Inductive Reasoning, Conditionals, and Belief Dynamics

Gabriele Kern-Isberner
Dept. of Computer Science, TU Dortmund, 44221 Dortmund, Germany
gabriele.kern-isberner@cs.tu-dortmund.de

Wolfgang Spohn
Dept. of Philosophy, University of Konstanz, 78457 Konstanz, Germany
wolfgang.spohn@uni-konstanz.de

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 which include conditionals as 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 as well as by ranking-theoretic reasoning based on so-called c-revision. We explain the philosophical perspective behind our approach, and we illustrate its constructive usefulness as well as its integrating power.

1 Introduction

In a familiar sense, inductive reasoning means deriving general 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 any

The idea for this joint paper originated from the First International Conference on Foundations, Applications and Theory of Inductive Logic at the University of Munich from October 12-14, 2022, supported by the Deutsche Forschungsgemeinschaft (DFG) under grant XXX. We thought that our presentations ("Some Strategic Considerations Concerning Inductive Logic", invited talk by Wolfgang Spohn, and "Inductive reasoning, conditionals, and belief revision" by Gabriele Kern-Isberner) there merged nicely.

kind of 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 view on the distinction between knowledge and beliefs is a pragmatic one, because its main goal is to model knowledge and behaviour of agents. So, knowledge often means only subjective knowledge which is more or less the same as belief. Here, we avoid discussing the precise nature of knowledge and belief and use the terms “knowledge” and “belief” interchangeably, just as the term “epistemic state”, the Greek origin of which refers to knowledge, stands for any kind of belief state.[1]

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 [49] and answer set programming [19]. 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* [26, 40] 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 [53, 55], like system Z [21] that allow for reasoning from conditional belief bases.

This paper aims at describing inductive reasoning in a broader context and in a more unified way, elaborating on connections to conditionals and belief change theory distinguishing clearly between background, or generic, beliefs and evidential, or contextual, information, a feature that is listed in [12] as one of three basic requirements a *plausible exception-tolerant inference system* has to meet. We build upon previous works, in particular [29, 33], and elaborate a general vision of inductive reasoning in the context of belief revision. While it has been well known that nonmonotonic reasoning and belief revision are “two sides of the same coin” [18],

[1]Philosophers are used to sharply distinguish knowledge and belief and have intensely discussed this distinction for more than 60 years. They mostly deny that knowledge is merely true belief or even merely justified true belief, see [25].

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 and which represents inductive reasoning from different background beliefs and under different contextual information. A core move in this methodology is to equip epistemic states with meta-structures supporting reasoning and revision, and to use conditionals for expressing beliefs in the first place.

Our approach allows for taking plausible propositional beliefs into account as well, namely by identifying a conditional $(A|\top)$, where $\top$ is a tautology, with the plausible belief A. Note that the statement "A is plausible" is not the same as saying "A is a (certain) fact", both from an epistemological and a knowledge representation point of view. While the latter statement is considered as factual evidence and takes only models of A into account, the first one also considers models of $\neg A$, but as less plausible, reflecting a more generic perspective. We show that our framework can indeed make a difference here.

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 [28]. So, we build upon AGM revision but go far beyond that by addressing iterative revision and conditional revision. Ranking functions implement total preorders by assigning natural numbers to the different layers of a total preorder and thus allow for calculating differences as a measure of plausibility which make it possible to reason in a way that is similar to probabilistic reasoning. As a proof of concept, we illustrate our formal framework in a probabilistic environment by the entropy principles, and in a qualitative/semi-quantitative environment by ranking functions and c-revision.

The outline of the paper is as follows: After recalling basic definitions and notations in Section 2, we discuss, in Section 3, the nature of epistemic states and their dynamics or revision and their fundamental connections to argumentation, inductive reasoning, and conditionals. We explain both, the philosophical perspective as well as how this perspective allows a detailed view of the interactions between inductive reasoning and belief revision. Section 4 then specifies our approach in probabilistic terms via the principles of optimum entropy and in ranking-theoretic terms via the method of so-called c-revision. In section 5, we want to exemplify the integrative power of our approach by comparing inductive reasoning as developed here with the so-called method of focusing that says how to apply beliefs to specific situations. Section 6 provides a brief conclusion.

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 ignature 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 be *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 Boolean combinations or nestings of conditionals are 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 [9] 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 [53]. 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 $P : \Omega \to [0, 1]$ with $\sum_{\omega \in \Omega} P(\omega) = 1$. The probability of a formula $A \in \mathcal{L}$ is given by $P(A) = \sum_{\omega \models A} P(\omega)$. Since $\mathcal{L}$ is finite, Ω is finite, too, and we only need additivity instead of σ-additivity. 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 $P(B|A) = x$ ($x \in [0, 1]$).

Ranking functions, also known as *ordinal conditional functions* (OCFs), $\kappa : \Omega \to \mathbb{N} \cup \{\infty\}$ with $\kappa^{-1}(0) \neq \emptyset$, were first introduced by Spohn [53]. They express degrees of plausibility of propositional formulas A by specifying degrees of disbeliefs of their negations $\overline{A}$. More formally, we have $\kappa(A) := \min\{\kappa(\omega) \mid \omega \models A\}$, so that $\kappa(A \vee B) = \min\{\kappa(A), \kappa(B)\}$. Hence, due to $\kappa^{-1}(0) \neq \emptyset$, at least one of $\kappa(A), \kappa(\overline{A})$ must be 0. Note that expressing absolutely certain beliefs is also possible by assigning the rank ∞ to all worlds falsifying those beliefs. A proposition A is believed if $\kappa(\overline{A}) > 0$ (which implies $\kappa(A) = 0$). Degrees of plausibility can also be assigned to conditionals by setting $\kappa(B|A) = \kappa(AB) - \kappa(A)$. Moreover, ranking functions can also be conditioned by propositions A via $\kappa|A(\omega) = \kappa(\omega) - \kappa(A)$, yielding a ranking function on the models of A. 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 $A\overline{B}$.[2] Ranking functions can be considered as qualitative counterparts of probability distributions. Their plausibility degrees may be taken as logarithmic order-of-magnitude abstractions of probabilities (cf. [20, 21]).

In the following, we take the concept of epistemic states for granted and elaborate on general notations that we use throughout this paper. So, 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, e.g., for ranking functions, $(\mathcal{L} \mid \mathcal{L})^* = (\mathcal{L} \mid \mathcal{L})$. 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 their 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. So, in Subsection 3.1, we first discuss the general relation between arguments and inductive reasoning on the one hand and the dynamics of epistemic states on the other, emphasizing the crucial role of conditionals in this context. In a kind of digression in Subsection 3.2, we more carefully distinguish between inference and arguments and argue that the latter build on relevance relations, which, however, play no further role in this paper. Subsection 3.3 discusses the relation between epistemic dynamics and inductive reasoning in a bit more detail. Subsection 3.4

[2] The full definition here is $\kappa \models (B|A)$ iff $\kappa(AB) < \kappa(A\overline{B})$, or $A \equiv \bot$. For sake of simplicity, we exclude conditionals including contradictions from consideration here.

forms the constructive core of this section. It spells out the various forms the interaction between inductive reasoning and belief revision may take, with particular emphasis on the background beliefs in the form of conditionals which are previously accepted and on the role of a principle called Coherence guiding iterated change. The discussion of epistemic dynamics or belief revision as a framework for inductive reasoning must include the issue where this dynamics may start from. This refers us back to some initial epistemic state as pondered in Subsection 3.5. So, our philosophical tour is topped off by a discussion how we might conceive of such initial epistemic states. It will turn out that we can do so very much in line with our previous discussion of background beliefs.

3.1 Arguments, reasoning, epistemic states and conditionals

Let us start with looking at what happens in arguing with one another. When we give an argument, we start from some hopefully shared premises and infer a conclusion, which is then hopefully shared as well. Or the argument may be only hypothetical, where the epistemic status of the premises is left open. An inference proceeds in the very same way from premises to a conclusion. Or we may say that we reason from the premises to the conclusion. Then we might call the premises the reasons for the conclusion. These are equivalent ways of describing what is going on in an argument.

We should slightly restrict our topic right away. When we talk about arguments, we refer only to descriptive, factual, or empirical reasoning, where premises as well as conclusion are descriptive or truth-evaluable. However, we believe that everything we discuss here applies *mutatis mutandis* to normative, deontic, or evaluative reasoning, where premises and conclusions may be normative sentences the truth-evaluability of which is at least doubtful. We indeed think that there are close parallels, see [58]. However, this is a different large field we do not enter here.

