Inductive reasoning, conditionals, and belief revision

Gabriele Kern-Isberner[0000−0001−8689−5391]

Dept. of Computer Science, TU Dortmund, 44221 Dortmund, Germany

Abstract. This paper presents a broad view on inductive reasoning by embedding it in theories of epistemic states, conditionals, and belief revision. More precisely, we consider inductive reasoning as a specific case of belief revision on epistemic states where three-valued conditionals are a basic means for representing beliefs. We present a general framework for inductive reasoning from conditional belief bases that also allows for taking background beliefs into account, and illustrate this by probabilistic reasoning based on optimum entropy.

Keywords: inductive reasoning · conditionals · belief revision · reasoning at optimum entropy

1 Introduction

In its original sense, inductive reasoning means deriving generic knowledge from given examples in a way that completes the example-based information concisely to make it applicable to other situations. In this paper, we take a bit broader view on inductive reasoning: we pursue the idea that inductive reasoning should be able to "generate" new beliefs from given beliefs and ideally, complete the beliefs of a human being as far as possible. This is a very common and basic problem in the area of knowledge representation in artificial intelligence. Here, it is usually assumed that knowledge and beliefs of a human being, or an agent, respectively, can be represented by a knowledge base, i.e., a finite set of formulas in a suitable logic, and that more knowledge and beliefs can be inferred from this base. In artificial intelligence, the distinction between knowledge and beliefs is vague, because its main goal is to model knowledge and behaviour of agents, so knowledge often means subjective knowledge, which is very close to beliefs. We do not want to enter this fundamental discussion here but will use the term beliefs throughout this paper as a synonym for subjective knowledge.

So, inductive reasoning should be able to extend the beliefs of a belief base in a non-trivial, principled way. Of course, the logic framework in which beliefs are represented plays a crucial role here. In the simple case of propositional logic, deduction, or more generally, a Tarski consequence operator would satisfy the general requirements of an inductive reasoning operator, and similarly for first-order predicate logic. Beyond classical logics, non-monotonic logics using so-called default rules, or rules with exceptions, provide more powerful inference operators, prominent approaches here are Reiter's default logic and answer set programming. Both are symbolic and able to infer formulas from belief bases of facts and rules. In quantitative logical settings, probability theory offers a rich semantic framework for nonmonotonic reasoning, and the principle of maximum entropy (*MaxEnt principle*) [7, 13] yields a most powerful inductive inference operator from probabilistic belief bases. There are also popular approaches using qualitative structures like (total) preorders, or semi-quantative methodologies based on Spohn's ordinal conditional functions, also called ranking functions [17], like system Z [6] that allow for reasoning from conditional belief bases.

This paper aims at describing inductive reasoning in a broad context where we elaborate on connections to conditionals and belief change theory, and where we are able to distinguish clearly between background, or generic, beliefs and evidential, or contextual, information, a feature that is listed in [3] as one of three basic requirements a *plausible exception-tolerant inference system* has to meet. We build upon previous works, in particular [9, 11], and elaborate a general vision of inductive reasoning in the context of belief revision. While is has been well known that nonmonotonic reasoning and belief revision are "two sides of the same coin" [5], the focus here is on inductive reasoning as a concept that merges techniques from both areas to bring forth a methodology in which reasoning and revision can interact in various ways to realise inductive reasoning from different background beliefs and under different contextual information. A core concept in this methodology are epistemic states which are equipped with meta-structures supporting reasoning and revision, and beliefs are expressed by conditionals in the first place. Note that, of course, also propositional beliefs are covered in our approach by identifying a conditional $(A|\top)$, where $\top$ is a tautology, with the plausible belief A. Interestingly, total preorders on possible worlds are meta-structures that provide a solid foundation for reasoning, revision, and conditionals, and indeed, they are a basic requirement for AGM revision [8]. So, we build upon AGM revision but go far beyond by addressing iterative revision and conditional revision.

The outline of the paper is as follows: We recall basic definitions and notations in Section 2 and discuss the nature of epistemic states in Section 3, also pointing out their connections to inductive reasoning, conditionals, and belief revision. We also exemplify inductive reasoning and belief revision in probabilistics via the principles of optimum entropy. Finally, we compare inductive reasoning/revision to applying beliefs to specific situations, which we call focusing in Section 4, and conclude in Section 5.

