# 1

# Beliefs, Belief Revision, and Noisy Sensors

JAMES DELGRANDE

ABSTRACT. In logical AI, an agent's beliefs are typically categorical, in that they are specified by a set of formulas. An agent may change its beliefs as a result of being informed in one fashion or another about some aspect of the world, or following the execution of some action. The areas of belief revision and reasoning about action deal with just such change in belief. However, most information about the real world is not categorical. While there are well-established accounts for accommodating noncategorical information via probability theory, it is worth asking whether probabilistic information may be reconciled with the logical accounts of belief change. We present such an account in this paper. An agent receives uncertain information as input and its probabilities, expressed as probabilities on possible worlds, are updated via Bayesian conditioning. A set of formulas among the (noncategorical) beliefs is identified as the agent's (categorical) belief set. This set is defined in terms of the most probable worlds such that the summed probability of these worlds exceeds a given threshold. The effect of this updating on the belief set is examined with respect to its appropriateness as a revision operator. It proves to be the case that a subset of the classical AGM belief revision postulates are satisfied. Most significantly, the success postulate is not guaranteed to hold. However it does hold after a sufficient number of iterations. Not is it the case that in revising by a formula consistent with the agent's beliefs, revision corresponds to expansion. On the other hand, limiting cases of the presented approach correspond to specific approaches to revision that have appeared in the literature.

It is a great pleasure to dedicate this paper to Hector Levesque on the occasion of his 60th birthday. While Hector's work has broadly focussed on representational aspects of an agent's beliefs together with accounts of reasoning – whether epistemic, nonmonotonic, limited, in a theory of action, or otherwise – it has certainly touched on many other areas over the years. This paper outlines a possible linking of two such areas, that of reasoning about noisy sensors on the one hand [Bacchus, Halpern, and Levesque 1999], and revision in the presence of (categorical) observations on the other [Shapiro, Pagnucco, Lespérance, and Levesque 2011].

## 1 Introduction

In logical AI, an agent is generally regarded as holding, or *believing*, some set of formulas to be true. As well, an agent's knowledge of a domain will most often be incomplete and inaccurate. Consequently, an agent must change its beliefs in response to receiving new information. *Belief revision* addresses the problem of how an agent may incorporate a new formula into its set of beliefs. That is, the the agent has some set of beliefs $K$ which are accepted as holding in the domain of application, and the agent is given a new formula $\phi$ which it is to incorporate into the set of beliefs. If $\phi$ is consistent but conflicts with $K$, some beliefs will have to be dropped from $K$ before $\phi$ can be added. The original and best-known approach to belief revision is called the *AGM approach* [Alchourrón, Gärdenfors, and Makinson 1985; Gärdenfors 1988], named after the developers of this framework. Subsequently, the area of belief revision has developed into an active area of research in KR&R [Peppas 2007].

However, information about the world is often not categorical, but is received only with a certain level of confidence. For example, an agent may make observations about the world via sensors, but such sensors may be inaccurate or may provide incorrect information. There are of course approaches for modifying an agent's beliefs in such a situation, most obviously via probability theory and using Bayesian conditioning (e.g. [Pearl 1988]). However, in this case formulas are held with attached (subjective) probabilities, and are not generally held as being absolutely true or false.

In one sense, these approaches seem to be addressing the same problem, since they both consider the change in an agent's beliefs in the presence of new information. Yet the approaches also seem to be quite different. In the case of belief revision, an agent *accepts* a certain set of beliefs as categorically holding, and another categorical belief is to be consistently incorporated into this set. This approach is essentially *qualitative*, since the sentences making up the agent's knowledge base are (simply) believed to be true. In the case of updating via Bayesian conditioning, beliefs are generally not held with certainty, but rather with varying levels or degrees of confidence. The task then is to modify these degrees of confidence, expressed as probabilities, as new evidence is received. Hence this approach is fundamentally *quantitative*.

This division into qualitative and quantitative approaches to belief represents two fundamentally different ways of dealing with uncertain information, often referred to as the *logicist* and *probabilist* camps. The distinction can be paraphrased as concerning on the one hand those approaches that adopt a proposition as holding, but hedged in the sense that one is prepared to give it, as opposed to accepting a proposition in a hedged fashion, in that it isn't held with certainty but just with some level of confidence [Kyburg 1994].

However, it can also be observed that categorical beliefs arise from noncategorical, hedged claims: in fact, it can be argued that *none* of our knowledge about the world is certain, yet we often – perhaps most often – act as if it were. Hence, not only do people act assuming that the sun will rise tomorrow, they will usually act as if their car is guaranteed to be where they parked it. Consequently, it is an interesting question to ask how one may move from a noncategorical account of a domain to a categorical account. As alluded to above, Kyburg, among others, has been occupied with this question. In this paper we take a different tack and ask whether an underlying non-categorical approach, based on subjective probabilities, may have something to say about the categorical approach of belief revision. That is, evidence about the real world is generally uncertain, and it is of interest to examine how such a setting may be reconciled with the assumptions underlying belief revision.

We begin with a simple model of an agent's beliefs, in which probabilities are associated with possible worlds which in turn characterise the agent's subjective knowledge. The agent's accepted, categorical beliefs are characterised by the set of worlds with highest probability such that the sum of the probabilities over those worlds exceeds a certain threshold. As new, uncertain information is received, the probabilities attached to worlds are modified and the set of accepted beliefs consequently changes. One can then examine the dynamics of these accepted beliefs to see how it accords with accounts of belief revision. Perhaps not surprisingly, only a subset of the AGM revision postulates are satisfied. Most notably, if a formula $\phi$ is received with an attached probability, it does not necessarily appear among the set of accepted beliefs. However, it proves to be the case that after some number of iterations of revision by $\phi$, $\phi$ will come to be believed. This makes intuitive sense: if one is informed of a formula with probability $> .5$, one may still not immediately believe $\phi$. However after repeated such reports one eventually accepts the formula. We also examine variants of the approach. It proves to be the case that two extant approaches to belief revision are closely related to instances of the approach developed here.

The next section reviews background material. Section 3 reviews the updating of probabilities first, by way of motivation, in terms of formulas and second in terms of probabilities on possible worlds. The following section motivates and defines the notion of *epistemic state* as used in the paper. Section 5 describes belief revision in this framework, including properties of the resulting revision operator and a comparison to related work. Section 6 gives a brief summary.

## 2 Background

### 2.1 Formal Preliminaries

Let $\mathcal{P} = \{a, b, \ldots\}$ be a finite set of atomic sentences, and let $\mathcal{L}$ be the language over $\mathcal{P}$ closed under the usual connectives $\neg$, $\wedge$, $\vee$, and $\supset$. The classical consequence relation is denoted $\vdash$; $\mathcal{C}n(A)$ is the set of logical consequences of a formula or set of formulas $A$; that is $\mathcal{C}n(A) = \{\phi \in \mathcal{L} \mid A \vdash \phi\}$. $\top$ stands for some arbitrary tautology and $\bot$ is defined to be $\neg\top$. Given two sets of formulas $A$ and $B$, $A + B$ denotes the *expansion* of $A$ by $B$, that is $A + B = \mathcal{C}n(A \cup B)$. Expansion of a set of formulas $A$ by a formula $\phi$ is defined analogously. Sentences $\phi$ and $\psi$ are *logically equivalent*, $\phi \equiv \psi$, iff $\phi \vdash \psi$ and $\psi \vdash \phi$. This also extends to sets of formulas. A propositional *interpretation* (or *possible world*) is a mapping from $\mathcal{P}$ to {true, false}. The set of interpretations of $\mathcal{L}$ is denoted $W_{\mathcal{L}}$. A *model* of a sentence $\phi$ is an interpretation $w$ that makes $\phi$ true according to the usual definition of truth, and is denoted by $w \models \phi$. We also write $W \models \phi$ if $w \models \phi$ for every $w \in W$. $Mod(A)$ is the set of models of the set of formulas $A$. $Mod(\{\phi\})$ is also written as $Mod(\phi)$. For $W \subseteq W_{\mathcal{L}}$, we denote by $\mathcal{T}(W)$ the set of sentences which are true in all elements of $W$; that is $\mathcal{T}(W) = \{\phi \in \mathcal{L} \mid w \models \phi \text{ for every } w \in W\}$.

A total preorder $\preceq$ is a reflexive, transitive binary relation, such that either $w_1 \preceq w_2$ or $w_2 \preceq w_1$ for every $w_1, w_2$. As well, $w_1 \prec w_2$ iff $w_1 \preceq w_2$ and $w_2 \not\preceq w_1$. $w_1 = w_2$ abbreviates $w_1 \preceq w_2$ and $w_2 \preceq w_1$. Given a set $S$ and total preorder $\preceq$ defined on members of $S$, we denote by $\min(S, \preceq)$ the set of minimal elements of $S$ in $\preceq$.