We know well enough what a deductive argument, inference or reason is. There are computational or syntactic and semantic versions. Their mark is (guaranteed) truth preservation. We know how the versions work and we understand their relation. The problem is: most of our inferences and arguments are not deductive or truth-preserving. They are inductive, nonmonotonic, or defeasible. As mentioned, we use "inductive" as a general term for all these kinds of reasoning. When deductive logic dominated formal philosophy, there was the idea that inductive arguments are simply elliptic. They reduce to deductive arguments when implicit premises are made explicit. However, this idea is definitely misguided. We must engage in theories of inductive reasoning on their own, because inductive reasoning involves also defeasible and tentative, even creative processes. It explores the field of rationality beyond deductive logic.

We emphasize that we use “inductive” as a general term for all these kinds of reasoning. In the 19th century, the ’inductive method’ referred to an inference from the particular to the general, the paradigmatic inference being so-called enumerative induction. However, we need all kinds of non-truth-preserving inferences, we need all possible ways to infer what the world is like beyond of what we observe. These ways often include an inductive inference in the traditional sense, an inference from the particular to the general. But we must not presuppose that it is always so. E.g., our forecast of the results of the next election are not based on any putative generalizations of voters’ behavior. So, we are well advised to accept our broad sense of inductive reasoning.

The traditional offers accounting for inductive reasoning are not so rich. There is a long strand of grasping probabilistic inductive reasoning, culminating perhaps in Carnap’s inductive logic [6, 7]. Of course, Bayesian methods are by now well entrenched in all of our scientific canons. Enumerative induction – it indeed looks simplistic– was rather criticized than explicated (but see [56]). The strongest historic attempt at explicating inductive reasoning beyond probabilistic methods are John Stuart Mill’s methods of induction. However, only with the rise of conditional logic, default logic, etc. do we see attempts at grasping inductive reasoning beyond probabilistic methods.

Alas, in the meantime, there is a plethora of diverging, incompatible, or incommensurable accounts of inductive reasoning, forming a large and confusing field that is very difficult to evaluate. Let us try to state some guidelines helping to steer clear in this field.

How might we approach the topic? We take the following observation to be basic: When we give an argument or provide reasons, we try to convince our interlocutors of the conclusion of the argument or of what the reasons entail, not by talking them into the conclusion, but by appealing to their reasoning capacities that make them hopefully infer the same conclusion as we infer. So, the point of giving arguments or providing reasons is to induce rational belief change. This includes the limiting case of stabilizing or confirming, i.e., not changing the epistemic state. In short, reasoning is about the dynamics of belief or about epistemic dynamics in general.

The social dimension, however, is not really essential in our view. Of course, arguing is a social activity, just as language in general. However, for a theory of inductive reasoning, this dimension seems negligible. Sure, one does not argue with oneself, but one is engaged in reasoning or making inferences by oneself all the time. Giving arguments and reasons to others presupposes to have reasons and to have worked them out by oneself, to be convinced by one’s own arguments. And the latter is about individual rational belief change or revision.

There is also a hypothetical variant. When I argue from some assumed premises, I work out what to rationally infer from them, i.e., what to believe given the premises. So, arguments are about conditional belief or, more neutrally, about conditional epistemic states. Indeed, conditional epistemic states and the dynamics of epistemic states are closely related. In simple conditionalization (which can be stated for various epistemic formats) they are even the same; the posterior epistemic state after an input is the same as the prior epistemic state conditional on this input. (It is not always this simple. Still, any other rule of epistemic change we know of is based on the notion of a conditional epistemic state, e.g., conditional probabilities, conditional ranks, or whatever.)

The Ramsey test utilizes this observation for the semantics of the conditional. Indeed, one might say that the semantics of the conditional is the focal point of all our theories of inductive reasoning. All of this establishes a fundamental and close relation between arguments, reasoning and inference on the one hand and epistemic states, their dynamics, and conditionals on the other hand. The direction of the relation – is one relatum more basic than the other? – is still open. We will discuss this.

In the context of inductive reasoning and belief revision we are discussing, we want to take a pragmatic view on epistemic states. We assume the representation of epistemic states to be equipped with some meta-structures allowing to perform reasoning and belief revision in suitable logical frameworks, and we expect them 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, i.e., no further thinking in the sense of exploiting given beliefs and information more deeply would yield a better result. Note that we use the term "revision" here in a general sense, as a synonym for any kind of epistemic dynamics integrating new information to one's current beliefs, i.e., as a super-concept also including update [28] or focusing [12]. When the specific change operator called revision in the AGM theory [1] is meant, we speak of "AGM revision", or specify this explicitly.

The Ramsey test directly utilizes the tight connection between reasoning and belief dynamics for stating a semantics of the conditional. A conditional is accepted in an epistemic state if after acceptance of the antecedent the consequent is accepted as well. In the meantime there are many variations of the Ramsey test. Thereby, we presuppose that epistemic states can evaluate conditionals to be accepted or not accepted. This is a crucial feature of modelling human's beliefs going beyond classical logic. We avoid saying that a conditional is *true* in an epistemic state, because we have above introduced conditionals not as binary but three-valued, and more importantly, because in the commonsense context of reasoning considered here conditionals do not work truth-functionally at all. Rather, 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.

In this paper, we pursue the following well-established variant of the Ramsey test: 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 account 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, thereby measuring the strength of a conditional, if allowed by the respective semantic framework. (In section 2, we have already provided a notation for this measure.)

All in all, it seems appropriate to say that the semantics of the conditional is the focal point of all our theories of inductive reasoning. To resume, our remarks establish a fundamental and close relation between (i) arguments, reasoning and inference, (ii) epistemic states and their dynamics, and (iii) the logic of conditionals. The direction of the relation – is one relatum more basic than the other? – is still open. We will discuss this.

Formally, the upshot of our informal discussion is that in symbolic resp. qualitative frameworks, 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:

$$\Psi \models (B|A) \quad \text{iff} \quad AB \prec_\Psi A\overline{B} \quad \text{iff} \quad A \mathrel{|\!\sim}_\Psi B \quad \text{iff} \quad \Psi * A \models B, \tag{1}$$

where Ψ is an epistemic state in $\mathcal{E}^*$, $\preceq_\Psi$ is a suitable relation expressing plausibility (or probability)[3], $\mathrel{|\!\sim}_\Psi$ 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 [8]). In quantitative frameworks, degrees of beliefs must be suitably added. More generally, we assume that $*$ can also deal with more complex beliefs given by sets of conditionals Δ such that $\Psi * \Delta \in \mathcal{E}^*$. We also adopt the success postulate of AGM theory [1], i.e., we presuppose that $\Psi * \Delta \models \Delta$, meaning that $\Psi * \Delta \models \delta$ for all conditionals δ in Δ. This also includes the case of revision by a (plausible) proposition A via identifying A with $(A|\top)$, as assumed above. Equation (1) reveals that both epistemic states and conditionals

[3]Note that in qualitative semantic environments, e.g., ranking functions, lesser means more plausible, and this also complies with the readings in nonmonotonic preferential inference. Hence we stick to this tradition here. Of course, for probabilities and also, e.g., possibilities, the scales are inverted, so $\preceq_\Psi$ must be interpreted via the numerical $>$-relation. Technically, $A \prec_\Psi B$ iff $A \preceq_\Psi B$ and not $B \preceq_\Psi A$.

are also carriers of strategic information that become effective for reasoning and revision.

3.2 Arguments, inference, and relevance

In arriving at equation (1) we have more or less equated arguments and inference or reasoning. But, as a kind of digression, we might be a bit more careful here. The aim of inference is establishing a conclusion from given premises. If the conclusion is not established with certainty given the premises, as it is in deductive inference, it should at least be more plausible as its opposite. This is what guides equation (1). However, at least intuitively, an argument does more. It provides a *reason for* the conclusion.

Let, e.g., B be the proposition that we will not reach the climate target of keeping global warming below 1,5 degrees. B is highly plausible according to our present epistemic state Ψ, much more plausible than its opposite that we do reach this target. Now let $A =$ "Tom Cruise wins a special Oscar award", which is epistemically entirely irrelevant to B and does not influence the plausibility of B. According to (1), we then have $A \mid\!\sim_\Psi B$. We might still say that B follows from A, because B holds anyway. But it would be odd to say that A is a reason or an argument for B. Or let $A =$ "the US build fifty solar power plants", which is epistemically negatively relevant to B. It diminishes the plausibility of B a little bit, but certainly not far enough to make it less plausible than its opposite. We think that much stronger measures would be needed to reach the climate target. So, according to (1), we still have $A \mid\!\sim_\Psi B$. But now it would be even odder to say that A is an argument or a reason for B. Rather, it is an argument or a reason *against* B, though too weak to undermine B.

Hence, let us define that A is an epistemic *reason for* B iff A is epistemically positively relevant[4] to B iff A raises the plausibility of B iff $(B|A) \prec_\Psi (B|\overline{A})$, where $\preceq_\Psi$ is suitably lifted to conditionals. And the point is that arguments must provide reasons in this sense. That is, an argument is a structure with a premise or premises and a conclusion such that the premise or the conjunction of the premises is positively relevant to the conclusion.[5]. How does this relate to our basic equation (1)?