2 Basics and notations

The propositional language $\mathcal{L}$ with formulas A, B is defined in the usual way by virtue of a finite signature Σ with atoms $a, b, \ldots$ and junctors $\wedge, \vee$, and $\neg$ for conjunction, disjunction, and negation, respectively. The $\wedge$-junctor is mostly omitted, so that AB stands for $A \wedge B$, and negation is usually indicated by overlining the corresponding proposition, i.e. $\overline{A}$ means $\neg A$. Literals are positive or negated atoms. The set of all propositional interpretations over Σ is denoted by Ω_Σ. As the signature will be fixed throughout the paper, we will usually omit the subscript and simply write Ω. Possible worlds are understood as a synonym for interpretations, and are usually represented by a complete conjunction of the corresponding literals, i.e., a conjunction mentioning all atoms of the signature such that exactly those atoms are negated that are evaluated to *false*. Also the satisfaction relation $\models$ between worlds and formulas is defined in the usual way: $\omega \models A$ iff ω evaluates A to *true*. In this case, we say ω is a model of A. The set of all models of A is denoted by $Mod(A)$. Then, $A \models B$ for two formulas $A, B \in \mathcal{L}$ if $Mod(A) \subseteq Mod(B)$.

$\mathcal{L}$ is extended to a conditional language $(\mathcal{L} \mid \mathcal{L})$ by introducing a conditional operator $|$: $(\mathcal{L} \mid \mathcal{L}) = \{(B|A) \mid A, B \in \mathcal{L}\}$. $(\mathcal{L} \mid \mathcal{L})$ is a flat conditional language, no nesting of conditionals is allowed. Conditionals $(B|A)$ with *antecedent* (or *premise*) A and *consequent* B are basically considered as three-valued entities in the sense of de Finetti [2] which can be verified ($\omega \models AB$), falsified ($\omega \models A\overline{B}$), or simply not applicable ($\omega \models \overline{A}$) in a possible world ω. So, they have to be interpreted within richer semantic structures such as *epistemic states* like probability distributions, or ranking functions [17]. In this paper, we choose both of these semantic frameworks to exemplify our approach.

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

Ordinal conditional functions (OCFs), (also called *ranking functions*) $\kappa : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ with $\kappa^{-1}(0) \neq \emptyset$, were introduced first by Spohn [17]. They express degrees of plausibility of propositional formulas A by specifying degrees of disbeliefs of their negations $\overline{A}$. More formally, we have $\kappa(A) := \min\{\kappa(\omega) \mid \omega \models A\}$, so that $\kappa(A \vee B) = \min\{\kappa(A), \kappa(B)\}$. A conditional $(B|A)$ is accepted in the epistemic state represented by κ, written as $\kappa \models (B|A)$, iff $\kappa(AB) < \kappa(A\overline{B})$, i.e. iff AB is more plausible than $A\overline{B}$.

In general, let Ψ be any epistemic state, specified by some structure that is found appropriate to express conditional beliefs from a suitable conditional language $(\mathcal{L} \mid \mathcal{L})^*$, in which conditionals may be equipped with quantitative degrees of belief, according to the chosen framework. For instance, for probability functions, $(\mathcal{L} \mid \mathcal{L})^* = (\mathcal{L} \mid \mathcal{L})^{prob} = \{(B|A)[x] \mid A, B \in \mathcal{L}, x \in [0,1]\}$, and in qualitative environments, $(\mathcal{L} \mid \mathcal{L})^* = (\mathcal{L} \mid \mathcal{L})$. Moreover, an entailment relation $\models$ is given between epistemic states and conditionals; basically, $\Psi \models (B|A)^*$ means that $(B|A)^*$ is accepted in Ψ, where acceptance is defined suitably. Let $\mathcal{E}^* = \mathcal{E}^*_\Sigma$ denote the set of all such epistemic states using $(\mathcal{L} \mid \mathcal{L})^*$ for representation of (conditional) beliefs. Moreover, epistemic states are considered as (epistemic) models of sets of conditionals $\Delta \subseteq (\mathcal{L} \mid \mathcal{L})^*$: $Mod^*(\Delta) = \{\Psi \in \mathcal{E}^* \mid \Psi \models \Delta\}$. As usual, $\Delta \subseteq (\mathcal{L} \mid \mathcal{L})^*$ is *consistent* iff $Mod^*(\Delta) \neq \emptyset$, i.e. iff there is an epistemic state which is a model of Δ.

3 Inductive reasoning based on epistemic states and belief revision

In this section, we develop our general approach to inductive reasoning as a special case of epistemic belief revision. Epistemic states serve as a mediator between reasoning and revision by providing both an epistemic background for reasoning and an ideal outcome of induction from and revision by (conditional) belief bases. First, we discuss the semantic structures of epistemic states that are required for this purpose; in particular, we emphasize the crucial role of conditionals in this context. Then we present the technical realisations of our approach on an abstract level. Finally, we elaborate on the different types of belief and information that our approach can handle.