Let $\rho : W_{\mathcal{L}} \mapsto [0, 1]$ be a function such that $0 \leq \rho(w) \leq 1$ and $\sum_{w \in W_{\mathcal{L}}} \rho(w) = 1$. $\rho$ is a *probability assignment* to worlds. We distinguish the function $\rho_\top$ where $\rho_\top(w) = \frac{1}{|W_{\mathcal{L}}|}$ for every world $w$, and we use $\rho_\bot$ to denote the (non-probability) assignment where $\rho_\bot(w) = 1$ for every world $w$. $\rho_\top$ can be used to characterise a state of ignorance for an agent, while $\rho_\bot$ is a technical convenience that will be used to characterise an inconsistent set of beliefs. Mention of probability assignments will include $\rho_\bot$ as a special case. These functions are extended to subsets of $W_{\mathcal{L}}$ by, for $W \subseteq W_{\mathcal{L}}$, $\rho(W) = \sum_{w \in W} \rho(w)$. Informally, $\rho(w)$ is the (subjective) probability that, as far as the agent knows, $w$ is the actual world being modelled; and for $W \subseteq W_{\mathcal{L}}$, $\rho(W)$ is the probability that the real world is a member of $W$. As will be later described, the function $\rho$ can be taken as comprising the major part of an agent's *epistemic state* [Darwiche and Pearl 1997; Peppas 2007]. The probability of a formula $\phi$ then is given by: $Pr_\rho(\phi) = \sum_{w \models \phi} \rho(w) = \rho(Mod(\phi))$. *Conditional probability* is defined, as usual, by $Pr_\rho(\phi|\psi) = Pr_\rho(\phi \wedge \psi)/Pr_\rho(\psi)$ and is undefined when $Pr_\rho(\psi) = 0$.

### 2.2 Belief revision

The *AGM approach* [Gärdenfors 1988] provides the best-known approach to belief revision. Belief change is described at the *knowledge level*, that is on an abstract level, independent of how beliefs are represented and manipulated. An agent's beliefs are modelled by a set of sentences, or *belief set*, closed under the logical consequence operator of a logic that includes classical propositional logic. Thus a belief set $K$ satisfies the constraint: $\phi \in K$ if and only if $K \vdash \phi$. $K$ can be understood as a partial theory of the world. $K_\perp$ is the inconsistent belief set (i.e. $K_\perp = \mathcal{L}$).

In the revision of $K$ by a formula $\phi$, the intent is that $\phi$ is to be incorporated into $K$ so that the resulting belief set is consistent whenever $\phi$ is consistent. If $\phi$ is inconsistent with $K$, revision will require the removal of beliefs from $K$ in order to retain consistency. In this approach, revision is a function from belief sets and formulas to belief sets. However, various researchers have argued that it is more appropriate to consider *epistemic states* (also called *belief states*) as objects of revision. An epistemic state $\mathcal{K}$ includes information regarding how the revision function itself changes following a revision. The belief set corresponding to belief state $\mathcal{K}$ is denoted $Bel(\mathcal{K})$. As well, we will use the notation $Mod(\mathcal{K})$ to mean $Mod(Bel(\mathcal{K}))$. Formally, a revision operator $*$ maps an epistemic state $\mathcal{K}$ and new information $\phi$ to a revised epistemic state $\mathcal{K}*\phi$. Then, in the spirit of [Darwiche and Pearl 1997], the AGM postulates for revision can be reformulated as follows:

($\mathcal{K}*1$) $Bel(\mathcal{K}*\phi) = Cn(Bel(\mathcal{K}*\phi))$

($\mathcal{K}*2$) $\phi \in Bel(\mathcal{K}*\phi)$

($\mathcal{K}*3$) $Bel(\mathcal{K}*\phi) \subseteq Bel(\mathcal{K}) + \phi$

($\mathcal{K}*4$) If $\neg\phi \notin Bel(\mathcal{K})$ then $Bel(\mathcal{K}) + \phi \subseteq Bel(\mathcal{K}*\phi)$

($\mathcal{K}*5$) $Bel(\mathcal{K}*\phi)$ is inconsistent, only if $\nvdash \neg\phi$

($\mathcal{K}*6$) If $\phi \equiv \psi$ then $Bel(\mathcal{K}*\phi) \equiv Bel(\mathcal{K}*\psi)$

($\mathcal{K}*7$) $Bel(\mathcal{K}*(\phi\wedge\psi)) \subseteq Bel(\mathcal{K}*\phi) + \psi$

($\mathcal{K}*8$) If $\neg\psi \notin Bel(\mathcal{K}*\phi)$ then $Bel(\mathcal{K}*\phi) + \psi \subseteq Bel(\mathcal{K}*(\phi\wedge\psi))$

The postulates express very basic properties for revision. Thus, the result of revising $\mathcal{K}$ by $\phi$ is an epistemic state in which $\phi$ is believed in the corresponding belief set (($\mathcal{K}*1$), ($\mathcal{K}*2$)); whenever the result is consistent, the revised belief set consists of the expansion of $Bel(\mathcal{K})$ by $\phi$ (($\mathcal{K}*3$), ($\mathcal{K}*4$)); the only time that $Bel(\mathcal{K})$ is inconsistent is when $\phi$ is inconsistent (($\mathcal{K}*5$)); and revision is independent of the syntactic form of the formula for revision (($\mathcal{K}*6$)). The last two postulates state that whenever consistent, revision by a conjunction corresponds to revision by one conjunct and expansion by the other.

Various constructions have been proposed to characterise belief revision. Katsuno and Mendelzon [1991] have shown that a revision can be characterised in terms of a total preorder on the set of possible worlds. For epistemic state $\mathcal{K}$, a *faithful ranking* on $\mathcal{K}$ is a total preorder $\preceq_\mathcal{K}$ on the possible worlds $W_\mathcal{L}$, such that for any possible worlds $w_1, w_2 \in W_\mathcal{L}$:

1. If $w_1, w_2 \models Bel(\mathcal{K})$ then $w_1 =_\mathcal{K} w_2$
2. If $w_1 \models Bel(\mathcal{K})$ and $w_2 \not\models Bel(\mathcal{K})$, then $w_1 \prec_\mathcal{K} w_2$

Intuitively, $w_1 \preceq_{\mathcal{K}} w_2$ if $w_1$ is at least as plausible as $w_2$ according to the agent. The first condition asserts that all models of the agent's knowledge are ranked equally, while the second states that the models of the agent's knowledge are lowest in the ranking. It follows directly from the results of [Katsuno and Mendelzon 1991] that a revision operator $*$ satisfies ($\mathcal{K}*1$)–($\mathcal{K}*8$) iff there exists a faithful ranking $\preceq_{\mathcal{K}}$ for an arbitrary belief state $\mathcal{K}$, such that for any sentence $\phi$:

$$Bel(\mathcal{K}*\phi) = \begin{cases} \mathcal{L} & \text{if } \vdash \neg\phi \\ \mathcal{T}(\min(Mod(\phi), \preceq_{\mathcal{K}})) & \text{otherwise} \end{cases}$$

Thus when $\phi$ is satisfiable, the belief set corresponding to $\mathcal{K}*\phi$ is characterised by the least $\phi$ models in the ranking $\preceq_{\mathcal{K}}$.

The AGM postulates do not address properties of iterated belief revision. This has led to the development of additional postulates for iterated revision; the best-known approach is that of Darwiche and Pearl [1997]. They propose the following postulates, adapted according to our notation:

**C1** If $\psi \vdash \phi$, then $Bel((\mathcal{K}*\phi)*\psi) = Bel(\mathcal{K}*\psi)$

**C2** If $\psi \vdash \neg\phi$, then $Bel((\mathcal{K}*\phi)*\psi) = Bel(\mathcal{K}*\psi)$

**C3** If $\phi \in Bel(\mathcal{K}*\psi)$, then $\phi \in Bel((\mathcal{K}*\phi)*\psi)$

**C4** If $\neg\phi \notin Bel(\mathcal{K}*\psi)$, then $\neg\phi \notin Bel((\mathcal{K}*\phi)*\psi)$

Darwiche and Pearl show that an AGM revision operator satisfies each of the Postulates (C1)–(C4) iff the way it revises faithful rankings satisfies the respective conditions:

**CR1** If $w_1, w_2 \models \phi$, then $w_1 \preceq_{\mathcal{K}} w_2$ iff $w_1 \preceq_{\mathcal{K}*\phi} w_2$

**CR2** If $w_1, w_2 \not\models \phi$, then $w_1 \preceq_{\mathcal{K}} w_2$ iff $w_1 \preceq_{\mathcal{K}*\phi} w_2$

**CR3** If $w_1 \models \phi$ and $w_2 \not\models \phi$, then $w_1 \prec_{\mathcal{K}} w_2$ implies $w_1 \prec_{\mathcal{K}*\phi} w_2$

**CR4** If $w_1 \models \phi$ and $w_2 \not\models \phi$, then $w_1 \preceq_{\mathcal{K}} w_2$ implies $w_1 \preceq_{\mathcal{K}*\phi} w_2$

Thus postulate (C1) asserts that revising by a formula and then by a logically stronger formula yields the same belief set as simply revising by the stronger formula at the outset. The corresponding semantic condition (CR1) asserts that in revising by a formula $\phi$, the relative ranking of $\phi$ worlds remains unchanged. The other postulates and semantic conditions can be interpreted similarly. Subsequently, other approaches for iterated revision have been proposed, including [Boutilier 1996; Nayak, Pagnucco, and Peppas 2003; Jin and Thielscher 2007]. While interesting, we do not consider them further since they add little to the exposition.