[4]Relevance is a multiply ambiguous notion. It is clear that in our context epistemic relevance as defined is the only pertinent kind of relevance.

[5]More precisely, this is the structure of a *single* argument. In order to assess a chain of arguments, we would have to study the conditions under which positive relevance spreads along the chain. And in order to study the interaction of arguments, how they defeat, rebut or undermine one another, we would have to study how positive relevance behaves under the augmentation of premises. Here, we do not pursue this study. However, the remark is to express our skepticism that this interaction can be studied in the abstract, as is done in argumentation theory, e.g., according

There are two ways to respond. First, we might discriminate between inference and arguments and still stick to (1) concerning inductive inference, while assigning this explication of an argument to the realm of argumentation theory. This is what we shall do here. Second we might strengthen our notion of inference and our semantics of conditionals by adding the positive epistemic relevance condition. Then we would define $\Psi \models (B|A)$ and $A \mathrel{|\!\sim}_{\Psi} B$ as $AB \prec_{\Psi} A\overline{B}$ *and* $(B|A) \prec_{\Psi} (B|\overline{A})$. Thereby we enter the topic of so-called relevance conditionals, which were recently studied in great detail, see, e.g., [51, 47, 50, 48]. However, we do not pursue here this line of thought.

Is positive epistemic relevance a good explication of the notion of a reason? In any case, it is a subjective explication, entirely dependent on the subject's epistemic state. For the majority of philosophers this is insufficient. They seek to gain a more objective notion of a reason, even of a good reason. For them, rational epistemic dynamics is driven then by those preconceived good reasons. However, they were constructively quite poor in specifying what good reasons are. And the history of inductive skepticism teaches that this might not be easy. With the subjective understanding, we have at least a workable precise explication of that notion suitable for theorizing. We approximate "good reason" here in an abstract way by considering logic-based reasoning methodologies that are equipped with qualitative meta-information allowing for expressing what is good and what is not. And it is still true that epistemic reasons drive epistemic dynamics. Indeed, our definition entitles us to reversely say that epistemic reasons are whatever drives rational epistemic dynamics.

3.3 Inductive reasoning and epistemic dynamics

So much about a possible amendment of the basic equivalence (1) by positive relevance considerations. In the sequel, however, let us just develop that equivalence. It still leaves open how precisely to understand the relation between arguments and inductive reasoning on the one hand and epistemic dynamics and belief revision on the other. In particular, it raises an issue of primacy: Are we first to spell out inductive logic and rational reasoning? And are we thereby to ground an account of rational epistemic dynamics? Or is it the other way around?

We are skeptical of substantially implementing the first direction. It has an objectivistic flair: there is a correct inductive logic, and our epistemic states have to follow it. However, in the tradition of inductive skepticism this objectivism is discredited. The chief witness is perhaps the decline of Carnap's program of inductive

to [14].

logic, which became weaker and weaker till it became almost indistinguishable from de Finetti's subjectivism; see [6, 7].[6]

We therefore favor the second direction: a theory of rational doxastic dynamics should be provided first, from which then an account of inductive reasoning should be derived. This is our first important specification of (1). This entails that any account of inductive reasoning must be based on a specific conception of epistemic states and their rational dynamics. This is made explicit here by using the symbol $\mid\!\sim_{\Psi}$ for the inference relation. Not all accounts pay heed to this maxim. Still, a variety of potentially suitable accounts remain.

A further observation is that the equivalences in (1) presuppose that epistemic states must be conceived as coming in grades that are (partially, weakly, or strictly) ordered. That is, an epistemic state Ψ must provide a plausibility ordering $\preceq_{\Psi}$ in the sense that there are faithful assignments (similar to [28, 8]) that associate suitable meta-structures with an epistemic state. This may exclude further candidate accounts (e.g., the representation of an epistemic state plainly as a set of beliefs or as a propositional knowledge base). Hence, the derivation of an account of inductive reasoning from a conception of epistemic states entails some substantial constraints.

Let us be a bit more specific concerning what we expect from the meta-structures associated with an epistemic state. 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 just their potential for a richer semantics. Rather, they definitely provide richer structures that computations for information processing might utilize. 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, both full group operations) have played a major role here. Although AGM might have marked the beginning of symbolic belief revision and of devising rational postulates for belief change, actually performing belief change has been done for a much longer time within the probabilistic framework. The first belief change operator ever is

[6]However, we should at least point to the efforts of Williamson [60]. See also our discussion in section 3.5.

probabilistic conditioning, and Jeffrey's rule [41] shows a possible way of incorporating even uncertain evidence. So, it is not because of the numbers that we should value probability theory, but because of the rich arithmetic structure that provides a powerful apparatus to express and process information (cf. also [41]). 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 in a way that takes the probabilities of all cases into account. Having once adopted such basic techniques, information processing becomes easy. However, ranking functions show similarly good properties, here we have the (group operation) addition instead of multiplication, and the minimum of ranks instead of addition of probabilities. The minimum is weaker than the addition, it is not a group operation and does not allow for exploiting numerical relationships in a way that addition does. For instance, consider two atoms A, B and the propositions $A \wedge B, A \wedge \neg B, \neg A \wedge B$. In probability theory, if we know $P(A) = P(B)$, we can conclude $P(A \wedge \neg B) = P(\neg A \wedge B)$ because $P(A \wedge B) + P(A \wedge \neg B) = P(A) = P(B) = P(A \wedge B) + P(\neg A \wedge B)$. But from $\min\{\kappa(A \wedge B), \kappa(A \wedge \neg B)\} = \kappa(A) = \kappa(B) = \min\{\kappa(A \wedge B), \kappa(\neg A \wedge B)\}$, we cannot conclude $\kappa(A \wedge \neg B) = \kappa(\neg A \wedge B)$. Technically, this has significant effects on reasoning and revision. On the other hand, evaluating minima is computationally and cognitively less demanding, which might be seen even as an advantage of ranking functions.

We have to clarify the relation between inductive reasoning and the dynamics of epistemic states and thus to specify (1) still further. There is a distinction regarding this dynamics that is often neglected, but seems important to us. On the one hand, there is an internal dynamics which takes place without any external stimulus or input. It consists in thinking, reasoning, calculating, working out consequences, etc. All this in some sense amounts to a temporary or only hypothetical change of an epistemic state. *Inference rules* then tell us how the internal dynamics should proceed. On the other hand, there is an external dynamics which is driven by some external input, information, evidence, experience, not merely in the sense that such external input somehow stimulates the internal dynamics – of course, it does –, but in the sense that the input demands a change of the prior into the posterior epistemic state, however and however incompletely this change is computationally realized. *Rules of doxastic change* then tell us what the posterior state should be depending on the prior state and the input. Note that the internal dynamics belongs to the external statics. Thinking, etc. does not count as doxastic change in the external sense. According to the internal dynamics, an epistemic state records the current state of computation. According to the external dynamics, an epistemic state is an ideal entity which sets a computational goal and may or may not be fully reached by the computations in internal dynamics. Typically, logics, in their many variants, deal with the internal dynamics, by specifying calculi, inference rules, etc., on the syntactic level. By contrast, Bayesianism, belief revision theory, etc. are about the external dynamics. Their dynamic rules specify the relation between prior state, input, and posterior state on a semantic level. Their primary point is not to give computational advice, even if this can often be easily derived.

The connection between the two dynamics is this: The internal dynamics works towards reaching the goal set by the external dynamics, i.e., the posterior state necessitated by the input. Again, the connection may be construed in two opposite ways. Either the input initiates an internal dynamics, the completion of which results in some posterior state, which is then the one the external dynamics aims at. Given the internal dynamics, we can say what completion means (roughly, that any query can be answered in a most informed way so that further thinking or computation does not result in further internal change). Or the input necessitates an external change which then governs the internal dynamics (as being one the completion of which leads to the necessitated result).

We think that the second construal is the one to be preferred. This is our second important specification of (1). For, how could the inference rules be justified within the first construal? By being consistent? By intuition? By some model theory unrelated to epistemic dynamics? No, the justification lies in fixing the goal of computation by specifying a rational external dynamics. The internal dynamics then serves this goal; it is only a means to this end.[7]

We admit that the distinction is often subtle. Conditionalization rules directly tell how to compute the posterior state from the prior state and the evidence to condition on. Or: What is the difference between Rational Monotony (an axiom of conditional logic and nonmonotonic inference) and the postulate K*8 (also called subexpansion and crucial in AGM belief revision theory)? Via the Ramsey Test, they are directly intertranslatable. Still, they have different places in the overall picture of doxastic dynamics.