3.1 Epistemic states and conditionals

In this paper, in the context of inductive reasoning and belief revision, we take a pragmatic view on epistemic states. We expect (the representation of) epistemic states to be equipped with some meta-structures which in suitable logical frameworks allow for performing reasoning and belief revision, and to be complete in the sense that answers to all possible queries (in the respective) framework can be generated, to the best of the human's beliefs. Note that we use the term "revision" here in a general sense, as a synonym for integrating new information to one's current beliefs, i.e., as a super-concept also including update [8] or focusing [3]. When the specific change operator called revision in the AGM theory is meant, we speak of "AGM revision", or specify this explicitly.

As a crucial feature to go beyond classical logic towards modelling of human's beliefs, we presuppose that epistemic states can evaluate conditionals to be accepted or not accepted. We avoid saying that a conditional is *true* or not in an epistemic state because, on the one hand, conditionals are not binary but three-valued, and, on the other hand, the understanding of conditionals in commonsense reasoning is not truth-functional at all. To accept a conditional, humans would expect a meaningful connection between antecedent and consequent. This is crucial for our approach to inductive reasoning because this connection can be used for reasoning in a way that captures human-like thinking. The basic idea is simple: A conditional $(B|A)$ is accepted if its verification AB is deemed to be more plausible, or probable, than its falsification $A\overline{B}$. The inherent connection between antecedent and consequent is taken into regard by considering A and B resp. A and $\overline{B}$ jointly when assessing plausibility, or probability. Beyond plain comparison, also degrees of plausibility, or probability, can be assigned to verification and falsification so as to measure the strength of a conditional, if the respective semantic framework allows for that.

The fundamental connection between epistemic states, conditionals, plausibility, (inductive) reasoning, and belief revision on which this paper relies can be roughly expressed by the following equivalences:

$$\Psi \models (B|A) \quad \text{iff} \quad AB \prec_\Psi A\overline{B} \quad \text{iff} \quad A \mid\!\sim_\Psi B \quad \text{iff} \quad \Psi * A \models B, \qquad (1)$$

where Ψ is an epistemic state in $\mathcal{E}^*$, $\preceq_\Psi$ is a suitable relation expressing plausibility (or probability)[1], $\mid\!\sim_\Psi$ is an inference relation based on Ψ, and $*$ is an epistemic (or iterative) revision operator that takes an epistemic state and a proposition and returns again an epistemic state (in the sense of [1]). More generally, we assume that $*$ can also deal with much more complex beliefs given by sets of conditionals Δ such that $\Psi * \Delta \in \mathcal{E}^*$, and we also adopt the success postulate of AGM theory here, i.e., we presuppose that $\Psi * \Delta \models \Delta$. This also includes the case of revising by a (plausible) proposition A via identifying A with $(A|\top)$. Equation (1) reveals that both epistemic states and conditionals are also carriers of strategic information that become effective for reasoning and revision. Our focus here is on the inference relation $\mid\!\sim_\Psi$, and basing it on an epistemic state Ψ helps clarifying formally what is understood by induction. Before we go into more details here, we need to make explicit more clearly what we expect from the meta-structures associated with an epistemic state.

Indeed, a purely qualitative preorder might be a suitable meta-structure that is associated with an epistemic state. Of course, there are more sophisticated representation frameworks, such as possibility theory, ranking functions, and probability functions. But also modal logical frameworks seem to be good candidates for representing epistemic states, or heterogeneous structures consisting of different components (with reasonable interactions between them) might prove useful. This is not necessarily a question of numerical or symbolic representation, both types of frameworks can be fine.

But when it comes to numbers it should be clear that the crucial point here is not that they may provide a richer semantics, but they definitely provide richer structures that calculations for information processing might follow. And this makes them quite distinguished candidates for epistemic states in the context of reasoning and belief change. It is not by accident that probability theory with its two independent arithmetic operators (addition and multiplication) has been playing a major role here. Although AGM might have marked the beginning of symbolic belief revision and of devising rational postulates for belief change, performing practical belief change has been done for a much longer time in the probabilistic framework. Presumably the first belief change operator ever is probabilistic conditioning, and Jeffrey's rule [14] shows a possible way of incorporating even uncertain evidence. So, it is not for the numbers that we should care about probability theory but for the rich arithmetic structure that provides a powerful apparatus to express and process information (cf. also [14]). Via the multiplication operator, (conditional) independencies (and hence monotonic inference behaviour) can be expressed, and its inverse operator, division, allows to easily transform one distribution into another at the occurrence of new information via conditioning. Furthermore, the addition operator takes care of disjunctive propositional information, e.g., to allow for reasoning by cases. Having once adopted such basic techniques, information processing becomes easy.