### 2.3 Related Work

In probability theory and related approaches, there has of course been work on incorporating new evidence to produce a new probability distribution. The simplest means of updating probabilities is via conditionalisation: If an agent holds $\phi$ with probability $q$, and so $Pr(\phi) = q$, and the agent learns $\psi$ with certainty, then one can define the updated probability $Pr'(\phi)$ via

$$Pr'(\phi) \ = \ Pr(\phi|\psi) \ = \ Pr(\phi \wedge \psi)/Pr_\rho(\psi).$$

Of course an agent may not learn $\psi$ with certainty, but rather may change its probability assignment to $\psi$ from $Pr(\psi)$ to a new value $Pr'(\psi)$. The question then is how probabilities assigned to other variables should be modified. Jeffrey [1983] proposes that for proposition $\phi$, the new probability should be given by what has come to be known as Jeffrey's Rule for updating probabilities:

$$Pr'(\phi) \;=\; Pr(\phi|\psi)Pr'(\psi) + Pr(\phi|\neg\psi)Pr'(\neg\psi).$$

So $Pr'(\psi) = q$ means that the agent has learned that the probability of $\psi$ is $q$. In particular, if the probability of $\psi$ is further updated to $Pr''(\psi)$ but it turns out that $Pr''(\psi) = Pr'(\psi)$, then the distributions $Pr'$ and $Pr''$ will coincide.

This is orthogonal to our goals here. Instead, we are interested in the case where we have some underlying proposition, say that a light is on, represented by $on$, and we are given an observation $Obs_{on}$, where $Obs_{on}$ has an attached probability. Then if the agent receives repeated observations that the light is on, the agent's confidence that $on$ is true will increase with each positive observation. Details are given in the next section; the main point here is that Bayes' Rule will be more appropriate in this case, where Bayes' Rule is given by:

$$Pr(\phi|\psi) \;=\; Pr(\psi|\phi)Pr(\phi)/Pr(\psi).$$

Previous research dealing with the intersection of belief revision and probability is generally concerned with revising a probability function. In such approaches, an agent's belief set $K$ is given by those formulas that have probability 1.0. These formulas with probability 1.0 are referred to as the *top* of the probability function. For a revision $K * \phi$, the probability function is revised by $\phi$, and the belief set corresponding to $K * \phi$ is given by the top of the resulting probability function. So such approaches allow the characterisation of not just the agent's beliefs, but also allow probabilities to be attached to non-beliefs. As will be subsequently described, this is in contrast to the present approach, in which an agent's categorical beliefs will generally have probability less than 1.

One difficulty with revising probability functions is the *non-uniqueness problem*, that there are many different probability functions that have $K$ as their top. Lindström and Rabinowicz [1989] consider various ways of dealing with this problem. Boutilier [1995] considers the same general framework, but rather focuses on issues of iterated belief revision. However the approach described herein addresses a different problem: a means of incorporating uncertain information into a given probability function is assumed, and the question addressed is how such an approach may be reconciled with AGM revision, or alternatively, how such an approach may be considered as an instance (or proto-instance) of revision. To this end, Gärdenfors [1988, Ch. 5] has also considered an extension of the AGM approach to the revision of probability functions; we discuss this work in detail after our approach has been presented.

With respect to qualitative, AGM-style belief revision, the approach at hand might seem to be an instance of an *improvement operator* [Konieczny and Pino Pérez 2008]. An improvement operator according to Konieczny and Pino Pérez is a belief change operator where new information isn't necessarily immediately accepted. However plausibility is increased and, after a sufficient number of iterations, the information will come to be believed. Interestingly, as we discuss later, the approach described here differs in significant ways from those of [Konieczny and Pino Pérez 2008].

The setting adopted here is similar to that of [Bacchus, Halpern, and Levesque 1999]: Agents receive uncertain information, and as a result alter their (probabilistic) beliefs about the world. However, the goals are quite different. [Bacchus, Halpern, and Levesque 1999] is concerned with an extension of the situation calculus [Levesque, Pirri, and Reiter 1998] to deal with noisy sensors. Consequently their focus is on a version of the situation calculus in which the agent doesn't hold just categorical beliefs, but also probabilistic beliefs. The main issue then is how to revise these probabilities in the presence of sensing and non-sensing actions. In contrast, the present paper is concerned with the possible role of probabilistic beliefs with respect to a (classical AGM-style) belief revision operator. We further discuss this and other related work once the approach has been presented.

## 3 Unreliable Observations and Updating Probabilities

An agent will make observations concerning a domain. These observations may be unreliable, in that a value may be incorrectly sensed or reported. We wish to update the probability assignment to possible worlds appropriately, given such a possibly-erroneous observation. Consider first a situation in which an agent observes or senses $\phi$ with a given probability $q > .5$. Our interpretation of this event is that $\phi$ is reported as being true, but that the probability is $1-q$ that the sensing is incorrect (and so the probability is $1-q$ that $\neg\phi$ is in fact the case).[1] Since $q > .5$, the agent's confidence in $\phi$ will increase. We can write $Pr(Obs_\phi|\phi) = q$ for the probability of observing that $\phi$ is true given that $\phi$ is in fact true. As well, the agent will also have some prior probability $Pr(\phi)$ that $\phi$ is true. We wish to compute the probability that $\phi$ is true given the new piece of evidence, $Pr(\phi|Obs_\phi)$. This can be determined by Bayes' Rule:

$$(1) \quad Pr(\phi|Obs_\phi) \;=\; \frac{Pr(Obs_\phi|\phi)Pr(\phi)}{Pr(Obs_\phi)} \;=\; \frac{Pr(Obs_\phi|\phi)Pr(\phi)}{Pr(Obs_\phi|\phi)Pr(\phi) + Pr(Obs_\phi|\neg\phi)Pr(\neg\phi)}$$

For example assume that the agent has no prior information about a light being on or not, and so $Pr(on) = .5$. As well, the light sensor is correct 80% of the time, and so $Pr(Obs_{on}|on) = .8$ while $Pr(Obs_{on}|\neg on) = .2$. Hence following an observation that the light is on, we would obtain:

$$\begin{aligned} Pr(on|Obs_{on}) &= \frac{Pr(Obs_{on}|on)Pr(on)}{Pr(Obs_{on}|on)Pr(on) + Pr(Obs_{on}|\neg on)Pr(\neg on)} \\ &= (.8 \times .5)/[(.8 \times .5) + (.2 \times .5)] \\ &= .8. \end{aligned}$$

Thus on observing that the light was on, the agent's (subjective) probability that the light was on would increase from .5 to .8. If the agent was to subsequently re-sense the light, and again sense that the light was on, its degree of belief would then be given by:

$$\begin{aligned} Pr(on|Obs_{on}) &= \frac{Pr(Obs_{on}|on)Pr(on)}{Pr(Obs_{on}|on)Pr(on) + Pr(Obs_{on}|\neg on)Pr(\neg on)} \\ &= (.8 \times .8)/[(.8 \times .8) + (.2 \times .2)] \\ &\approx .94. \end{aligned}$$

[1] The case of non-binary valued sensing is straightforward and adds nothing of additional interest with respect to the problem at hand; see for example [Bacchus, Halpern, and Levesque 1999] for how this can be handled.

**Observations and possible worlds:** The preceding discussion reviews how a probability assignment to formulas may be updated given new information. Since we have a finite language and a finite set of possible worlds, it is straightforward to extend this to updating probabilities attached to worlds, and hence updating a probability function $\rho$. Since a world may be associated with the conjunction of literals true at that world, we can repeat the steps in the preceding section, but with respect to worlds.

Consider a situation in which we observe $\phi$ with probability $q$. As before, $Obs_\phi$ is true if $\phi$ is observed and false otherwise. We wish to update the probability that a world $w$ is the real world given this additional piece of information. That is, if $\phi$ is true at $w$ then the probability attached to $\rho$ will increase, and decrease if $\phi$ is false at $w$. For a world $w \in W_{\mathcal{L}}$, we have the prior probability assignment $\rho(w)$. We can use this in probability expressions by letting $Pr_\rho(w)$ be understood such that the occurrence of $w$ in $Pr_\rho(\cdot)$ stands for the (finite) conjunction of literals true in $w$, and thus $Pr_\rho(w) = \rho(w)$. Then, again with Bayes' rule we have:

$$(2) \quad Pr_\rho(w|Obs_\phi) \;=\; \frac{Pr_\rho(Obs_\phi|w)Pr_\rho(w)}{Pr_\rho(Obs_\phi)}$$

Thus on the left side of the equality, we wish to determine the (updated) probability of $w$, given that $\phi$ is observed. For the numerator on the right hand side, $Pr_\rho(Obs_\phi|w)$ is the probability of observing $\phi$ given that one is in $w$; this is just $q$ if $w \models \phi$ and $1-q$ otherwise. $Pr_\rho(w)$ is just $\rho(w)$. For the denominator, we have that

$$\begin{aligned} Pr_\rho(Obs_\phi) &= Pr_\rho(Obs_\phi|\phi)Pr_\rho(\phi) + Pr_\rho(Obs_\phi|\neg\phi)Pr_\rho(\neg\phi) \\ &= q \times \rho(Mod(\phi)) \;+\; (1-q) \times \rho(Mod(\neg\phi)). \end{aligned}$$

This justifies the following definition.