Let us summarize our two important claims so far: When we want to get a hold on inductive logic, we must start from an account of the rational dynamics of

[7]The work of Pollock [43] is characteristic for this opposition. [42] specifies argument types, and [43] then states rules for the interaction of arguments like the weakest link principle or the no-accrual-of-reasons principle. All of this belongs to the internal dynamics in our sense. From this, an account of belief revision, of the external dynamics, is inferred precisely by running this mechanism of argument types and rules of interaction on the new input till it comes to rest (if it does); see [44]. For a detailed criticism of the entire procedure in the direction indicated see [54].

epistemic states, which in turn presupposes a graded notion of a conditional doxastic state. Indeed, we must start from an account of the rational external dynamics of epistemic states, which sets the goal for the internal dynamics and thus for inference, reasoning, and argumentation.

3.4 The interaction of inductive reasoning and belief revision

If we understand inductive reasoning as completing partial (conditional) 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 [21], or c-representations [32], associating to each consistent set of conditionals a ranking function [53]. 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 possibly 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 [8], 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. What can we say about this change operator $*$, i.e., about the rational dynamics of epistemic states?

The most natural and the most wide-spread picture is that this dynamics is a kind of Markov process: The prior doxastic state and the (total) evidence (in between) determine the posterior doxastic state – according to rules of doxastic change that count as rationally justified. This is Markovian in the general sense that the prior state is supposed to encode a history of changes and hence this history can influence the current change process only through that prior state.

There is a very rich discussion about the rules governing rational epistemic change. In probability and ranking theory, there are rules of conditionalization, simple, generalized Jeffrey, and auto-epistemic conditionalization. There are reflection principles governing doxastic change in both theories. There is minimization of relative entropy, which has an analogue in ranking theory. And so on. We will see that our approach based on equations (1), (2), and (3) leads to formal constraints on doxastic change and its interaction with inductive reasoning that narrows down the range of suitable epistemic frameworks.

Let us add just three general remarks: First, there are many proposals for modelling epistemic states; we have mentioned a few of them in the course of this paper. In probability theory and also in ranking theory the discussion about rules of epistemic change is most elaborate. It would be desirable that it is carefully worked out also within other models, since stating a dynamics is imperative for any representation of epistemic states.

Second, doxastic states are not only about 'eternal' or context-independent propositions, which are usually taken as the only objects of epistemic states, but also, indeed essentially, about indexical or context-dependent propositions, which use, e,g„ "I", "now", and "here", the reference of which can only be determined in context. How do rules of doxastic change apply to them, and how do they interact with rules for 'eternal' propositions? These questions seem to receive only local attention. See, e.g., [57] and [15]. In our approach, indexical information can be part of the contextual beliefs, while 'eternal' information in the sense of generic beliefs would be part of the background beliefs.

Third, all the rules we mentioned concern learning or improving one's epistemic state. This is perhaps the only case that is relevant for the sciences. Still, it is a restriction. There are other kinds of epistemic change, and a theory of rationality should attend to them, too. In particular, we are thinking here of forgetting. As such, forgetting befalls us, there are not rational and irrational ways of forgetting. But there are rational ways of responding to forgetting. Not anything goes after having forgotten something; see, e.g., [57]. Some conceptual considerations and technical results about the role that ranking functions can play in the context of forgetting can be found, e.g., in [34, 35, 3]. However, in the present context which is about the basic connection between epistemic dynamics, induction, and conditionals, this is just a side remark.

So much about the change or revision operator $*$ by itself. Let us return to this basic connection. Given that $\Psi_\Delta = ind(\Delta)$ has been built up inductively from a belief base Δ, and that $\mathcal{I}$ will usually be only partial information about some current context, the following questions arise immediately: How do *ind* and $*$ interact? Which (maybe completely different) roles do Ψ_Δ, Δ and $\mathcal{I}$ play in this scenario?

We first focus on 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.

Let us now look more closely at the first question, the interaction between *ind* and $*$. 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 that we will call *conservative revision* because more prior information (i.e., $\mathcal{I}$) is preserved. Note that conservative revision generalizes Jeffrey's rule [27] to the case where several observations are processed at the same time, without presupposing that the observations are exclusive.

- 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 in *learning*. In the first case, if $\mathcal{I}'$ is fully compatible with Δ, $\mathit{ind}(\Delta \cup \mathcal{I}') * \mathcal{I}$ would be a proper solution. If $\mathcal{I}'$ contradicts (parts of) Δ, then $\Psi_\Delta * \mathcal{I}' = \mathit{ind}(\Delta) * \mathcal{I}'$ would provide suitable background beliefs, and $(\mathit{ind}(\Delta) * \mathcal{I}') * \mathcal{I}$ would be the result of the revision problem.

Therefore, we argue that the distinction between revision and update [28], 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 one might also discriminate 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 $\mathit{ind}(\Delta)$ to $\mathit{ind}(\Delta \cup \mathcal{I}')$ makes the revision of background beliefs possible. For more formal investigations of the differences between conservative revision and update, and for a reconciliation with AGM theory, see [33].

So, starting from an induction perspective, we developed scenarios which are similar to the ones considered in [10] for belief revision: Belief Revision as Defeasible Inference (BRDI), considering a specific case at hand, can be realized as conditioning in our framework, or more generally, via an update where the context has changed. Belief Revision as Prioritized Merging (BRPM), which collects several pieces of uncertain evidence about a case, is realized via conservative revision; please note that it is also possible to apply merging operators instead of simple set union if one wishes to do so. And finally, Revision of Background Knowledge by Generic Information (RBKGI), where the background knowledge is modified by new pieces of (generic) information often in the form of conditionals is also dealt with extensively in Section 3.4.

However, and in contrast to that paper, a main point of our approach is that these "scenarios of belief revision" are not unrelated, but can be realized coherently and naturally (with many interactions) in a rich framework of epistemic revision (see Section 3.4). Our approach is not about technical artefacts that coincidentally bring forth useful results but is grounded on philosophical considerations that clearly show that (inductive) nonmonotonic reasoning and belief revision are not just "two sides of the same coin", but that inductive reasoning is an integral part of epistemic revision in a conditional framework where principles of inductive reasoning follow more general principles of revising epistemic states by conditional beliefs.

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 might assume some uniform epistemic state Ψ_u as a starting point (but see also the discussion on prior and initial states in Section 3.5 below):

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

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 perfectly coherent with the revision operator and allows us to embed inductive reasoning in a richer methodology.

This embedding has two further important advantages: First, revision methodologies may immediately yield 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 4, we briefly present the principle of minimum cross-entropy for probabilities and, a bit more extensively, the c-revisions for ranking functions as suitable methodologies on the base of which inductive reasoning in the respective semantic frameworks can be realised in a straightforward way.

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 $\Psi_\Delta = \mathit{ind}(\Delta)$, 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 $\Psi_\Delta * \mathcal{I} = \mathit{ind}(\Delta) * \mathcal{I}$, whereas conservative revision via $\mathit{ind}(\Delta \cup \mathcal{I})$ 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, conservative revision $\mathit{ind}(\Delta \cup \mathcal{I})$ 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, i.e., we would have $\Psi_\Delta * (\mathcal{I} \circ \mathcal{I}')$ with a merging operator $\circ$. This would give rise to a variant of conservative revision which is neither truly conservative nor prioritized, we leave this for future work. 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 made possible by considering inductive reasoning within belief revision frameworks, and provides perfect grounds for a richer 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 was considered only in very few of the current belief revision frameworks and introduced under the name *Coherence* in [29], where it plays a crucial role for characterizing the principle of minimum cross entropy, but actually goes back to [52]

(Coherence) $\Psi * (\Delta_1 \cup \Delta_2) = (\Psi * \Delta_1) * (\Delta_1 \cup \Delta_2)$.[8]

[8]Coherence of revision corresponds to path independence of contraction, which was introduced by [24]. Both postulates deal with the iterated revision resp. contraction by sets of propositions where one set is a subset of the other. They express how the overall result can be computed from

(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 *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. On the contrary, these two revised epistemic states will usually 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 (or prior) 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 conservative 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:

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

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 see also our remarks on merging on p. 20). Pieces of information on different levels need not be consistent; here later or more reliable ones may override those on previous levels.

intermediate results, i.e., from intermediate points on the revision resp. contraction path.

3.5 Prior, initial and a priori epistemic states

The above picture of doxastic dynamics does not only appeal to rules of change, it also generates a regress to earlier doxastic states. Even allowing background and context to do their work only reiterates the regress. We reason forward from a certain background and in a certain context. But background and context generate an epistemic state from a still more prior state. So, where does this regress lead to? After all, the epistemic states do not come from an infinite past. Let us say that the regress comes to an end at an *initial* epistemic state from where the dynamics starts. But this is only a label. The question is whether we can characterize the initial state in some reasonable way.