[1] Note that $A \prec_\Psi B$ iff $A \preceq_\Psi B$ and not $B \preceq_\Psi A$

3.2 Inductive reasoning and belief revision

If we understand inductive reasoning as completing partial beliefs (as specified in a belief base Δ) as best as possible, then its result should be an epistemic state Ψ_Δ:

$$\Psi_\Delta = ind(\Delta), \tag{2}$$

where *ind* is some inductive reasoning mechanism; we also say that Δ is *inductively represented* by Ψ via *ind*, or that Δ *inductively generates* Ψ. For instance, Δ may be a set of conditionals, and *ind* might be specified by system Z [6], or c-representations [10], associating to each consistent set of conditionals a ranking function [17]. Inductive reasoning from Δ is then implemented by reasoning from $\Psi = ind(\Delta)$ via the conditionals being accepted in Ψ. That is, *ind* realises *model-based inductive reasoning*.

But this cannot be the end of the story. The mind of a human being is always evolving and changing by learning, or receiving new information $\mathcal{I}$ in general, where $\mathcal{I}$ can just be a fact, more complex contextual information also including conditionals (e.g., when we enter a new country, different compliance rules apply), or even trigger some deeper learning processes. Starting a new inductive reasoning process each time when we receive new information would make our beliefs incoherent, $\Psi = ind(\Delta)$ and $\Psi' = ind(\mathcal{I})$ might be completely unrelated (except for that they have been built up by the same inductive reasoning formalism). Integrating new information $\mathcal{I}$ into existing beliefs represented by an epistemic state Ψ is exactly the task of (epistemic or iterated) belief revision [1], returning a new epistemic state Ψ' after revising Ψ by $\mathcal{I}$:

$$\Psi_\Delta * \mathcal{I} = ind(\Delta) * \mathcal{I} = \Psi' \tag{3}$$

Note that we use $*$ here in a generic sense as a placeholder for a suitable change operator. Regarding that $\Psi_\Delta = ind(\Delta)$ has been built up inductively from a belief base Δ, and that also $\mathcal{I}$ will also be only partial information on some current context usually, the following questions naturally arise immediately: How do *ind* and $*$ interact? What (maybe completely different) roles do Ψ_Δ, Δ and $\mathcal{I}$ play in this scenario?

We first discuss the second question by analysing different qualities of beliefs with respect to the roles they play in the reasoning process. Roughly, we can distinguish between background, or generic, and evidential, or contextual knowledge, as well as between explicit and implicit beliefs. From background or generic knowledge, the agent takes beliefs which hold in general and of which she can make use of in different situations. For instance, the current beliefs of an agent getting up on a usual Monday morning might be different from those on a usual Sunday, but presumably his generic background has not changed much. The evidential resp. contextual information $\mathcal{I}$ she receives might include that it is Monday and raining, and that due to new construction areas she has to take some detours when going to work. We prefer the attribute "contextual" to "evidential" in the following, since this information may relate not only to a specific situation and can be much more complex than some evidential facts. For instance, the temporal scope of context may be one hour or one week, the scope may refer to a specific house or to a whole country, or it may contain information on abstract contexts, such as holidays or working environments. Assuming that $\Psi_\Delta = ind(\Delta)$ expresses background beliefs, incorporating contextual information cannot be done simply via the "union" of Ψ_Δ and $\mathcal{I}$ (whatever this might be), or by the union of Δ and $\mathcal{I}$ because this would ignore the different natures of background beliefs and contextual information. The agent's new epistemic state should rather arise from the adaptation of Ψ_Δ to contextual information. This is expressed by (3), but only as a base case when we start reasoning from a belief base including our core background beliefs. However, this process must be iterative, i.e., $\Psi = \Psi_\Delta$ may more generally be the result of such a revision $\Psi = \Psi_{prior} * \mathcal{I}_{prior}$, or new information $\mathcal{I}'$ arrives that triggers a new change process $(\Psi_\Delta * \mathcal{I}) * \mathcal{I}'$, so that (3) evolves to the iterative change problem

$$(\Psi_\Delta * \mathcal{I}) * \mathcal{I}' = (ind(\Delta) * \mathcal{I}) * \mathcal{I}'. \tag{4}$$

And here, three essentially different reasoning resp. revision scenarios are possible (note that the $*$-operators are just placeholders to be specified adequately):