DEFINITION 1. Let $\rho$ be a probability assignment to worlds. Let $\phi \in \mathcal{L}$ and $q \in [0,1]$. Let $\eta \;=\; \rho(Mod(\phi)) \times q \;+\; \rho(Mod(\neg\phi)) \times (1-q)$.

Define the probability assignment $\rho(\phi, q)$ by:

$$\begin{aligned} \rho(\phi, q) &= \rho_\perp \qquad\qquad\qquad\quad \text{if } \eta = 0; \quad \text{otherwise:} \\ \rho(\phi, q)(w) &= \begin{cases} (\rho(w) \times q)/\eta & \text{if } w \models \phi \\ (\rho(w) \times (1-q))/\eta & \text{if } w \not\models \phi \end{cases} \end{aligned}$$

Thus, for probability function $\rho$, a new probability function $\rho(\phi, q)$ results after sensing $\phi$ with probability $q$. Observe that $\rho(\phi, q)(w)$ in Definition 1 corresponds to $Pr_\rho(w|Obs_\phi)$ in (2). If $\eta = 0$, the updated probability assignment involves accepting with certainty (i.e. $q = 1$) an impossible proposition ($\rho(Mod(\phi)) = 0$), or rejecting with certainty a necessarily true proposition. In either case, an incoherence state of affairs ($\rho_\perp$) results.

**Example:** Consider Table 1. The first column lists possible worlds in terms of an assignment of truth values to atoms, where $\overline{a}$ stands for $\neg a$. The second column gives an initial probability function, while the next three columns show how $\rho$ changes under different updates. At the outset $Pr(a) = .5$, $Pr(b) = .6$, and $Pr(c) = .5$. Following an observation of $a$ with reliability .8, we obtain that $Pr(a) = .8$, $Pr(b) = .6$, and $Pr(c) = .5$. If we iterate the process and again observe $a$ with the same reliability, the probabilities become $Pr(a) = .9412$, $Pr(b) = .6$, and $Pr(c) = .5$. Thus the probability of $a$ increases, and the probability of $b$ and $c$ varies depending on the probabilities assigned to individual worlds;

| $Worlds$ | $\rho$ | $\rho(a,.8)$ | $\rho(a,.8)(a,.8)$ | $\rho(a,.8)(b,.8)$ |
|---|---|---|---|---|
| $a,b,c$ | .150 | .240 | .2824 | .3333 |
| $a,b,\overline{c}$ | .150 | .240 | .2824 | .3333 |
| $a,\overline{b},c$ | .100 | .160 | .1882 | .0556 |
| $a,\overline{b},\overline{c}$ | .100 | .160 | .1882 | .0556 |
| $\overline{a},b,c$ | .150 | .060 | .0176 | .0972 |
| $\overline{a},b,\overline{c}$ | .150 | .060 | .0176 | .0972 |
| $\overline{a},\overline{b},c$ | .100 | .040 | .0118 | .0139 |
| $\overline{a},\overline{b},\overline{c}$ | .100 | .040 | .0118 | .0139 |

Table 1. Example of Updating Probabilities of Worlds

in the example they happen to be unchanged. Last if we sense $a$ and then $b$, in both cases with reliability .8 we obtain that $Pr(a) = .7778$, $Pr(b) = .8610$, and $Pr(c) = .5$.

## 4 Epistemic States: A Model of Categorical Belief based on Noncategorical Belief

This section presents the notion of *epistemic state* as it is used in the present approach. We first discuss intuitions then give the formal details.

### 4.1 Intuitions

An agent's epistemic state $\mathcal{K}$ is given by a pair $(\rho, c)$, where $\rho$ is a probability assignment over possible worlds and $c$ is a *confidence level*. The probability function captures what the agent knows about the world. We wish to say that an agent *accepts* a belief represented by a formula just if, in some fashion, the probability of the formula exceeds the confidence level $c$. Thus, the agent accepts a formula if its probability is "sufficiently high". This notion of *acceptance* is nonstandard, in that an accepted belief will be categorical yet its associated probability may be less than 1. This is in contrast to the approaches combining probability and revision described in Section 2.3, where the agent's categorical beliefs have probability 1. This also is in contrast with [Bacchus, Halpern, and Levesque 1999], where non-beliefs have probability 0. In any case, for us an accepted belief is one that is categorical, in that the agent may act under the assumption that it is true, yet it is also noncategorical, in that its probability is less than 1, and hence it can be given up following a revision. The issue then becomes one of suitably defining the worlds characterising an agent's accepted beliefs.

The most straightforward way of defining acceptance is to say that a formula $\phi$ is accepted just if $\rho(Mod(\phi)) \geq c$. This leads immediately to the *lottery paradox* [Kyburg 1961]. This problem is that for any $c < 1.0$ one can construct a scenario where $p_1, \ldots, p_n$, along with $\neg p_1 \vee \ldots \vee \neg p_n$ are all accepted. But the set consisting of these formulas is of course inconsistent. The resolution proposed here is to focus instead on the set of possible worlds characterising an agent's beliefs. That is, the agent's (categorical) beliefs will be identified with a subset of possible worlds in which the set of beliefs is true. The issue then is to determine the appropriate subset of possible worlds. To this end, we suggest that the appropriate set of worlds is comprised of those worlds of greatest probability such that the probability of the set exceeds $c$. Since the agent's accepted beliefs are characterised by a unique set of worlds, the lottery paradox doesn't arise.

The assumption that worlds with higher probability are to be preferred to those with lower probability for characterising an agent's beliefs can be justified by (at least) two arguments. First, if an agent had to commit to a single world being the real world, then it would choose a world $w$ for which the probability $\rho(w)$ was maximum; if it had to commit to $n$ worlds, then it would choose the $n$ worlds with highest probability. Similarly, if it were to choose the most likely set of worlds containing the real world, such that the probability of the set exceeded a certain bound (here $c$), then it would choose the set of worlds of maximal probability that meets or exceeds $c$. Since there is nothing that distinguishes worlds beyond their probability, if $\rho(w) = \rho(w')$ then if $w$ is in this set then so is $w'$.

A second argument is related to a principle of *informational economy*: It seems reasonable to assume that, given a set of candidate belief sets, an agent will prefer a set that gives more information over one that gives less. This is the case here. In general there will be more than one set of worlds where the probability of the set exceeds $c$. The set composed of worlds of maximal probability is generally also the set with the least number of worlds, which in turn will correspond to the belief set with the maximum number of (logically distinct) formulas. So this approach commits the agent to the maximum set of accepted beliefs, where the overall probability of the set exceeds $c$.[2] Such a set may be said to have the greatest *epistemic content* among the candidate belief sets.

Thus an epistemic state consists principally of a probability function on possible worlds. Via an assumption of maximality of beliefs (or maximal epistemic content), and given the confidence level $c$, a set of accepted beliefs is defined. So this differs significantly from prior work, in that an accepted formula will generally have an associated probability that is less than 1.0. Arguably this makes sense: for example, I believe that my car is where I left it this morning, in that I act as if this was a true fact even though I don't hold that it is an absolute certainty that the car is where I left it. If pressed, I would be happy to attach a probability to the possibility of my car not being where I left it, but I would continue to act as if it were (simply) true that my car was where I left it. Moreover, of course, I am prepared to revise this belief if I receive information to the contrary.

### 4.2 Epistemic States: Formal Details

In this subsection we define our notion of epistemic state, and relate it to *faithful rankings* that have been used to characterise AGM revision.

DEFINITION 2. $\mathcal{K} = (\rho, c)$ is an *epistemic state*, where:

- $\rho$ is probability assignment to possible worlds and
- $c \in (0, 1]$ is a *confidence level*.

As described, an epistemic state characterises the state of knowledge of the agent, both its (contingent) beliefs as well as, implicitly, those beliefs that it would adopt or abandon in the presence of new information. We need to also define the agent's *belief set* or beliefs about the world at hand. This is most easily done by first defining the worlds that characterise the agent's belief set, and then defining the belief set in terms of these worlds.

DEFINITION 3. For epistemic state $\mathcal{K} = (\rho, c)$, the set of worlds characterising the agent's belief set, $Mod(\mathcal{K}) \subseteq W_{\mathcal{L}}$, is the least set such that:

If $\rho = \rho_{\perp}$ then $Mod(\mathcal{K}) = \emptyset$; otherwise:

[2] These notions of course make sense only in a finite (under equivalence classes) language, which was assumed at the outset.

1. $\rho(Mod(\mathcal{K})) \geq c$,
2. If $w \in Mod(\mathcal{K})$ and $w' \notin Mod(\mathcal{K})$ then $\rho(w) > \rho(w')$.

$Mod(\cdot)$ is uniquely characterised; in particular we have that if $\rho(w) = \rho(w')$ then $w \in Mod(\mathcal{K})$ iff $w' \in Mod(\mathcal{K})$.

DEFINITION 4. For epistemic state $\mathcal{K}$, the agent's accepted (categorical) beliefs, $Bel(\mathcal{K})$, are given by

$$Bel(\mathcal{K}) \;=\; \{\phi \mid Mod(\mathcal{K}) \models \phi\} \;=\; \mathcal{T}(Mod(\mathcal{K})).$$

Thus, an agent accepts a sentence if it is sufficiently likely, and a sentence is "sufficiently likely" if it is true in the set of most plausible worlds such that the probability of the set exceeds the given confidence level. Since an agent's beliefs are characterised by a single set of possible worlds, the lottery paradox doesn't arise.