The initial state is crucially important, because it and the course of experience fixes all further doxastic states, at least when the learning rules are deterministic. All our rational learning strategies are already encapsulated in this initial state. So, what can we say about it? This is an old and intriguing philosophical issue. In an absolute sense, 17th century philosophy spoke of innate ideas. This was the kind of preformation of our mind discussed between empiricism and rationalism in those times. The talk of innate ideas can certainly not be taken literally. They did not refer to the newborn baby's doxastic state. The discussion advanced with Kant. He may be taken to suggest that the initial state in an absolute sense consists of a priori knowledge. For him, apriority was an epistemological, not a genealogical category. In today's terminology, we may call the initial state conceived in this absolute way the all-embracive *ur-prior*. Alternatively, we may have a low-key relative understanding of the initial state. Then it is just posited to be initial for a given application at hand, where we are at the beginning of an invesigation, and not intended as an all-purpose ur-prior. Let us call this an application-relative conception of initiality. It is definitely closer to current practice, but perhaps not the foundational response we are looking for.

There is a philosophical debate whether the initial state is rationally unique or whether various initial states may be rationally permitted.[9] The opposition is not designed for an application-relative understanding of the initial state. Prima facie, any kind of constraints on the initial state may be imposed depending on the application at hand, and so the issue of Uniqueness does not really arise. Philosophers rather discuss the issue regarding the absolute ur-prior. No doubt, Uniqueness may look attractive. Sometimes, the debate between Uniqueness and Permissiveness is taken to be a symmetric one. Each side has to advance arguments for its claim. In our view, however, the burden of proof is only on the defenders of Uniqueness. It's not that the defenders of Permissiveness have to positively show that various

[9]See, e.g, [23] and [37].

ur-priors are equally rational. They only see no reason to be presently convinced of Uniqueness. Given the fate of Carnap's inductive logic [6, 7], the constructive outlook of Uniqueness is indeed dim. Carnap started as a defender of Uniqueness. However, he immediately recognized that his so-called Wittgenstein function was a total failure. Then he came up with his so-called λ-continuum of inductive methods, and in [6, 7], he ended up with some symmetry principles, which were still extremely permissive.

Note also that Uniqueness would entail an alternative picture of the doxastic dynamics. Any doxastic state is then only a function of the total evidence since the initial time. We only need to take stock of the accumulating evidence. References to any intermediate doxastic states are no longer required. This picture is no longer Markovian. This is how objective Bayesianism as propounded by Williamson [60, 38] conceives of inductive logic. He certainly pursues only a modest application-relative understanding of the initial state, but he thoroughly applies maximum-entropy reasoning both to the initial state as well as to how the total evidence changes the initial state. (The evidence need not be propositional, but can provide any kind of constraints on the posterior state.) Objective Bayesianism is certainly the most constructive attempt to establish Uniqueness. However, let's not further discuss its prospects.

Let us rather look a bit more closely at the initial state by ourselves. We mentioned its relation to the a priori, if taken in the absolute sense. However, so far we referred only to the Kantian a priori. In Kantian terms it is characterized by absolute necessity and generality. We better do not engage into Kant exegesis. A better and indeed fitting characterization is in the present terms of belief dynamics: A priori propositions are just those believed under any circumstances, *whatever the evidence*. Therefore, it is apt to call them unrevisably a priori. Certainly, every initial state must respect this kind of apriority. However, it cannot fully characterize initiality, since it is inductively barren. It cannot tell anything about inductive inference and thus misses what we are after here.

There are strong suggestions in the literature that there are also weaker relative or contextual notions of apriority (see, e.g., [45, 16]). One idea is to relativize apriority to the concepts we have. E.g., Kant may have been right about the apriority of Euclidean geometry, but only as long as there was no other conception of physical space.[10] This kind of apriority may be aptly called defeasible. Such are the beliefs or, in general, the features of initial doxastic states *before any evidence*, which may change afterwards. Apriority in this sense does not entail truth, such beliefs may turn out to be false. The historically first clear example for defeasible apriority is the

[10] This is the suggestion of Putnam (1962, pp. 372f.)

so-called principle of ignorance dictating symmetric or uniform probabilities, which are defeasible, sensitive to experience. Of course, we know by now how elusive the principle of ignorance is. If uniformity, as assumed in (6), is a feature of the initial epistemic state, it may need heavy qualification.

So far, though, defeasible apriority is just another label for the initial epistemic state. Let us at least give some hints how we might say a bit more about it. Above we said that a background may contain, or a context may provide, a lot of conditional information, which is then used for inductive inference. We have indicated how this may technically work in Section 3.4, where we also showed that the property of (Coherence) allows for reducing the impact of the concrete form of the initial state on future revisions significantly. We suggest that the same is true for the initial state. The idea is this:

A subject's doxastic state, however modelled, operates on a given algebra of propositions which are generated from a given conceptual field. This field need not consist of all concepts the subject possesses. It may be a small field just grasping the application at hand. But the subject must master those concepts; she cannot have doxastic attitudes towards propositions she does not understand. Then we might conceive of an initial doxastic state about (an algebra of propositions generated by) a given conceptual field as consisting just of what is required to master this field, but without any further information or evidence concerning those propositions. This would explicate the phrase that the initial state (concerning this field) is one the subject is in before acquiring any (relevant) evidence.

Of course, the explication can't mean that the subject has had no experience whatsoever in such an initial state. She needs a lot of experience in order to acquire any concept at all. However, it is hard to say how much information exactly she must have gathered in order to count as possessing a certain concept. Therefore, the explication is bound to be vague. Still, we think that the notion of a stereotype introduced by Putnam [46] is useful here. One must have learned the relevant stereotype in order to be said to possess a concept. One masters the concept of a dog only if one believes that some (ostensive) paradigms are dogs, that dogs have a certain variable size and shape, that they bark, that they have four legs and a tail, and so on. All this is not unrevisably a priori, it may turn out false in specific cases. But it comes along with the concept of a dog and may thus be called defeasibly a priori.

Such stereotypes are ubiquitous. Our prime examples are dispositional concepts. Reduction sentences, e.g., "an object, when put in water, is soluble if and only if it dissolves", are stereotypes. A disposition typically shows its manifestation. But it may be present, while the manifestation fails, and the other way around.

Thus, reduction sentences are defeasibly a priori conditionals[11]. This is now not just an application-relative a priori, but more strictly a concept-relative a priori. And it promises to answer our quest for a characterization of initiality. Formally, however, it works like any presupposed background or contextual information, be it in conditional or unconditional form. This is how the present point connects up with the explanations in the previous sections.

4 Proofs of concept: Reasoning on optimum entropy and with ranking functions

In a purely qualitative setting, epistemic states can be represented by systems of spheres [39], or simply by a preorder on $\mathcal{L}$ (which is mostly induced by a preorder on worlds). However, for this type of epistemic states, no methodologies are available to date which can handle the complex scenarios of inductive reasoning and belief revision that we sketched in Section 3. Therefore, we choose probabilities and ranking functions as illustrations of our general concept. We briefly describe two well-known revision methodologies in these two frameworks which induce approaches to inductive reasoning from conditional belief bases that have also attracted much attention: The probabilistic principle of minimum cross-entropy (with the principle of maximum entropy as the method for inductive reasoning), and c-revisions of ranking functions (with c-representations allowing for inductive reasoning). Since there is already a vast literature on the entropy principles while ranking functions and c-revisions are less well-known, we focus on the latter approach here. Note that c-revisions have been introduced in a more general form in [30, 32], for the sake of ease of notation, we only use a simplified version of c-revisions here which is nevertheless able to capture all aspects of our approach.

Both revision operators are linked by the property that they both satisfy the *principle of conditional preservation*, as specified e.g. in [30, 32]. This principle makes use of the arithmetic structures underlying probabilities and rankings and allows a very accurate and precise handling of conditional information under belief change. We do not go into technical details here, but the structural similarity between the operators is obvious (see equations (8) and (9) below, keeping in mind that ranks can be understood as logarithmic order-of-magnitude abstractions of probabilities, hence exponents become factors, and products turn into sums when going from probabilities to rankings).

[11] The point is elaborated in [55], sect. 13.3.

4.1 Reasoning on optimum entropy

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 rehearse them fully here but refer in particular to [40, 29, 32]. For a (consistent) set of probabilistic conditionals Δ, the principle of maximum entropy selects the unique probability distribution $ME(\Delta)$ 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 $P *_{ME} \Delta$ that is a model of Δ and has minimal information distance to P, thus realizing probabilistic belief revision. We refer to both principles as the *ME-principles.* The crucial equation for understanding and analyzing ME-revision is given by

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

with the α_i's being exponentials of the Lagrange multipliers, one for each conditional in Δ, and have to be chosen properly to ensure that $P *_{ME} \Delta$ satisfies all conditionals in Δ 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 [29]. If P_u is a suitable uniform distribution, both ME-principles are related via $ME(\Delta) = P_u *_{ME} \Delta$. 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) [52]. Hence the ME-methodology is quite a perfect example to illustrate all concepts and relationships presented in this paper in a probabilistic framework.