- First, the context to which $\mathcal{I}$ refers has evolved, and $\mathcal{I}'$ is information on this new context for which, however, $\mathcal{I}$ is still relevant. This scenario is often referred to as *updating*. Then the two $*$-operators in (4) would be of the same type, and $\Psi_\Delta * \mathcal{I}$ would be changed to $(\Psi_\Delta * \mathcal{I}) * \mathcal{I}'$. A modification of this scenario applies if the contexts to which $\mathcal{I}$ and $\mathcal{I}'$ refer are completely unrelated, but the agent uses the same background beliefs Ψ_Δ for reasoning, then we would end up with $\Psi_\Delta * \mathcal{I}'$.
- Second, $\mathcal{I}'$ refers to the same context as $\mathcal{I}$. In this case, $\mathcal{I}$ and $\mathcal{I}'$ should be considered to be on the same level, and we would obtain $\Psi_\Delta * (\mathcal{I} \cup \mathcal{I}')$. This is a typical case of *belief revision* in the AGM-sense.
- Third, $\mathcal{I}'$ enriches or modifies background beliefs, i.e., it affects the basis from which reasoning with the information $\mathcal{I}$ is performed. This is what happens when *learning*. In the first case, if $\mathcal{I}'$ is fully compatible with Δ, $ind(\Delta \cup \mathcal{I}') * \mathcal{I}$ would be a proper solution. If $\mathcal{I}'$ contradicts (parts of) Δ, then $\Psi_\Delta * \mathcal{I}' = ind(\Delta) * \mathcal{I}'$ would provide suitable background beliefs, and $(ind(\Delta) * \mathcal{I}') * \mathcal{I}$ would be the result of the revision problem.

Therefore, we argue that the distinction between revision and update [8], and also the relation between belief change and learning is not just a technical issue, but has to be made on a conceptual and modelling level. The involved revision operators $*$ might respect such differences, but from the discussion above it becomes clear that also differences can be made by different ways of applying one and the same revision operator $*$ in different scenarios, also involving inductive reasoning. While (3) claims that involving belief revision is necessary for a coherent perspective of inductive reasoning, the third of the cases elaborated above shows how inductive reasoning can affect belief revision: Changing $ind(\Delta)$ to $ind(\Delta \cup \mathcal{I}')$ makes the revision of background beliefs possible. For more formal investigations of the differences between AGM-like revision and update, and for a reconciliation with AGM theory, please see [11].

Elaborating further on this intimate connection between inductive reasoning and belief revision, we might even envisage inductive reasoning involving background beliefs expressed by an epistemic state Ψ_{bk}, i.e., $\Psi = ind_{\Psi_{bk}}(\Delta)$, and then inductive reasoning from Δ might be realised by revision:

$$\Psi = ind_{\Psi_{bk}}(\Delta) = \Psi_{bk} * \Delta. \tag{5}$$

And when no background beliefs are available or relevant, we assume some uniform epistemic state Ψ_u as a starting point:

$$ind = ind_{\Psi_u}. \tag{6}$$

This implements inductive reasoning from epistemic states thoroughly via epistemic belief revision because this approach yields

$$\Psi_\Delta = ind(\Delta) = \Psi_u * \Delta. \tag{7}$$

This means that each epistemic revision operator that is able to handle complex information Δ induces an inductive inference operator. This makes inductive reasoning coherent, as explained above, and allows us to embed inductive reasoning in a richer methodology.

This embedding has two further important advantages: First, revision methodologies may yield immediately mechanisms of inductive reasoning and suitable quality criteria. Second, splitting up inductive reasoning clearly into its inductive mechanism, its involved background beliefs, and context-based beliefs makes formalisms more explicit and more broadly (and flexibly) applicable. However, only very few approaches to epistemic revision with sets of conditionals exist; in Section 3.4, we briefly present the principle of minimum cross-entropy for probabilities as a suitable methodology on the base of which inductive reasoning in the respective semantic frameworks can be realised in a straightforward way.

3.3 Different types of beliefs

Our approach to inductive reasoning via belief revision sketched above also distinguishes between explicit beliefs in a belief base, and implicit beliefs derivable in an epistemic state. The necessity of such a distinction is quite obvious in a belief change scenario, since implicit resp. derived beliefs are more easily changed than explicit beliefs. Having to give up explicit beliefs not only needs more effort, but it is quite a different thing. Formally, if $\Psi_\Delta = ind(\Delta)$, and the new information $\mathcal{I}$ is in conflict with Δ, e.g., $\Delta \cup \mathcal{I}$ is inconsistent, then we are still able to perform revision in the sense of updating via $\Psi_\Delta * \mathcal{I} = ind(\Delta) * \mathcal{I}$, whereas revision as genuine revision in the AGM sense via $ind(\Delta \cup \mathcal{I})$ would not be possible. If the agent comes to know that an explicit belief is (presumably) false, she might react more reluctant to incorporate it, trying perhaps to collect more evidence etc. If finally, she is ready to believe the new information, there are three possibilities: In the first case, the new information $\mathcal{I}$ might contradict the derived beliefs in Ψ_Δ but is nevertheless consistent with Δ, AGM revision $ind(\Delta \cup \mathcal{I})$

would be a suitable option. In the second case, the agent acknowledges that her previous explicit beliefs were erroneous before, in which case she has to perform a proper belief base change by applying merging techniques which are able to resolve conflicts[2]. In the third case, the agent admits that the current context has changed, and she has to adapt her beliefs to these changes, in which case one would find some updating process appropriate. Summarizing, our approach to inductive reasoning is able to deal with (and properly distinguish between) generic, background and contextual beliefs, on the one hand side, and explicit and implicit beliefs, on the other. This is possible by considering inductive reasoning within belief revision frameworks, and provides perfect grounds for a rich methodology that ensures coherence over different reasoning scenarios.