This then describes the static aspects of an epistemic state. For the dynamic aspects (i.e. revision) it will be useful to distinguish those formulas that are *possible*, in the sense that they are *conceivable*, which is to say they have a non-zero probability. We use $Poss_{\mathcal{K}}(\phi)$ to indicate that, according to the agent involved, $\phi$ is possible; that is, there is possible world $w$ such that $w \models \phi$ and $\rho(w) > 0$. $Poss_{\mathcal{K}}(\cdot)$ can be axiomatised as the modality $\Diamond$ in the modal logic S5 [Hughes and Cresswell 1996]. We have the simple consequence:

PROPOSITION 5. *If not* $Poss_{\mathcal{K}}(\phi)$ *then* $\neg\phi \in Bel(\mathcal{K})$

The probability assignment to possible worlds defines a ranking on worlds, where worlds with higher probability are lower in the ranking:

DEFINITION 6. For given $\rho$, define $rank_\rho(w)$ for every $w \in W_{\mathcal{L}}$ by:

1. $rank_\rho(w) = 0$ if $\nexists w'$ such that $\rho(w') > \rho(w)$
2. Otherwise, $rank_\rho(w) = 1 + \max\{rank_\rho(w') : \rho(w') > \rho(w)\}$.

Lastly, we can define a *faithful ranking* (as given in Section 2) to relate the ranking defined here to rankings used in belief revision:

DEFINITION 7. The *faithful ranking* $\preceq_{\mathcal{K}}$ is given by:

1. If $w_1, w_2 \models Bel(\mathcal{K})$ then $w_1 =_{\mathcal{K}} w_2$
2. If $w_1 \models Bel(\mathcal{K})$ and $w_2 \not\models Bel(\mathcal{K})$ then $w_1 \prec_{\mathcal{K}} w_2$
3. Otherwise if $rank_\rho(w_1) \leq rank_\rho(w_2)$ then $w_1 \preceq_{\mathcal{K}} w_2$.

Thus, a faithful ranking on worlds can be defined in a straightforward manner from an epistemic state as given in Definition 2. The first two conditions stipulate that we have a faithful ranking. The third condition ensures that we have total preorder that conforms to the probability assignment for those worlds not in $Mod(\mathcal{K})$. Clearly this faithful ranking suppresses detail found in $\rho$. First, quantitative information is lost in going from Definition 2 to Definition 6. Second, gradations in an agent's beliefs are lost: worlds in $Mod(\mathcal{K})$ may have varying probabilities, yet in the corresponding faithful ranking given in Definition 7, all worlds in $Mod(\mathcal{K})$ are ranked equally. Consequently, the notion of epistemic state as defined here is a richer structure than that of a faithful ordering.

## 5 Belief Dynamics in a Probabilistic Framework

We next consider how this approach fits with work in belief revision. A natural way to define the revision of an epistemic state $\mathcal{K} = (\rho, c)$ by $\phi$ with reliability $q$ is to set $\mathcal{K} * (\phi, q) = (\rho(\phi, q), c)$. Of course, revision so defined is a ternary function, as opposed to the usual expression of revision as a binary function, $\mathcal{K} * \phi$. There are various ways in which this mismatch may be resolved. First, we could simply regard revision in a probabilistic framework as a ternary function, with the extra argument giving the reliability of the observation. This is problematic, with regards to our aims, since a ternary operator represents a *quantitative* approach, where the degree of support $q$ of $\phi$ is taken into account. In contrast, AGM revision is *qualitative*, in that for a revision $\mathcal{K} * \phi$, it is the (unqualified) formula $\phi$ that is a subject of revision. This clash then highlights the main issue of this paper: a probabilistic approach is intrinsically quantitative, while standard approaches to belief revision are inherently qualitative.

In re-considering revision $*$ as a binary function, the intent is that in expressing $\mathcal{K}*(\phi, q)$ as a binary function $\mathcal{K} * \phi$, we want to study *properties* of the function $*$ without regard to specific values assigned to $q$. Consequently, we assume that the reliability of a revision is some fixed probability $q$. Given that the reliability is fixed, we can drop the probability argument from a statement of revision, and simply write $\mathcal{K} * \phi$. We later also consider the situation where the reliability of observations may vary.

Revision by $\phi$ is intended to *increase* the agent's confidence in $\phi$, and so for $\mathcal{K} * \phi$ it is understood that the probability of $\phi$ is greater than 0.5. Since revision corresponds to the incorporation of *contingent* information, it is reasonable to assume that nothing can be learned with certainty, and so we further assume that $q < 1$.[3] Consequently, in what follows, we assume that the reliability of a revision is a fixed number $q$ in the range $(0.5, 1.0)$.

DEFINITION 8. Let $q \in (0.5, 1.0)$ be fixed. Let $\mathcal{K} = (\rho, c)$ be an epistemic state and $\phi \in \mathcal{L}$. Define the revision of $\mathcal{K}$ by $\phi$ by:

$$\mathcal{K} * \phi = (\rho(\phi, q), c)$$

Clearly one needs to know the value of $q$ (along with $\mathcal{K}$ and $\phi$) before being able to determine $\mathcal{K}*\phi$. However, without knowing the value of $q$, one can still investigate properties of the class of revision functions, which is our goal here. Other aspects of the definition are discussed below, in the discussion of postulates. We first revisit our previous example.

**Example (continued):** Consider again Table 1, and assume that our initial epistemic state is given by $\mathcal{K} = (\rho, 0.9)$. At the outset, $Bel(\mathcal{K}) = \mathcal{C}n(\top)$. If the probability associated with the world given by $\{\overline{a}, \overline{b}, \overline{c}\}$ was .05, with the balance distributed uniformly across other possible worlds, we would have $Bel(\mathcal{K}) = \mathcal{C}n(a \vee b \vee c)$.

We obtain that $Bel(\mathcal{K} * a) = \mathcal{C}n(a \vee b)$, and $Bel(\mathcal{K} * a * a) = \mathcal{C}n(a)$. We also obtain $Bel(\mathcal{K}*a*b) = \mathcal{C}n(a \vee b)$ and (not illustrated in Table 1) $Bel(\mathcal{K}*a*b*b) = \mathcal{C}n(b)$. So, not unexpectedly, for repeated iterations, the resulting belief set "converges" toward accepting the iterated formula, with the results being biased by the initial probability distribution.

[3]It might be pointed out that a tautology can be learned with absolute certainty. However, it can be pointed out in return that a tautology is in fact *known* with certainty, so the probability being 1 or less makes no difference. In any case, we later examine the situation where $q = 1$.

### 5.1 Properties of Probability-Based Belief Revision

Recall that $Poss_{\mathcal{K}}(\phi)$ indicates that, according to the agent, $\phi$ is possible, in that there is $w$ such that $w \models \phi$ and $\rho(w) > 0$. $\mathcal{K} *^n \phi$ stands for the $n$-fold iteration of $\mathcal{K} * \phi$, that is:

$$\mathcal{K} *^n \phi \;=\; \begin{cases} \mathcal{K} * \phi & \text{if } n = 1 \\ (\mathcal{K} *^{n-1} \phi) * \phi & \text{otherwise} \end{cases}$$

We obtain the following results; numbering corresponds to the AGM revision postulates.

THEOREM 9. *Let $\mathcal{K}$ be an epistemic state and $\phi$, $\psi \in \mathcal{L}$.*

($\mathcal{K} * 1$) $Bel(\mathcal{K} * \phi) = Cn(Bel(\mathcal{K} * \phi))$

($\mathcal{K} * 2a$) *If $Poss_{\mathcal{K}}(\phi)$ then $\phi \in Bel(\mathcal{K} *^n \phi)$ for some $n > 0$*

($\mathcal{K} * 2b$) *If $\mathcal{K} \neq \mathcal{K}_\perp$ and $\phi \in Bel(\mathcal{K})$ then $\phi \in Bel(\mathcal{K} * \phi)$*

($\mathcal{K} * 2c$) *If $\mathcal{K} \neq \mathcal{K}_\perp$ and not $Poss_{\mathcal{K}}(\phi)$ then $Bel(\mathcal{K} * \phi) = Bel(\mathcal{K})$*

($\mathcal{K} * 5$) *$Bel(\mathcal{K} * \phi)$ is consistent.*

($\mathcal{K} * 6$) *If $\phi \equiv \psi$ then $Bel(\mathcal{K} * \phi) \equiv Bel(\mathcal{K} * \psi)$*