4.2 Ordinal c-revision

Transferring the basic ideas underlying the ME-principles to the framework of ranking functions brings us to *c-revisions* and *c-representations* [32].

A(n *ordinal*) *c-revision operator* $*_c$ returns for each ranking function κ and each consistent set Δ of conditionals a ranking function $\kappa *_c \Delta$ that satisfies the principle of conditional preservation, as specified in [30, 32].

Again, by applying the c-revision approach to the uniform prior, i.e., the ranking function κ_u with $\kappa_u(\omega) = 0$ for all $\omega \in \Omega$, we obtain quite easily very well-behaved inductive inference operations on default (or conditional) bases called *c-representations.*

Formally, the c-revision methodology[12] provides approaches to revision of ranking functions by sets of conditionals, and inductive reasoning from conditional belief bases (also by taking background beliefs into account) according to (7) and (5) via the following schemata:

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

such that

$$\kappa_i^- > \min_{\omega \models A_i B_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-) - \min_{\omega \models A_i \overline{B}_i} (\kappa(\omega) + \sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-) \tag{10}$$

for revision and inductive reasoning with background beliefs, and

$$\kappa_\Delta(\omega) = \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B_i}}} \kappa_i^- \tag{11}$$

such that

$$\kappa_i^- > \min_{\omega \models A_i B_i} (\sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-) - \min_{\omega \models A_i \overline{B}_i} (\sum_{\substack{j \neq i \\ \omega \models A_j \overline{B}_j}} \kappa_j^-) \tag{12}$$

for inductive reasoning $\mathit{ind}(\Delta) = \kappa_\Delta$. Because the principle of conditional preservation constitutes the main building principle for both the ME-principles and the c-revision methodology, c-representations resp. c-revisions can be considered as suitable translations of the ME-principles to the framework of ranking functions (see also [20, 5]). Note that c-revisions also satisfy Coherence, when considering the whole family of c-revisions of a specific revision problem; for technical details, please see [36]. The paper [36] also elaborates on relations to other properties of iterated revision, in particular to the Darwiche-Pearl postulates [8].

For representing an epistemic state or revising it, any c-representation resp. c-revision of a conditional belief base may be chosen, all of them share the same good properties, see, e.g., [32, 2]. For practical applications, c-representations resp. c-revisions with pareto-minimal parameters κ_i^- are usually chosen, which however are not uniquely determined in general. C-representations generalize the approach of system Z^* [20] which is defined only for so-called minimal-core knowledge bases, without relying on a probabilistic foundation. For minimal-core knowledge bases, minimal c-representations are unique and coincide with system Z^*.

[12] We consider a simplified version here which is sufficient for the purposes of this paper.

This framework of c-revisions resp. inductive c-representations usually takes plausible beliefs in the form of conditionals (which also cover plausible facts by identifying $(A|\top)$ with a plausible fact A) but can also deal with certain information by assigning ∞ to all falsifying worlds. We illustrate this for the case of c-revising κ with a certain fact which we denote by A^∞. According to (9), we obtain

$$\kappa *_c A^\infty(\omega) = \kappa_0 + \kappa(\omega) + \begin{cases} 0 & \text{if } \omega \models A \\ \infty & \text{if } \omega \not\models A \end{cases}$$

with $\kappa_0 = -\min\{\min_{\omega \models A} \kappa(\omega), \infty\} = -\min\{\kappa(A), \infty\} = -\kappa(A)$, hence

$$\kappa *_c A^\infty(\omega) = \begin{cases} \kappa|A(\omega) & \text{if } \omega \models A \\ \infty & \text{if } \omega \not\models A. \end{cases} \tag{13}$$

Therefore, $\kappa *_c A^\infty$ can be considered as an extension of Spohn's conditioning of ranking functions [53]. If a certain fact A^∞ is part of a belief base, this can be handled similarly by the general approach via (9) and (10) by assigning ∞ to all models of $\overline{A}$. Note that this might influence the minima in (10) and lead to a different revision result, please see also Section 5.

The following example illustrates inductive reasoning and all three change scenarios from Section 3.4 – update, conservative revision, and learning – via the c-change methodology in the framework of ranking functions and qualitative conditionals.

Example 1. *Suppose we have the propositional atoms f -* flying*, b -* birds*, p -* penguins*, w -* winged *animals, k -* kiwis*, d -* doves*. Let the set Δ consist of the following conditionals:*

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

Moreover, we assume strict knowledge, i.e., absolute certainty of the fact that penguins, kiwis, and doves are pairwise exclusive*, which amounts to considering only those worlds as possible that make at most one of $\{p, k, d\}$ true; all other worlds have infinite rank.*

As initial epistemic state, we represent inductively Δ via a c-representation (11) obtaining $\kappa_\Delta = \kappa_u *_c \Delta$ as current epistemic state (cf. Figure 1). The calculation of the parameters κ_i^- according to (12) is straightforward: For each r_i with $i \in$

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

Figure 1: Epistemic state κ_Δ as result of inductive reasoning from Δ in Example 1

$\{1, 4, 5, 6\}$, *and for each of the two minima occurring in (12), respectively, we can choose worlds that do not falsify any (other) conditional from* Δ*; e.g., for* r_1*, choose* $\overline{p}\overline{k}dbfw$ *for the first minimum over the models of* bf*, and* $\overline{p}\overline{k}\,\overline{d}b\overline{f}w$ *for the second minimum over the models of* $b\overline{f}$*. So for each of these* κ_i^-*, both minima are evaluated to* 0*, and we have* $\kappa_i^- > 0$*. We choose all parameters minimally, so we obtain*

$$\kappa_i^- = 1 \textit{ for } i \in \{1, 4, 5, 6\}.$$

The calculation of κ_2^-, κ_3^- *is a bit more complicated. First note that due to* p, k, d *being exclusive, only the models in the leftmost column of Figure 1, i.e., the penguin worlds, are relevant for this calculation. For* κ_2^-*, we compute*

$$\begin{array}{rcl} \kappa_2^- & > & \min\{\kappa_3^-, \kappa_3^- + \kappa_4^-, \kappa_1^-, \kappa_1^- + \kappa_4^-\} - \min\{\kappa_3^-, 0\} \\ & = & \min\{\kappa_3^-, \kappa_1^-\} - 0, \end{array}$$

and because we set $\kappa_1^- = 1$*, we have* $\kappa_2^- > \min\{\kappa_3^-, 1\}$*. Similarly, for* κ_3^- *we obtain* $\kappa_3^- > \min\{\kappa_2^-, 1\}$*. From both inequalities, we can conclude that each of* κ_2^-, κ_3^- *must be at least* 1*, and choosing them minimally yields*

$$\kappa_2^- = \kappa_3^- = 2.$$

Using these parameters defines κ_Δ *according to (11).*

It can be checked easily that κ_Δ *yields the conditional beliefs that penguin-birds do not fly (*$\kappa_\Delta \models (\overline{f}|pb)$ *because of* $\kappa_\Delta(pb\overline{f}) = 1 < 2 = \kappa_\Delta(pbf)$*), and that also penguins are expected to have wings (*$\kappa_\Delta \models (w|p)$ *because of* $\kappa_\Delta(pw) = 1 < 2 = \kappa_\Delta(p\overline{w})$*). So, c-representations do not suffer from the so-called* drowning problem*, particularly a problem of system Z [21]. Moreover, also both kiwis and doves inherit the property* *of having wings from their superclass* birds*, due to* $\kappa_\Delta(kw) = 0 < 1 = \kappa_\Delta(k\overline{w})$ *and* $\kappa_\Delta(dw) = 0 < 1 = \kappa_\Delta(d\overline{w})$.