Furthermore, we mention an axiom for iterated revision that is particularly suitable to express coherence in the above sense, but which has been considered only in very few of the current belief revision frameworks and has been introduced under the name *Coherence* in [9] where it plays a crucial role for characterizing the principle of minimum cross entropy, but actually goes back to [16]:

(Coherence) $\Psi * (\Delta_1 \cup \Delta_2) = (\Psi * \Delta_1) * (\Delta_1 \cup \Delta_2)$.

(Coherence) demands that adjusting any intermediate epistemic state $\Psi * \Delta_1$ to the full information $\Delta_1 \cup \Delta_2$ should result in the same epistemic state as adjusting Ψ by $\Delta_1 \cup \Delta_2$ in one step. The rationale behind this axiom is that if the new information drops in in parts, changing any intermediate state of belief by the full information should result unambigously in a final belief state. So, it guarantees the change process to be *logically coherent*.

Note that (Coherence) does not claim that $(\Psi*\Delta_1)*\Delta_2$ and $(\Psi*\Delta_1)*(\Delta_1\cup\Delta_2)$ are the same, just to the contrary – these two revised epistemic states will be expected to differ in general, because the first is not supposed to maintain prior contextual information, Δ_1, whereas the second should do so, according to success. However, (Coherence) can help ensuring independence of parts of the history that serves as background beliefs for inductive reasoning. In the situation described by (5) where we reason inductively from Δ with background beliefs Ψ_{bk}, imagine that we still are aware of the last conditional information Δ_0 that shaped Ψ_{bk}, i.e., $\Psi_{bk} = \Psi_1 * \Delta_0$, which would be mandatory to be able to distinguish among the different scenarios sketched above. But in general, it will be the case that Ψ_{bk} and Δ_0 do not determine Ψ_1 uniquely, so that there may be a different Ψ_2 satifying also $\Psi_{bk} = \Psi_1 * \Delta_0 = \Psi_2 * \Delta_0$. For updating Ψ_{bk}, this is irrelevant because only Ψ_{bk} matters. However, for AGM-like revision, we would like to compute $\Psi_{bk} * \Delta = \Psi_1 * (\Delta_0 \cup \Delta)$, but also $\Psi_2 * (\Delta_0 \cup \Delta)$ would be a suitable candidate. Here (Coherence) guarantees that the resulting epistemic state would be the same:

$$\Psi_1 * (\Delta_0 \cup \Delta) = (\Psi_1 * \Delta_0) * (\Delta_0 \cup \Delta) = (\Psi_2 * \Delta_0) * (\Delta_0 \cup \Delta) = \Psi_2 * (\Delta_0 \cup \Delta).$$

[2] Note that this would also be possible in our general framework, however, we leave this for future work here to not distract from the main focus of this paper

This makes clear that in our conceptual framework of inductive reasoning in the context of belief revision, integrating background beliefs and different pieces of information can be done in different but coherent ways. This means, having to deal with different pieces of information, the crucial question is not whether one information is more recent than others, but which pieces of information should be considered to be on the same level, i.e., belonging to the same type of belief (background vs. contextual), or referring to the same context (which may, but is not restricted to be, of temporal type). Basically, pieces of information on the same level are assumed to be compatible with one another, so simple set union will return a consistent set of formulas (please also see footnote 2). Pieces of information on different levels do not have to be consistent, here latter, or more reliable ones may override those on previous levels.

3.4 Reasoning on optimum entropy and with OCFs

We briefly illustrate the concepts presented in this section by inductive reasoning and revision with probabilities and ranking functions.