**Proof:** ($\mathcal{K} * 1$) follows directly from Definition 4. For ($\mathcal{K} * 2a$), it follows from Definition 1 that if $0 < Pr_\rho(\phi) \leq 1$ then $Pr_\rho(\phi) < Pr_{\rho(\phi,q)}(\phi) \leq 1.0$, and so if we iterate a revision by $\phi$, the probability of $\phi$ monotonically increases, with upper bound 1.0. It follows that for some $n$, $Mod(\mathcal{K} *^n \phi) \subseteq Mod(\phi)$, and so for some $n > 0$, $\phi \in Bel(\mathcal{K} *^n \phi)$. ($\mathcal{K} * 2b$) and ($\mathcal{K} * 2c$) have prerequisite condition that $\mathcal{K}$ is not the incoherent epistemic state. ($\mathcal{K} * 2b$) is obvious. For ($\mathcal{K} * 2c$), if $\mathcal{K} \neq \mathcal{K}_\perp$ then if there are no $\phi$-worlds with non-zero probability, then Definition 1 can be seen to leave the probability function unchanged. For ($\mathcal{K} * 5$) it can be seen from the definitions that if $\mathcal{K} \neq \mathcal{K}_\perp$, then there will be worlds with a non-zero probability, and so $Mod(\mathcal{K}) \neq \emptyset$ in Definition 3, and so $Bel(\mathcal{K})$ is well defined in Definition 4 and specifically $Bel(\mathcal{K}) \neq \mathcal{L}$. In particular, in the case of a revision by an inconsistent formula $\phi$, we get that $\mathcal{K} * \phi = \mathcal{K}$: All $\phi$ worlds (of which there are none) share in the probability $q$, and all $\neg\phi$ worlds share in the probability $1 - q$. The result is normalised, leaving the probabilities unchanged. If $\mathcal{K} = \mathcal{K}_\perp$, then in Definition 1 we get that $\eta \neq 0$, $0 < q < 1$, and so the probability assignment $\rho(\phi, q) \neq \rho_\perp$, and so in Definition 3 we obtain that $Mod(\mathcal{K}) \neq \emptyset$. Postulate ($\mathcal{K} * 6$) holds trivially, but by virtue of the fact that the reliability of an observation of $\phi$ is the same as that of $\psi$. ■

The weaker version of postulate ($\mathcal{K} * 2$), given by ($\mathcal{K} * 2a$), means that an agent will accept that $\phi$ is true after a sufficient number of iterations (or "reports") of $\phi$. Hence, despite the absence of other AGM postulates, the operator $*$ counts as a revision operator, since the formula $\phi$ will eventually be accepted, provided that it is possible. Note that if a formula $\phi$ is not possible, then from our earlier (non-revision) result

If not $Poss_{\mathcal{K}}(\phi)$ then $\neg\phi \in Bel(\mathcal{K})$

together with ($\mathcal{K} * 5$), we have that $\phi$ can never be accepted. As well, if $\phi$ is accepted, it will continue to be accepted following revisions by $\phi$ ($\mathcal{K} * 2b$). This last point would seem to be obvious, but is necessitated by the absence of a postulate of success and the absence of $(\mathcal{K} * 4)$. If a formula $\phi$ is deemed to be not possible, but the agent is not in the incoherent state $\mathcal{K}_\perp$, $(\mathcal{K} * 2c)$ shows that revising by $\phi$ leaves the agent's belief set unchanged. While this may seem noncontentious, [Makinson 2011] discusses the case where it may be meaningful to condition on a contingent formula whose probability is zero. However, such a situation appears to rely on an underlying infinite domain.

It can be noted that $K_\perp$ plays no interesting role in revisions; this is reflected by $(\mathcal{K} * 5)$, which asserts that no revision can yield $\mathcal{K}_\perp$. Hence an epistemic state can be inconsistent only if the original assignment of probabilities to worlds is the absurd probability assignment $\rho_\perp$. Any subsequent revision will have $Bel(\mathcal{K}_\perp * \phi) \neq \mathcal{L}$. In particular if $\phi$ is $\perp$ then $\rho_\perp(\phi, q) = \rho_\top$ and so $Bel(\mathcal{K}_\perp * \perp) = \mathcal{C}n(\top)$.

Postulate $(\mathcal{K}*5)$ is quite strong, in that it imposes no conditions on the original epistemic state $\mathcal{K}$ or the formula for revision. If one begins with the inconsistent epistemic state $\mathcal{K}_\perp$, then revision is defined as being the same as a revision of the epistemic state of complete ignorance $\rho_\top$. This is pragmatically useful: from $\mathcal{K}_\perp$, if one revises by a formula $\phi$ where $Pr(\phi) \neq 0$, then analogous to the AGM approach, one arrives at a consistent belief state. This also goes beyond the AGM postulate $(K * 5)$, since if $\phi$ is held to be impossible (i.e. there are no worlds with nonzero probability in which $\phi$ is true), then there will be worlds in which $\neg\phi$ is true and with nonzero probability, and so revision yields meaningful results, in particular yielding the epistemic state with probability function $\rho_\top$.

Postulate $(\mathcal{K}*6)$ holds trivially, given the assumption that the reliability of an observation of $\phi$ is the same as that of $\psi$. This assumption is, of course, limiting, and in the case where observations may be made with differing degrees of reliability, the postulate would not hold. It can be noted that in the case where observations may be made with differing degrees of reliability, the postulate can be replaced by the weaker version:

If $\phi \equiv \psi$ then $Bel(\mathcal{K} * \phi) \subseteq Bel(\mathcal{K} * \psi)$ or $Bel(\mathcal{K} * \psi) \subseteq Bel(\mathcal{K} * \phi)$.

We next consider those postulates that don't hold, and why they fail to hold. For a counterexample for $(\mathcal{K} * 3)$, let $\mathcal{P} = \{a, b\}$, $\mathcal{K} = (\rho, 0.97)$, and $\rho$ is given as follows:

$$\begin{array}{ll} \rho(\{a,b\}) = .96 & \rho(\{a,\neg b\}) = .02, \\ \rho(\{\neg a,b\}) = .01 & \rho(\{\neg a,\neg b\}) = .01 \end{array}$$

$Bel(\mathcal{K})$ is characterised by $\{a, b\}$, $\{a, \neg b\}$, $Bel(\mathcal{K}) = \mathcal{C}n(a)$, and so $Bel(\mathcal{K})+a = \mathcal{C}n(a)$. If we revise by $a$ with confidence .8, we get

$$\begin{array}{ll} \rho(a,.9)(\{a,b\}) \approx .9746 & \rho(a,.9)(\{a,\neg b\}) \approx .0203 \\ \rho(a,.9)(\{\neg a,b\}) \approx .0025 & \rho(a,.9)(\{\neg a,\neg b\}) \approx .0025 \end{array}$$

Thus $Bel(\mathcal{K} * a)$ is characterised by $\{a, b\}$, i.e. $Bel(\mathcal{K} * a) = \mathcal{C}n(a \wedge b) \neq \mathcal{C}n(a) = Bel(\mathcal{K}) + a$. This illustrates a curious phenomenon: In the counterexample we have that $Bel(\mathcal{K}) = \mathcal{C}n(a)$ yet $Bel(\mathcal{K} * a) = \mathcal{C}n(a \wedge b)$. In revising by $a$, the probability of worlds given by $\{a, b\}$, $\{a, \neg b\}$ both increase, but that of $\{a, b\}$ increases so that its probability exceeds the confidence level $c$, and so it alone characterises the agent's set of accepted beliefs. We discuss this behaviour later, once we have presented the approach as a whole.

For $(\mathcal{K} * 4)$, it is possible to have formulas $\phi$ and $\psi$ such that $\phi$ and $\psi$ are logically independent, $Bel(\mathcal{K}) = \mathcal{C}n(\phi)$ and $Bel(\mathcal{K} * \psi) = \mathcal{C}n(\psi)$, thus contradicting the postulate. To see this, consider where $\mathcal{P} = \{a, b\}$, and $\mathcal{K} = (\rho, 0.9)$ and where:

$$\begin{array}{ll} \rho(\{a,b\}) = .46 & \rho(\{a,\neg b\}) = .46, \\ \rho(\{\neg a,b\}) = .06 & \rho(\{\neg a,\neg b\}) = .02 \end{array}$$

$Bel(\mathcal{K})$ is characterised by $\{a, b\}$, $\{a, \neg b\}$, i.e. $Bel(\mathcal{K}) = \mathcal{C}n(a)$, and so $Bel(\mathcal{K}) + b = \mathcal{C}n(a \wedge b)$. If we revise by $b$ with confidence .9, we get

$$\begin{array}{ll} \rho(b, .9)(\{a, b\}) \approx .802 & \rho(b, .9)(\{a, \neg b\}) \approx .089 \\ \rho(b, .9)(\{\neg a, b\}) \approx .105 & \rho(b, .9)(\{\neg a, \neg b\}) \approx .004 \end{array}$$

Since $\rho(b, .9)(\{a, b\}) + \rho(b, .9)(\{\neg a, b\}) > .9 = c$, we get that $Bel(\mathcal{K} * b)$ is characterised by $\{a, b\}$, $\{\neg a, b\}$ and so $Bel(\mathcal{K}) = \mathcal{C}n(b)$.

This illustrates another interesting point: not only does the postulate fail but, unlike ($\mathcal{K} * 2$), it may fail over any number of iterations. For the example given, the probability of the world given by $\{a, b\}$ will converge to something just over .88, which is below the given confidence level of $c = 0.9$. Since $\mathcal{C}n(a, b)$ is the result of expansion in the example, this shows that $\mathcal{C}n(a, b)$ will never come to be accepted. Similar remarks hold for ($\mathcal{K} * 8$).

($\mathcal{K} * 7$) doesn't hold for the same reason ($\mathcal{K} * 3$) doesn't. Substituting $\top$ for $\phi$ in ($\mathcal{K} * 7$) in fact yields ($\mathcal{K} * 3$). Similarly, ($\mathcal{K} * 8$) doesn't hold for the same reason that ($\mathcal{K} * 4$) doesn't. Substituting $\top$ for $\phi$ in ($\mathcal{K} * 8$) yields ($\mathcal{K} * 4$).

