\bar{f}), \kappa_H(\bar{h}) + \kappa_B(\bar{b}|\bar{h}) + \kappa_F(f) + \kappa_S(s|\bar{b}\bar{f})\} = 0$ .

On the semantical side, the global OCF of an inductive OCF-network may not be admissible with respect to the original knowledge base. For the car start example, e.g., we

**Table 3** System Z and minimal c-representation for the extended Tweety example

$\omega$	$pbfw$	$pb\bar{f}\bar{w}$	$p\bar{b}f\bar{w}$	$p\bar{b}\bar{f}w$	$p\bar{b}\bar{f}\bar{w}$												
$\kappa_{\mathcal{R}}^Z(\omega)$	2	2	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1
$\kappa_{\mathcal{R}}^c(\omega)$	2	3	2	4	4	4	4	2	2	2	2	1	1	1	1	1	1



**Fig. 5** OCF-network of the car start example

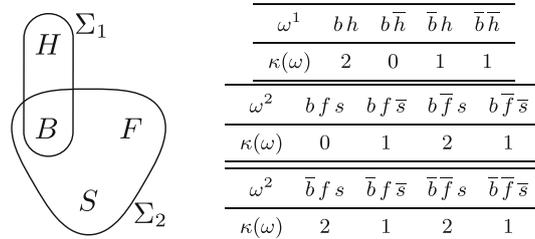
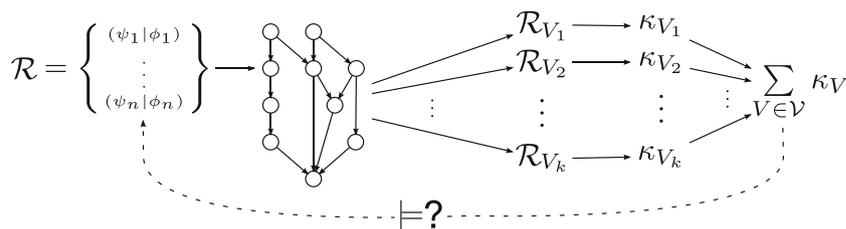
check if the conditional ( $slf$ ) is accepted by the global ranking function by comparing  $\kappa(fs) = 0$  to  $\kappa(f\bar{s}) = 1$  and obtain that, for this example, the global OCF accepts the conditional. As stated in [16], this cannot be guaranteed because there may be a conditional ( $\psi|\phi$ ) in the knowledge base which is accepted *locally* (which is ensured by the local approaches), but the global OCF does not accept this conditional. Formally this means we would have  $\kappa_V \models (\psi|\phi)$ , that is,  $\kappa_V(\phi\psi) < \kappa_V(\phi\bar{\psi})$ , for a  $V \in \mathcal{V}$  but  $\kappa(\phi\psi) \not< \kappa(\phi\bar{\psi})$ , i.e.,  $\kappa \not\models (\psi|\phi)$ . Figure 6 illustrates this problem.

Another approach from probabilities that originates from methods for computing probability distributions according to the principle of maximum entropy, is the approach of local event groups, or LEGs [24, 26]. An LEG is a set of variables with a probability distribution. For using this approach with OCFs, we define OCF-LEGs as sets of variables with an associated OCF. An OCF-LEG network [6] is a hypertree  $(\Sigma, \langle \Sigma_i, \kappa_i \rangle_{i=1}^k)$  of OCF-LEGs  $\langle \Sigma_i, \kappa_i \rangle$  such that  $\Sigma = \bigcup_{i=1}^k \Sigma_i$  with a global ranking function  $\kappa$  that coincides with the local ranking functions  $\kappa_i$  on the respective subsets of variables  $\Sigma_i$ . The *separators*  $\mathbb{S}_i$  are defined as  $\mathbb{S}_i = \Sigma_i \cap \bigcup_{j=1}^{i-1} \Sigma_j$ . This network decomposes the global ranking function such that we have

$$\kappa(\omega) = \sum_{i=1}^m \kappa_i(\Sigma_i(\omega)) - \sum_{i=1}^m \kappa_i(\mathbb{S}_i(\omega)), \tag{3}$$

(“factorisation property”) according to [6, Theorem 1]. In these networks, each OCF-LEG is conditionally independent from its neighbours given the intersection of both sets [6, Proposition 2]. This gives us that inferences, and thus especially the acceptance of conditionals by the global OCF, can be calculated locally in the OCF-LEG instead of consulting the global OCF [6, Corollary 1]. Figure 7 shows an OCF-LEG network for the car start example with local OCFs generated with c-representations.