The principles of maximum entropy and minimum cross-entropy are powerful methodologies for inductive reasoning and belief revision in probabilistics. Due to lack of space, we cannot recall them fully here but refer in particular to [13, 9, 10]. For a (consistent) set of probabilistic conditionals $\varDelta$, the principle of maximum entropy selects the unique probability distribution $ME(\varDelta)$ with maximum entropy, and if prior information P is given, then the principle of minimum cross-entropy selects (under mild consistency conditions) a unique probability distribution $P *_{ME} \varDelta$ that is a model of $\varDelta$ and has minimal information distance to P, thus realizing probabilistic belief revision. The crucial equation for understanding and analyzing ME-revision is given by

$$P *_{ME} \varDelta(\omega) = \alpha_0 P(\omega) \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \alpha_i^{1-x_i} \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \alpha_i^{-x_i}, \tag{8}$$

with the α_i's being exponentials of the Lagrange multipliers, one for each conditional in $\varDelta$, and have to be chosen properly to ensure that $P *_{ME} \varDelta$ satisfies all conditionals in $\varDelta$ with the associated probabilities. α_0 is simply a normalizing factor. For a complete axiomatization of the principle of minimum cross-entropy within the scope of probabilistic revision by conditional-logical postulates, see [9]. If P_u is a suitable uniform distribution, both ME-principles are related via $ME(\varDelta) = P_u *_{ME} \varDelta$. This means that ME is an inductive reasoning mechanism derived from a belief revision operator in the sense of (7), and $*_{ME}$ realises inductive reasoning from general background beliefs P in the sense of (5). Let us further note that ME-revision also satisfies (Coherence) [16]. Hence the ME-methodology is quite a perfect example to illustrate all concepts and relationships presented in this paper in a probabilistic framework.

Transferring the basic ideas underlying the ME-principles to the framework of ranking functions brings us to *c-revisions* and *c-representations* [10]. Formally, the c-revision methodology provides approaches to revision of ranking functions

κ by consistent sets Δ of conditionals, and inductive reasoning from conditional belief bases (also by taking background beliefs into account) according to (7) and (5) via the following schema:

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

where the parameters κ_i^- have to be chosen suitably to ensure that $\kappa *_c \Delta \models \Delta$, and κ_0 is a normalization factor. A c-representation of Δ is obtained from that by choosing the uniform prior $\kappa_u(\omega) = 0$ for all $\omega \in \Omega$. C-revisions satisfy (Coherence), but only when considered as a family of revisions (for more technical details, please see [12]).

4 Focusing and Conditioning

Focusing means applying generic knowledge to a reference class appropriate to describe the context of interest (cf. [3]). As this reference classe is assumed to be specified by factual information and indicates a shift in context (to that reference class), focusing should be performed by updating the current epistemic state to *factual* information which is certain, i.e. with probability 1. It can easily be shown that for ME-change, updating with such information results in conditioning the prior epistemic state, and indeed, conditioning is usually considered to be the proper operation for focusing.

However, in a probabilistic setting, conditioning has been used for revision, too [4, 3]. So revision and focusing are often supposed to coincide in the framework of Bayesian probabilities though they differ conceptually: revision is not only *applying knowledge*, but means incorporating a new constraint so as to *change knowledge*. Due to this conceptual mismatch, paradoxes have been observed. Gärdenfors investigated *imaging* as another proper probabilistic change operation [4]. Dubois and Prade argued that the assumption of having a uniquely determined probability distribution to represent the available knowledge at best is responsible for that flaw [3]).

However, we will show that in our framework, it is easily possible to treat revision as different from focusing without giving up the assumption of having a single, distinguished epistemic state as a base for inferences. The following proposition reveals the difference between revision by a certain information A, and focusing to A by conditioning; the proofs are straightforward using (8) but tedious.

Proposition 1. *Let P be a distribution, $\Delta \subseteq (\mathcal{L} \mid \mathcal{L})^{prob}$ a (P-consistent[3]) set of probabilistic conditionals, and suppose $A[1]$ to be a certain probabilistic fact.*

(i) Focussing on A, i.e., updating P with $A[1]$ is done by ME-revision and yields $P *_{ME} \{A[1]\} = P(\cdot|A)$; in particular, $(P *_{ME} \Delta) *_{ME} A[1] = (P *_{ME} \Delta)(\cdot|A)$.

[3] Δ is P-consistent if there is a distribution Q with $Q \models \Delta$ and $Q(\omega) = 0$ whenever $P(\omega) = 0$.

(ii) AGM-revising $P *_{ME} \mathit{\Delta}$ *with* $A[1]$ *yields* $P *_{ME} (\mathit{\Delta} \cup \{\mathrm{A}[1]\}) = P(\cdot|\mathrm{A}) *_{ME} \mathit{\Delta}$.

We illustrate that the correct usage of focusing and revision in the probabilistic framework helps resolving well-known paradoxes by considering an example that motivated the application of alternative approaches to uncertain reasoning like Dempster-Shafer theory [15].