### 5.2 Variants of the Approach

We next examine three variants of the approach. In the first, observations are made with certainty. This variation coincides with an extant approach in belief revision. As well it has close relations to Gärdenfors' revision of probability functions; a discussion of the relation with this latter work is deferred to the next section. In the second variant, observations are made with near certainty; again this variant corresponds with an extant approach in belief revision. In the last variant, informally possible worlds that characterise an agent's beliefs are retained after a revision if there is no reason to eliminate them.

**Certain Observations** Consider where observations are certain, and so the (binary) revision $\mathcal{K} * \phi$ corresponds to $\mathcal{K} * (\phi, 1.0)$. Clearly, if $\rho(w) = 0$, then $\rho(\phi, 1.0)(w) = 0$ for any $\phi$; that is, if a world had probability 0, then no observation is going to alter this probability. As well, if $w \models \neg\phi$ then $\rho(\phi, 1.0)(w) = 0$. So in a revision by $\phi$ with certainty, any $\neg\phi$ world will receive probability 0, and by the previous observation, this probability of 0 will remain fixed after subsequent revisions.

Thus in this case, revision is analogous to a form of *expansion*, but with respect to the epistemic state $\mathcal{K}$. So following a revision by $\phi$, all $\neg\phi$ worlds are discarded from the derived faithful ranking. This corresponds to revision in the approach of [Shapiro, Pagnucco, Lespérance, and Levesque 2011], where their account of revision is embedded in an account of reasoning about action. For their approach, a plausibility ordering over worlds is given at each world. Observations are assumed to be correct; thus an observation of $\phi$ means that $\neg\phi$ is impossible in the current world, and so all $\neg\phi$ worlds are discarded. This also means that an observation of $\phi$ followed by $\neg\phi$ yields the inconsistent epistemic state. This result can be justified by the argument that, if $\phi$ is observed with certainty, then if the world does not change, then it is *impossible* for $\neg\phi$ to be observed. In this approach, postulates $(K * 1) -$ (K*4), and $(K * 6)$ are shown to be satisfied.

**Near-Certain Observations** Consider where the binary revision $\mathcal{K} * \phi$ is defined to be $\mathcal{K} * (\phi, 1.0 - \epsilon)$, where $\epsilon$ is "sufficiently small" compared to the probabilities assigned by $\rho$. If the minimum and maximum values in the range of $\rho$ are $min_\rho$ and $max_\rho$, then "sufficiently small" would mean that

$$\max\{\rho(w) \mid w \in Mod(\neg\phi)\} \times \epsilon \ < \ \min\{\rho(w) \mid w \in Mod(\phi)\} \times (1.0 - \epsilon).$$

Thus for $\rho' = \rho(\phi, 1.0 - \epsilon)$ we would have for $w \models \phi$, $w' \not\models \phi$ that $\rho'(w) > \rho'(w')$. This yields *lexicographic revision* [Nayak 1994] in which, in revising by $\phi$, every $\phi$ world is ranked below every $\neg\phi$ world, but the relative ranking of $\phi$ worlds (resp. $\neg\phi$ worlds) is retained. In this approach, all AGM revision postulates hold.

**Retaining Confirmed Possible Worlds** The present approach clearly falls within belief revision, since under reasonable conditions a formula will become accepted. However, it has notable weaknesses compared to the AGM approach; in particular the postulates $(\mathcal{K} * 3)$, $(\mathcal{K} * 4)$, $(\mathcal{K} * 7)$, and $(\mathcal{K} * 8)$ all fail. In the case of $(\mathcal{K} * 4)$ and $(\mathcal{K} * 8)$ this seems unavoidable. However, an examination of $(\mathcal{K} * 3)$ and $(\mathcal{K} * 7)$ shows a curious phenomenon. Consider $(\mathcal{K} * 3)$: In the counterexample presented, the agent believed that the real world was among the set of worlds $\{\{a, b\}, \{a, \neg b\}\}$. On revising by $a$, the agent believed that the real world was among the set of worlds $\{\{a, b\}\}$, which is to say, that $\{a, b\}$ was the real world. But this means is that $\{a, \neg b\}$ was considered to be possibly be the real world according to the agent, but on receiving confirmatory evidence (viz. revision by $a$), this world was dropped from the characterising set. But arguably if $w$ may be the actual world according to the agent, and the agent learns $\phi$ where $w \models \phi$, then it seems that the agent should still consider $w$ as possibly being the actual world.

The reason for this phenomenon is clear: The probability of other worlds (in the example, given by $\{a, b\}$) becomes large enough following revision so that the "dropped" world isn't required in making up $Mod(\mathcal{K})$. To counteract this phenomenon, it seems reasonable to assume that if an agent considers a world to be possible, then it remains possible after confirmatory evidence. To this end, the approach can be modified so that one keeps track of worlds considered possible by the agent, where these are the worlds characterising the agent's contingent beliefs. An epistemic state now would be a triple $(\rho, c, B)$ where $B \subseteq W_{\mathcal{L}}$ characterises the agent's belief set following a revision by $\phi$ with probability $q$. Thus after revising by $\phi$, the new value of $B$ would be given by:

$$Mod((\rho(\phi, q), c, B)) \;\cup\; (Mod((\rho, c.B)) \cap Mod(\phi)).$$

In this case postulates $(K * 3)$ and $(K * 7)$ also hold.

### 5.3 Iterated Belief Revision

Turning to iterated revision, it proves to be the case that three of the Darwiche-Pearl postulates fail to hold. However, the reason that these postulates do not hold is not a result of the probabilistic approach per se, but rather is a result of the expression of a belief set in terms of possible worlds.

THEOREM 10. *Let $\mathcal{K}$ be an epistemic state with associated revision operator $*$. Then $\mathcal{K}$ satisfies* **C3**.

**Proof:** For **C3**, we obtain that the semantic condition **CR3** holds: If $w_1 \models \phi$ and $w_2 \not\models \phi$, then $w_1 \prec_{\mathcal{K}} w_2$ implies that $\rho(w_1) < \rho(w_2)$ from which it follows that $\rho(\phi, q)(w_1) < \rho(\phi, q)(w_2)$ and so $w_1 \prec_{\mathcal{K} * \phi} w_2$. By the same argument as [Darwiche and Pearl 1997, Theorem 13], we get that **C3** is satisfied. ■

$\mathcal{K}$ does not necessarily satisfy **C1**, **C3**, and **C4**. Consider **C1**, and let $\mathcal{P} = \{a, b\}$, $\mathcal{K} = (\rho, 0.9)$, and where:

$$\rho(\{a, b\}) = .85 \qquad \rho(\{a, \neg b\}) = .06,$$
$$\rho(\{\neg a, b\}) = .05 \qquad \rho(\{\neg a, \neg b\}) = .04$$

$Bel(\mathcal{K})$ is characterised by $\{a, b\}$, $\{a, \neg b\}$ and so $Bel(\mathcal{K}) = \mathcal{C}n(a)$. If we revise by $a$ with confidence .7, we get

$$\begin{array}{ll} \rho(a, .7)(\{a, b\}) \approx .891 & \rho(a, .7)(\{a, \neg b\}) \approx .063 \\ \rho(a, .7)(\{\neg a, b\}) \approx .023 & \rho(a, .7)(\{\neg a, \neg b\}) \approx .018 \end{array}$$

Thus $Bel(\mathcal{K}) = Bel(\mathcal{K} * a)$. If we again revise by $a$ with confidence .7, we get

$$\begin{array}{ll} \rho(a, .7)(\{a, b\}) \approx .917 & \rho(a, .7)(\{a, \neg b\}) \approx .065 \\ \rho(a, .7)(\{\neg a, b\}) \approx .010 & \rho(a, .7)(\{\neg a, \neg b\}) \approx .008 \end{array}$$

Since $\rho(a, .9)(\{a, b\}) > .9 = c$, so $Mod(\mathcal{K} * a * a) = \{a, b\}$. Hence $Bel(\mathcal{K} * a) \neq Bel(\mathcal{K} * a * a)$, thereby violating **C1**.[4] Other postulates fail for analogous reasons.

It is worth considering why most of the iteration postulates fail. Interestingly, for the semantic conditions, **CR1** – **CR4**, if expressions of the form $w_1 \prec_{\mathcal{K}} w_2$ are replaced by expressions of the form $\rho(w_1) \leq \rho(w_2)$, then the modified conditions hold in the current approach. That is, it is easily verified that all of the following hold:

THEOREM 11.

**PCR1** *If* $w_1, w_2 \models \phi$, *then* $\rho(w_1) \leq \rho(w_2)$ *iff* $\rho(\phi, q)(w_1) \leq \rho(\phi, q)(w_2)$.

**PCR2** *If* $w_1, w_2 \not\models \phi$, *then* $\rho(w_1) \leq \rho(w_2)$ *iff* $\rho(\phi, q)(w_1) \leq \rho(\phi, q)(w_2)$.

**PCR3** *If* $w_1 \models \phi$ *and* $w_2 \not\models \phi$, *then* $\rho(w_1) < \rho(w_2)$ *implies* $\rho(\phi, q)(w_1) < \rho(\phi, q)(w_2)$.