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

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

The updated epistemic state $\kappa^* = \kappa_\Delta *_c \{(\overline{w}|d)\}$ *is a c-revision of* κ_Δ *by* $\{(\overline{w}|d)\}$ *and can be obtained from* κ_Δ *via (11) by setting* $\kappa^*(\omega) = \kappa_\Delta(\omega) + 2$ *for any* ω *with* $\omega \models dw$ *and setting* $\kappa^*(\omega) = \kappa_\Delta(\omega)$ *otherwise.*

While the conservatively revised state $\kappa_{\Delta'}$*, by construction, still represents the six conditionals that have been known before (and, of course, the new conditional), it can be verified easily that the updated state* κ^* *only represents the five conditionals* $(f|b)$*,* $(b|p)$*,* $(\overline{f}|p)$*, and* $(w|b)$*,* $(b|k)$*, but it no longer satisfies* $(b|d)$ *because* $\kappa^*(bd) = \kappa^*(\overline{b}d) = 1$ *- since* birds *and* wings *have been plausibly related by the conditional* $(w|b)$*, the property of not having wings casts (reasonably) doubt on doves being birds. Moreover, the agent is now also uncertain about the ability of doves to fly, as also* $\kappa^*(fd) = \kappa^*(\overline{f}d) = 1$*. This illustrates that explicitly stated prior beliefs are kept under conservative revision, but might be given up under update. It can be easily checked that these effects would have been the same when c-revising* $\kappa_{\Delta'}$ *instead of* κ_Δ*, since the agent's beliefs on kiwis and doves do not interfere. Moreover, note that if the agent became aware of having missed to represent the conditional belief* $(\overline{w}|k)$ *in the new world after the occurrence of the mysterious illness, still* $\kappa_{\Delta'} *_c (w|b)$ *would be the most adequate result, because here background beliefs are affected, the agent has* learned $(\overline{w}|k)$ *by conservatively revising* κ_Δ.

5 Focusing and Conditioning

Focusing means applying generic knowledge to a reference class appropriate to describe the context of interest (cf. [12]). As this reference class 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. It can easily be shown that both for ME-change and for ordinal c-change, updating with such information results in conditioning the prior epistemic state (see Propositions 2 and 3 below), and indeed, conditioning is usually considered to be the proper operation for focusing. We share this view in this paper, i.e., in our framework, focusing is done via conditioning resp. updating with certain facts.

However, conditioning has been used for revision, too [17, 12]. So revision and focusing are often supposed to coincide 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 [17]. 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, and they recommend to use upper and lower probabilities to permit a proper distinction (cf. [12]).

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 result of revision and a base for inferences. Making use of ME-revision for probabilities, and c-revisions for ranking functions, respectively, it is indeed possible to realize this conceptual difference appropriately. To make this clear, we have to consider belief changes induced by some certain information A, that is, we learn proposition A with certainty. For probabilities, this means that we assign probability 1 to A, while for ranking functions, we assign rank ∞ to $\overline{A}$, see (13) (which implies particularly that A has rank 0). The following two propositions reveal the difference between revision by a certain information A, as realized according to (8) resp. (9) and (10), and focusing to A by conditioning; the proofs are straightforward but tedious, using the mentioned equations.

Proposition 2. *Let P be a distribution, $\Delta \subseteq (\mathcal{L} \mid \mathcal{L})^{prob}$ a (P-consistent[13]) set of probabilistic conditionals, and suppose $A[1]$ to be a certain probabilistic fact.*

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

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

$(P *_{ME} \Delta)(\cdot|A)$.

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

An analogical statement holds for focussing and conservative revision for ranking functions.

Proposition 3. *Let* κ *be a ranking function,* $\Delta \subseteq (\mathcal{L} \mid \mathcal{L})$ *a (*κ*-consistent*[14]*) set of conditionals, and suppose* A *to be a certain fact.*

(i) Focussing κ *on* A*, i.e., updating* κ *with the certain fact* A *via c-revision is done by conditioning and yields* $\kappa *_c A^\infty(\omega) = \kappa|A(\omega)$ *for models* ω *of* A*; in particular,* $(\kappa *_c \Delta) *_c A^\infty = (\kappa *_c \Delta)|A$ *on the models of* A*.*

(ii) Conservatively revising $\kappa *_c \Delta$ *with the certain fact* A *yields* $\kappa *_c (\Delta \cup \{\mathrm{A}^\infty\}) = (\kappa *_c \mathrm{A}^\infty) *_c \Delta$ *(which coincides with* $(\kappa|A) *_c \Delta$ *on the models of* A*) if the same parameters* κ_i^- *are chosen for both c-revisions.*

We present a ranking function adaptation of a probabilistic example from [33].

Example 4. *A psychologist has been working with addicted people for a couple of years. His experiences concerning the propositions*

a : *addicted to alcohol*
d : *addicted to drugs*
y : *being young*

can be summarized by the ranking function κ *as given in (14), serving as the initial epistemic state of the psychologist here.*

ω	$\kappa(\omega)$	$\kappa_1^*(\omega)$	$\kappa *_c y^\infty(\omega)$	$\kappa_2^*(\omega)$	$\kappa_3^*(\omega)$
ady	4	4	3	2	2
$ad\overline{y}$	4	4	∞	∞	∞
$a\overline{d}y$	3	3	2	1	1
$a\overline{d}\overline{y}$	0	0	∞	∞	∞
$\overline{a}dy$	1	6	0	3	4
$\overline{a}d\overline{y}$	4	9	∞	∞	∞
$\overline{a}\overline{d}y$	2	2	1	0	0
$\overline{a}\overline{d}\overline{y}$	3	3	∞	∞	∞

(14)

[14] Δ is κ-consistent if there is a ranking function κ' with $\kappa' \models \Delta$ and $\kappa'(\omega) = \infty$ whenever $\kappa(\omega) = \infty$.

The following conditionals can be entailed from κ:

$$(\overline{d}|a), (\overline{a}|d), (\overline{a}|y), (a|\overline{y}), (d|y), (\overline{d}|\overline{y}).$$

These conditionals express that when focussing on drugs and/or alcohol, usually, people are not addicted to both, and that young people are usually addicted to drugs but not to alcohol, while for older people, it is the other way round.

Now the psychologist is going to change his job: He will be working in a clinic where addictions to both alcohol and drugs are not uncommon, more precisely, people being addicted to drugs tend to also being addicted to alcohol. So, when starting to work in the new environment, the psychologist c-revises his initial epistemic state κ by $\Delta_1 = \{(a|d)\}$, yielding $\kappa_1^* = \kappa *_c \Delta_1$ (with minimal parameter).

After having spent a couple of days in the new clinic, the psychologist realized that this clinic is for young people only, i.e., he has overlooked the certain fact y that only young people are present in his new working context. He conservatively revises κ_1^* by y^∞, yielding $\kappa_2^* = \kappa *_c \Delta_2$ with $\Delta_2 = \{(a|d), y^\infty\}$. Hence, according to Proposition 3, he obtains $\kappa_2^* = (\kappa *_c y^\infty) *_c \Delta_1$.

Note that κ_2^* is different from the ranking function that the psychologist would have obtained by focusing *his beliefs represented by κ_1^* on a young person; in that case, he would have updated κ_1^* by y^∞, yielding $\kappa_3^* = \kappa_1^* *_c y^\infty$ (which coincides with $\kappa_1^*|y$ on the models of y). Clearly, κ_2^* and κ_3^* are different (though the differences are only small).*

Propositions 2 and 3, as well as Example 4 show that, in a (generalized) framework of inductive reasoning including belief revision, a proper distinction between focusing and revision is not only possible, but even mandatory. This difference is akin to the one between "conditioning" and "constraining" elaborated by Voorbraak [59] for classes of probability functions (for a criticism of conditioning *sets* of probability measures, cf. [22]). It is interesting to note that this difference between focusing and (genuine) revision that is possible in our framework also allows for making a basic difference between revising by factual evidence vs. revising by more generic pieces of information. Example 4 nicely illustrates this difference when incorporating the information "all people are young" (κ_2^*) vs. "a specific person is young" (κ_3^*).

However, a proper distinction between focusing and (general) revision is hard to make in most frameworks. For instance, in probabilistics, conditioning is often perceived as the main operation for adjusting to new information (revision) in Bayesian approaches, and hence coincides with focusing, which may lead to unintuitive results. We illustrate this by discussing an example from [13] below. In that paper, a Baysian analysis was performed with conditioning as the major (and only) probabilistic change operation, and it was argued that modelling ignorance via uniform distributions can lead to counter-intuitive results. From that example, we extract the main points for modelling it in the frameworks considered here to highlight the erroneous effects that conditioning may have when applied too plainly.

Example 5 (adapted from [13])**.** *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 man (Peter or Paul), or a lady (Mary), but he does not know about the outcome of the tossing. So, initially, the explicit beliefs of Smith are given by* $\Delta_1 = \{(Peter \vee Paul)[0.5], Mary[0.5]\}$*, and his initial epistemic state can be calculated via the principle of maximum entropy:* $P_1 = ME(\Delta_1)$*. 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* Δ_2 *by conditioning (which corresponds to the usual Bayesian update), 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 (admittedly) unintuitive result is neither an argument against uniform priors, nor against maximum entropy or probability theory in general, but caused by the confusion between focusing and revision. Incorporating Δ_2 *by conditioning would be seen as focusing in our framework, which seems unappropriate because we do not focus on (the reference class of) Peter not being the culprit, but should understand* Δ_2 *as an additional piece of information on the same level as* Δ_1 *because both refer to exactly the same context, namely, the murder, and deliberating on possible delinquents. This is exactly what conservative revision does. So, the correct change operation here would be conservative revision instead of conditioning, which results in computing* $P_3 = ME(\Delta_1 \cup \Delta_2)$*. Now, in fact, we obtain* $P_3(Mary) = P_3(Paul) = 0.5$*, as expected.*

Therefore, our approach shows that the problem addressed in the paper [13] is not with uniform priors but with reducing probabilistic belief revision to Bayesian conditioning, and that the problem can be solved within probabilistic reasoning if the context of information (not necessarily only temporal meta-information) is properly taken into a account and if a richer epistemic framework of revision is used where inductive reasoning is an integral part of.