Example 1. In a well-known example, Peter, Paul, and Mary are killers one of whom has been hired by Big Boss to commit a murder. Police Inspector Smith knows that Big Boss has first tossed a coin to decide whether it should be a male (Peter or Paul), or a female (Mary), but he does not know about the outcome of the tossing. So, initially, the explicit beliefs of Smith are given by $\mathit{\Delta}_1 = \{(Peter \vee Paul)[0.5], Mary[0.5]\}$, and his initial epistemic state can be calculated via the principle of maximum entropy: $P_1 = ME(\mathit{\Delta}_1)$. It is straightforward to see that $P_1(Mary) = 0.5, P_1(Paul) = P_1(Peter) = 0.25$.

Now Smith comes to know that Peter has been arrested right before the murder, so he could not have committed the crime. This piece of information can be encoded by $R_2 = \{\neg Peter[1]\}$. When incorporating $\mathit{\Delta}_2$ by an update operation (which amounts to a conditioning here), the new epistemic state would be $P_2 = P_1(\cdot|\neg Peter)$, and hence the new beliefs concerning Paul and Mary would be $P_2(Mary) = \frac{2}{3}$, and $P_2(Paul) = \frac{1}{3}$. This seems to be unintuitive, as it gives undue precedence to Mary. However, this flaw is neither an argument against maximum entropy, nor against probability theory in general, but caused by the confusion between focusing and revision. The correct change operation here is revision as simultaneous change, which amounts to computing $P_3 = ME(\mathit{\Delta}_1 \cup \mathit{\Delta}_2)$. Now, in fact, we obtain $P_3(Mary) = P_3(Paul) = 0.5$, as expected.

A statement analogical to Proposition 1 holds for focussing and AGM-revision, in particular c-revision, for OCFs.

5 Conclusion

The aim of this paper is to describe inductive reasoning from conditional belief bases in a rich epistemic framework that takes epistemic states and conditionals as basic encodings of information. Allowing inductive reasoning from background beliefs (in the form of belief bases or epistemic states) leads us naturally to consider also belief revision. More boldly, our main claim here is that inductive reasoning can be considered as a special case of epistemic belief revision. In this way, a coherent and homogeneous approach to inductive reasoning is possible that allows us to realize different forms of inductive reasoning via AGM-like revision, updating, and focusing. We presented a general, abstract framework based on epistemic states and conditionals, and illustrated our ideas both for ordinal and probabilistic environments. We also showed how commonly known paradoxes can be avoided in our framework.

References

1. A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
2. B. de Finetti. La prévision, ses lois logiques et ses sources subjectives. In *Ann. Inst. H. Poincaré*, volume 7. 1937. English translation in *Studies in Subjective Probability*, ed. H. Kyburg and H.E. Smokler, 1964, 93-158. New York: Wiley.
3. D. Dubois and H. Prade. Non-standard theories of uncertainty in plausible reasoning. In G. Brewka, editor, *Principles of Knowledge Representation*. CSLI Publications, 1996.
4. P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, Mass., 1988.
5. P. Gärdenfors. Belief revision and nonmonotonic logic: Two sides of the same coin? In *Proceedings European Conference on Artificial Intelligence, ECAI'92*, pages 768–773. Pitman Publishing, 1992.
6. M. Goldszmidt and J. Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84:57–112, 1996.
7. E. Jaynes. *Papers on Probability, Statistics and Statistical Physics*. D. Reidel Publishing Company, Dordrecht, Holland, 1983.
8. H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings Second International Conference on Principles of Knowledge Representation and Reasoning, KR'91*, pages 387–394, San Mateo, Ca., 1991. Morgan Kaufmann.
9. G. Kern-Isberner. Characterizing the principle of minimum cross-entropy within a conditional-logical framework. *Artificial Intelligence*, 98:169–208, 1998.
10. G. Kern-Isberner. A thorough axiomatization of a principle of conditional preservation in belief revision. *Annals of Mathematics and Artificial Intelligence*, 40(1-2):127–164, 2004.
11. G. Kern-Isberner. Linking iterated belief change operations to nonmonotonic reasoning. In G. Brewka and J. Lang, editors, *Proceedings 11th International Conference on Knowledge Representation and Reasoning, KR'2008*, pages 166–176, Menlo Park, CA, 2008. AAAI Press.
12. G. Kern-Isberner and D. Huvermann. What kind of independence do we need for multiple iterated belief change? *J. Applied Logic*, 22:91–119, 2017.
13. J. Paris. *The uncertain reasoner's companion – A mathematical perspective*. Cambridge University Press, 1994.
14. J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, Ca., 1988.
15. G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ, 1976.
16. J. Shore and R. Johnson. Properties of cross-entropy minimization. *IEEE Transactions on Information Theory*, IT-27:472–482, 1981.
17. W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics, II*, pages 105–134. Kluwer Academic Publishers, 1988.