**Fig. 6** Illustration of the admissibility problem of inductive OCF-networks



**Fig. 7** OCF-LEG network of the car start example

OCF-LEG networks can be set up inductively from a conditional knowledge base without language restrictions [6, 24, 26]. Here, each conditional defines a hyperedge which contains all variables of which literals are contained in either the premise or the conclusion of the conditional. After this step, a covering hypertree of this hypergraph is set up using standard graph methods, and local knowledge bases are formed with respect to the resulting hyperedges. From these local knowledge bases local OCFs are calculated using one of the inductive methods described in Sect. 3.2 and finally, the global OCF of the network is calculated according to the decomposition property (3), see [6] for details.

It is ensured that the global OCF is admissible with respect to each local knowledge base and hence with the original knowledge base. Thus OCF-LEG networks overcome the admissibility problem of OCF-networks, while the network approach still reduces time and space complexity as stated for network approaches in general.

### 5 Conclusion

In this article we pointed out the relevance of conditional logics for modelling human reasoning adequately, as we illustrated with the Wason selection task, and described various approaches from the field of artificial intelligence to process conditional information. We recalled preferential models as qualitative semantics and ordinal conditional functions as (semi-) quantitative semantics for inductively generating models from knowledge bases, the latter being sets of conditionals that encode if-then-sentences. We reported on three established approaches, structural inference, System Z and c-representations, which turned out to

generate three preferential inference relations that are different in general but complementary in various respects.

We combined the discussed inductive OCF-methods with a network approach to obtain a semi-qualitative variant of Bayesian networks, so called OCF-networks. This approach inherits major properties from Bayesian networks (e.g., conditional independence, decomposition) and allows for inductive construction but fails to ensure that the resulting OCF represents the original knowledge base completely. The afterwards reported approach of OCF-LEG networks uses hypergraphs instead of directed acyclic graphs. With OCF-LEG network we were able to overcome the problems that OCF-networks have with partial knowledge bases, this approach allows us to construct undirected graph structures based on knowledge bases with a complete inductive representation of the knowledge bases. This network has a decomposition property by which the global OCF can be constructed by local calculations and also comes with an independence property, which together allow to shift most calculations to small local ranking components. An implementation of the OCF-LEG approach called AIRCONDITIONALS—Automated Inductive Reasoning with Conditionals is publicly accessible<sup>2</sup>.

### References

1. Adams EW (1975) The Logic of Conditionals: An Application of Probability to Deductive Logic. Synthese Library. Springer Science+Business Media, Dordrecht

<sup>2</sup> <http://airconditionals.cs.tu-dortmund.de>.