6 Conclusion

The central claim of this paper was that inductive reasoning can be considered as a special case of epistemic belief revision. We should study the latter in order to deliver an account of the former. This allowed us to present a general, abstract framework based on epistemic states and conditionals and to show how a coherent and homogeneous approach to inductive reasoning is possible realizing different forms of inductive reasoning via conservative revision, updating, and focusing, where all change operations are realized via the same revision operator, but applied in different ways. In particular, we could describe inductive reasoning from conditional belief bases in a rich epistemic framework that takes epistemic states and conditionals as basic encodings of information. We could thus capture how to inductively reason from background beliefs in the form of belief bases or epistemic states. We illustrated our ideas both for ordinal and probabilistic environments and finally showed how commonly known paradoxes can be avoided in our framework.

We presented two semantical frameworks that allow for implementing these ideas as a proof of concept: the principles of optimum entropy in probabilistics, and c-representations/c-revisions for ranking functions. For future work, it would be interesting to see what other semantical frameworks can be used resp. extended to realize the cornerstones of our framework as described in Section 3, in particular in Section 3.4. Possibility theory [11] seems to be a most promising candidate here because it is similar to ranking functions, at least in its product-based form [4]. First steps towards elaborating this have been taken in [31, 30] but more needs to be done to fill out the complete framework. Moreover, we mentioned Gärdenfors' *imaging* [17] as another probabilistic change operation that has interesting applications. However, it is still not clear how imaging can be integrated in our approach, this is also part of our ongoing work.

References

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

[2] C. Beierle, C. Eichhorn, G. Kern-Isberner, and S. Kutsch. Properties and interrelationships of skeptical, weakly skeptical, and credulous inference induced by classes of minimal models. *Artif. Intell.*, 297:103489, 2021.

[3] C. Beierle, G. Kern-Isberner, K. Sauerwald, T. Bock, and M. Ragni. Towards a general framework for kinds of forgetting in common-sense belief management. *KI – Zeitschrift für Künstliche Intelligenz: Special Issue on Intentional Forgetting*, 33(1):57–68, 2019.

[4] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proceedings 3th International Conference on Principles of Knowledge Representation and Reasoning KR'92*, pages 673–684, 1992.

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

[6] R. Carnap. A basic system of inductive logic. In R. Carnap and R. C. Jeffrey, editors, *Studies in Inductive Logic and Probability*, volume I, pages 33–165. University of California Press, Berkeley, 1971/1980.

[7] R. Carnap. A basic system of inductive logic. In R. C. Jeffrey, editor, *Studies in Inductive Logic and Probability*, volume II, pages 7–155. University of California Press, Berkeley, 1971/1980.

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

[9] 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.

[10] D. Dubois. Three scenarios for the revision of epistemic states. *Journal of Logic and Computation*, 18(5):721–738, 2008.

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

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

[13] D. Dubois, H. Prade, and P. Smets. Representing partial ignorance. *IEEE Trans. Syst. Man Cybern. Part A*, 26(3):361–377, 1996.

[14] P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[15] A. Egan and M. G. Titelbaum. Self-locating beliefs. In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy.* Winter 2022 edition, 2022.

[16] H. Field. The a prioricity of logic. *Proceedings of the Aristotelian Society*, 96(1):359–379, 1996.

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

[18] 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.

[19] M. Gelfond and N. Leone. Logic programming and knowledge representation – the A-prolog perspective. *Artificial Intelligence*, 138:3–38, 2002.

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

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

[22] A. Grove and J. Halpern. Updating sets of probabilities. In *Proceedings Fourteenth Conference on Uncertainty in AI*, pages 173–182, 1998.

[23] B. Hedden. *Reasons Without Persons: Rationality, Identity, and Time.* Oxford University Press, Oxford, 2015.

[24] M. Hild and W. Spohn. The measurement of ranks and the laws of iterated contraction. *Artificial Intelligence*, 172(10):1195–1218, 2008.

[25] J. J. Ichikawa and M. Steup. The analysis of knowledge. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Summer 2018 edition, 2018.

[26] E. Jaynes. *Papers on Probability, Statistics and Statistical Physics.* D. Reidel Publishing Company, Dordrecht, Holland, 1983.

[27] R. Jeffrey. *The logic of decision.* University of Chicago Press, Chicago, IL, 1983.

[28] 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.

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

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

[31] G. Kern-Isberner. Representing and learning conditional information in possibility theory. In *Proceedings 7th Fuzzy Days, Dortmund, Germany*, pages 194–217. Springer LNCS 2206, 2001.

[32] 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.

[33] 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.

[34] G. Kern-Isberner, T. Bock, C. Beierle, and K. Sauerwald. Axiomatic evaluation of epistemic forgetting operators. In R. Bartak and K. Brawner, editors, *Proceedings of the 32nd International FLAIRS Conference, FLAIRS-32*, pages 470–475, Palo Alto, CA, 2019. AAAI Press.

[35] G. Kern-Isberner, T. Bock, K. Sauerwald, and C. Beierle. Belief change properties of forgetting operations over ranking functions. In A. Nayak and A. Sharma, editors, *Proceedings of the 16th Pacific Rim International Conference on Artificial Intelligence PRICAI 2019*, number 11670 in Lecture Notes in Artificial Intelligence, pages 459–472. Springer, 2019.

[36] 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.

[37] M. Kopec and M. G. Titelbaum. The uniqueness thesis. *Philosophy Compass*, 11(4):189–200, 2016.

[38] J. Landes, S. Rafiee Rad, and J. Williamson. Determining maximal entropy functions for objective bayesian inductive logic. *Journal of Philosophical Logic*, 52(2):555–608, 2023.

[39] D. Lewis. *Counterfactuals.* Harvard University Press, Cambridge, Mass., 1973.

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

[41] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, San Mateo, Ca., 1988.

[42] J. L. Pollock. *Nomic Probability and the Foundations of Induction.* Oxford University Press, Oxford, 1990.

[43] J. L. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person.* MIT Press, Cambridge, MA, 1995.

[44] J. L. Pollock and A. S. Gillies. Belief revision and epistemology. *Synthese*, 122:69–92, 2000.

[45] H. Putnam. The analytic and the synthetic. In H. Feigl and G. Maxwell, editors, *Scientific Explanation, Space, and Time*, volume 3, pages 358–397. University of Minnesota Press, Minneapolis, 1962.

[46] H. Putnam. The meaning of 'meaning'. In K. Gunderson, editor, *Language, Mind, and Knowledge*, volume 7, pages 131–193. University of Minnesota Press, Minneapolis, 1975.

[47] E. Raidl. Definable conditionals. *Topoi*, 40(1):87–105, 2021.

[48] E. Raidl and H. Rott. Towards a logic for 'because'. *Philosophical Studies*, 2023.

[49] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[50] H. Rott. Difference-making conditionals and the relevant Ramsey test. *Review of Symbolic Logic*, 15(1):133–164, 2022.

[51] M. Sezgin, G. Kern-Isberner, and H. Rott. Inductive reasoning with difference-making

conditionals. In M. Martinez and I. Varcinczak, editors, *Proceedings of the 18th International Workshop on Non-Monotonic Reasoning, NMR 2020*, 2020.

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

[53] 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.

[54] W. Spohn. A brief comparison of Pollock's defeasible reasoning and ranking functions. *Synthese*, 131:39–56, 2002.

[55] W. Spohn. *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford University Press, 2012.

[56] W. Spohn. Enumerative induction. In C. Beierle, G. Brewka, and M. Thimm, editors, *Computational Models of Rationality: Essays Dedicated to Gabriele Kern-Isberner on the Occasion of Her 60th Birthday*, pages 96–114. College Publications, London, 2016.

[57] W. Spohn. The epistemology and auto-epistemology of temporal self-location and forgetfulness. *Ergo*, 4(13):359–148, 2017.

[58] W. Spohn. Defeasible normative reasoning. *Synthese*, 197:1391–1428, 2020.

[59] F. Voorbraak. Probabilistic belief expansion and conditioning. Technical Report LP-96-07, Institute for Logic, Language and Computation, University of Amsterdam, 1996.

[60] J. Williamson. *In Defence of Objective Bayesianism*. Oxford University Press, Oxford, 2010.

Received August 2023