2. Benferhat S, Cayrol C, Dubois D, Lang J, Prade H (1993) Inconsistency Management and Prioritized Syntax-Based Entailment. In: Proceedings of the International Joint Conferences on Artificial Intelligence. pp. 640–647
3. R Da Silva Neves, Bonnefon JF, Raufaste E (2002) An Empirical Test of Patterns for Nonmonotonic Inference. *Ann Math Artif Intell* 34(1–3):107–130
4. Dietz EA, Hölldobler S, Ragni M (2013) A computational logic approach to the abstract and the social case of the selection task. In: Proceedings of the 11th International Symposium on Logical Formalizations of Commonsense Reasoning
5. Donald N, Cross CB (2002) Conditional logic. In: Gabbay DM, Guenther F (eds) *Handbook of Philosophical Logic*, 2nd edn., vol. 4. Kluwer Academic Publishers, Dordrecht, pp. 1–98
6. Eichhorn C, Kern-Isberner G (2014) LEG Networks for Ranking Functions. In: Logics in Artificial Intelligence. In: Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA'14), Lecture Notes in Computer Science, vol. 8761. pp. 210–223
7. Eichhorn C, Kern-Isberner G (2015) Using inductive reasoning for completing OCF-networks. In: FLAIRS Uncertain Reasoning special issue (in print), *Journal of Applied Logic*
8. Furbach U, Schon C (2015) Deontic Logic for Human Reasoning. In: Eiter T, Strass H, Truszczynski M, Woltran S (eds) *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation—Essays Dedicated to Gerhard Brewka on the Occasion of His 60th Birthday*, no. 9060 in Lecture Notes in Artificial Intelligence. Springer International Publishing, Wiesbaden
9. Goldszmidt M, Pearl J (1991) System-Z+: A formalism for reasoning with variable-strength defaults. In: Proceedings of the ninth National conference on Artificial intelligence (AAAI'91), vol. 1. pp 399–404
10. Goldszmidt M, Pearl J (1996) Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artif Intell* 84(1–2):57–112
11. Halpern JY (2005) *Reasoning About Uncertainty*. MIT Press, Cambridge
12. Kern-Isberner G (2001) Conditionals in Nonmonotonic Reasoning and Belief Revision—Considering Conditionals as Agents. No. 2087 in *Lecture Notes in Computer Science*. Springer, Berlin
13. Kern-Isberner G (2004) A thorough axiomatization of a principle of conditional preservation in belief revision. *Ann Math Artif Intell* 40:127–164
14. Kern-Isberner G, Eichhorn C (2012) A structural base for conditional reasoning. In: *Human Reasoning and Automated Deduction—KI 2012 Workshop Proceedings*. pp. 25–32
15. Kern-Isberner G, Eichhorn C (2013) Intensional combination of rankings for OCF-networks. In: Proceedings of the 26th International FLAIRS Conference. pp. 615–620
16. Kern-Isberner G, Eichhorn C (2013) OCF-networks with missing values. In: Proceedings of the 4th Workshop on Dynamics of Knowledge and Belief. pp. 46–60
17. Kern-Isberner G, Eichhorn C (2014) Structural inference from conditional knowledge bases. In: Unterhuber M, Schurz G (eds) *Studia Logica Special Issue Logic and Probability: Reasoning in Uncertain Environments*, vol 102(4). Springer Science+Business Media, Dordrecht
18. Kraus S, Lehmann DJ, Magidor M (1990) Nonmonotonic reasoning, preferential models and cumulative logics. *Artif Intell* 44(1–2):167–207
19. Kuhnmünch G, Ragni M (2014) Can Formal Non-monotonic Systems Properly Describe Human Reasoning? *Proc Cognitive Sci Conf 2014*:1806–1811
20. Lehmann DJ, Magidor M (1992) What does a conditional knowledge base entail? *Artif Intell* 55(1):1–60
21. Lewis D (1976) Probabilities of conditionals and conditional probabilities. *Philos Rev* 85:297–315
22. Lukasiewicz T (2005) Weak nonmonotonic probabilistic logics. *Artif Intell* 168(1–2):119–161
23. Makinson D (1994) General patterns in nonmonotonic reasoning. In: Gabbay DM, Hogger CJ, Robinson JA (eds) *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol 3. Oxford University Press, New York, pp 35–110
24. Meyer CH (1998) Korrektes Schließen bei unvollständiger Information: Anwendung des Prinzips der maximalen Entropie in einem probabilistischen Expertensystem. 41. Peter Lang Publishing Inc, Pieterlen (**in German**)
25. Oaksford M, Chater N (1995) Information gain explains relevance which explains the selection task. *Cognition* 57:97–108
26. Pearl J (1988) *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann Publishers Inc., San Francisco
27. Pearl J (1990) System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning. In: Proceedings of the 3rd conference on Theoretical aspects of reasoning about knowledge (TARK'90). pp. 121–135
28. Pfeifer N, Kleiter GD (2005) Coherence and nonmonotonicity in human reasoning. *Synthese* 146(1–2):93–109
29. Sperber D, Cara F, Girotto V (1995) Relevance theory explains the selection task. *Cognition* 57:31–95
30. Spohn W (1988) Ordinal conditional functions: A dynamic theory of epistemic states. In: Harper W, Skyrms B (eds) *Causation in Decision, Belief Change and Statistics*, The University of Western Ontario Series in Philosophy of Science, vol 42. Springer Science+Business Media, Dordrecht, pp 105–134
31. Spohn W (2012) *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford University Press, New York
32. Thimm M (2014) Tweety—A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In: Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR'14)
33. Thorn PD, Eichhorn C, Kern-Isberner G, Schurz G (2015) Qualitative Probabilistic Inference with Default Inheritance for Exceptional Subclasses. In: *progic 2015: The Seventh Workshop on Combining Probability and Logic*
34. Wason PC (1968) Reasoning about a rule. *Q J Exp Psychol* 20(3):273–281
35. Wason PC, Johnson-Laird P (1990) *Psychology of Reasoning: structure and Content*. Harvard University Press



**Christian Eichhorn** received his diploma degree in Computer Science from the University of Technology (TU) in Dortmund in 2011. He is currently enrolled as a Ph.D. student at the University of Technology (TU) in Dortmund.



**Gabriele Kern-Isberner** is a professor of computer science at the University of Technology (TU) in Dortmund. In 1985, she received her Ph.D. in mathematics from the University of Dortmund, and in 2000, she did her habilitation in computer science at the FernUniversität in Hagen. She has been working both in qualitative and quantitative methods for knowledge representation and reasoning, and building bridges between these two fields. She is inter-

ested in particular in multidisciplinary work with philosophers and psychologists.