**PCR4** *If* $w_1 \models \phi$ *and* $w_2 \not\models \phi$, *then* $\rho(w_1) \leq \rho(w_2)$ *implies* $\rho(\phi, q)(w_1) \leq \rho(\phi, q)(w_2)$.

**Proof:** Straightforward from Definition 1. ■

The problem is that our faithful ranking (Definition 7) doesn't preserve the ordering given by $\rho$. In particular, if $w_1$, $w_2 \in Mod(\mathcal{K})$ then $w_1 =_{\mathcal{K}} w_2$ in the derived faithful ranking, while most often we will have $\rho(w_1) \neq \rho(w_2)$. Essentially, in moving from values assigned via $\rho$ to the faithful ranking, gradations (given by probabilities) among worlds in $Mod(\mathcal{K})$ are lost. That is, in a sense, the probabilistic approach provides a finer-grained account of an epistemic state than is given by a faithful ranking on worlds, in that models of the agent's belief set also come with gradations of belief.

### 5.4 Relation with Other Work

**Other Approaches to Revision** We have already discussed the relation of the approach to [Shapiro, Pagnucco, Lespérance, and Levesque 2011] and [Nayak 1994].

The work in belief change that is closest to that described here is that of *improvement operators* [Konieczny and Pino Pérez 2008], where an an improvement operator is a belief change operator in which new information isn't necessarily immediately accepted, but where the plausibility is increased. Thus after a sufficient number of iterations, the information will come to be believed. The general idea of this approach then is similar to the present approach. As well, in both approaches the success postulate does not necessarily hold, so new information is not necessarily immediately accepted. However, beyond failure of the success postulate, the approaches have quite different characteristics.

[4] In terms of **CR1**, we have $\{a, b\} =_{\mathcal{K}*a} \{a, \neg b\}$ but $\{a, b\} \prec_{\mathcal{K}*a*a} \{a, \neg b\}$, thereby violating **CR1**.

In the postulate set following,[5] $\circ$ is an improvement operator, and $\times$ is defined by: $\mathcal{K} \times \phi = \mathcal{K} \circ^n \phi$ where $n$ is the first integer such that $\phi \in Bel(\mathcal{K} \circ^n \phi)$.

**(I1)** There exists $n$ such that $\phi \in Bel(\mathcal{K} \circ^n \phi)$

**(I2)** If $\neg\phi \notin Bel(\mathcal{K})$ then $Bel(\mathcal{K} \times \phi) \equiv Bel(\mathcal{K}) + \phi$

**(I3)** $Bel(\mathcal{K} \circ \phi)$ is inconsistent, only if $\nvdash \neg\phi$

**(I4)** If $\phi_i \equiv \psi_i$ for $1 \leq i \leq n$ then $Bel(\mathcal{K} \circ \phi_1 \circ \ldots \circ \phi_n) \;\equiv\; Bel(\mathcal{K} \circ \psi_1 \circ \ldots \circ \psi_n)$

**(I5)** $Bel(\mathcal{K} \times (\phi \wedge \psi)) \subseteq Bel(\mathcal{K} \times \phi) + \psi$

**(I6)** If $\neg\psi \notin Bel(\mathcal{K} \times \phi)$ then $Bel(\mathcal{K} \times \phi) + \psi \subseteq Bel(\mathcal{K} \times (\phi \wedge \psi))$

To show the approaches are independent, it suffices to compare ($\mathcal{K}*3$)/($\mathcal{K}*4$) with **(I2)**. According to **(I2)**, after some number of iterations of an improvement operator, the resulting belief set will correspond to expansion of the original belief set by the formula in question. However, there are cases in which neither $(K * 3)$ nor $(\mathcal{K} * 4)$ are satisfied regardless of the number of iterations. Similar comments apply to $(\mathcal{K} * 7)$ and $(\mathcal{K} * 8)$, and **(I5)** and **(I6)**, respectively. The need for the extended postulate for irrelevance of syntax for epistemic states **I4** was noted in [Booth and Meyer 2006]. In the present approach $(\mathcal{K} * 6)$ suffices.

**Other Related Work** As discussed in Section 2, earlier work dealing specifically with revision and probability has been concerned with revising probability functions. Thus, [Gärdenfors 1988; Lindström and Rabinowicz 1989; Boutilier 1995] deal with extensions to the AGM approach for revising probability functions. In these approaches there is a probability function associated with possible worlds, but where the agent's belief set is characterised by worlds with probability 1.0. For a revision $K * \phi$, $\phi$ represents new *evidence*, and the probability function is revised by $\phi$. The belief set corresponding to $K * \phi$ then is the set of propositions with probability 1.0. In contrast, in the approach at hand, an agent's *accepted* beliefs are characterised by a set of possible worlds whose overall probability in the general case will be less than 1.0. In a sense then there is finer granularity with regards the present approach, since the worlds characterising a belief set may have varying probability. As well, for us if a formula $\phi$ has probability of 1.0, then it cannot be removed by subsequent revisions; a formula is accepted as true if its probability is sufficiently high, although it may potentially be revised away. This arguably confirms to intuitions, in that if a formula is held with complete certainty then it *should* be immune from revisions.

It was noted that [Bacchus, Halpern, and Levesque 1999] presents the same general setting in which an agent receives possibly-unreliable observations. However, the concern in this paper was to update probabilities associated with worlds and then to use this for reasoning about dynamic domains expressed via the situation calculus. The approach at hand employs the same method for updating probabilities but addresses the question of how this may be regarded as, or used to formulate, an approach to belief revision. The present approach also has finer granularity, in that in [Bacchus, Halpern, and Levesque 1999] non-beliefs are given by worlds with probability 0; in the approach at hand, non-beliefs are those that fall outside the set of accepted beliefs, and may have non-zero probability. Again, arguably the present approach conforms to intuitions, since if a formula is held to be impossible then it seems it should forever remain outside the realm of revision.

[5] [Konieczny and Pino Pérez 2008] follow [Katsuno and Mendelzon 1991], where the result of revision is a formula, not a belief set. We rephrased the [Konieczny and Pino Pérez 2008] postulates in terms of belief sets. In particular, ($\mathcal{K}*3$) and ($\mathcal{K}*4$) correspond to **(I2)**, while ($\mathcal{K}*7$) and ($\mathcal{K}*8$) correspond to **(I5)** and **(I6)** respectively.

## 6 Conclusion

We have explored an approach to beliefs and belief revision, based on an underlying model of uncertain reasoning. With few exceptions, research in belief revision has dealt with categorical information in which an agent has some set of beliefs and the goal is to incorporate a formula into this set of beliefs. However, most information about the real world is not categorical, and arguably no non-tautological belief may be held with complete certainty. To accommodate this, one alternative is to adopt a purely probabilistic framework for belief change. However, such a framework ignores the fact that an agent may well *accept* a formula as being true, even if this acceptance is tentative, or hedged in some fashion. So another alternative, and the one followed here, is to begin with a probabilistic framework, but also define a set of formulas that the agent accepts. Revision can then be defined in this framework, and the effect of revision on the agent's accepted beliefs examined.

To this end we assumed that an agent receives uncertain information as input, and the agent's probabilities on possible worlds are updated via Bayesian conditioning. A set of formulas among the (noncategorical) beliefs is then identified as the agent's (categorical) belief set. We show that a subset of the AGM belief revision postulates are satisfied by this approach. Most significantly, though not surprisingly, the success postulate is not guaranteed to hold, though it is after a sufficient number of iterations. As well, it proves to be the case that in revising by a formula consistent with the agent's beliefs, revision does not necessarily correspond to expansion. As another point of interest, of the postulates for iterated revision that we considered, only **C3** holds. This is because, even though the updating of the probability assignment $\rho$ satisfies all of the corresponding semantic conditions, the induced faithful ordering $\prec_{\mathcal{K}}$ does not. Last, although the approach shares motivation and intuitions with improvement operators, these approaches have different properties.

There are two ways that these results may be viewed with respect to classical AGM-style belief revision. On the one hand, it could be suggested that the current approach simply provides a revision operator that is substantially weaker than given in the AGM approach and approaches to iterated revision. On the other hand, the AGM approach and approaches to iterated revision have been justified by appeals to rationality, in that it is claimed that *any* rational agent should conform to the AGM postulates and, say, the Darwiche/Pearl iteration postulates. Thus, to the extent that the presented approach is rational, this would appear to undermine the rationale of these approaches, at least in the case of uncertain information.

## References


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

Bacchus, F., J. Halpern, and H. Levesque [1999]. Reasoning about noisy sensors and effectors in the situation calculus. *Artificial Intelligence 111*(1-2), 171–208.

Booth, R. and T. A. Meyer [2006]. Admissible and restrained revision. *Journal of Artificial Intelligence Research 26*, 127–151.

Boutilier, C. [1995]. On the revision of probabilistic belief states. *Notre Dame Journal of Formal Logic 36*(1), 158–183.

Boutilier, C. [1996]. Iterated revision and minimal revision of conditional beliefs. *Journal of Logic and Computation 25*, 262–305.

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

Gärdenfors, P. [1988]. *Knowledge in Flux: Modelling the Dynamics of Epistemic States*. Cambridge, MA: The MIT Press.

Hughes, G. and M. Cresswell [1996]. *A New Introduction to Modal Logic*. London and New York: Routledge.

Jeffrey, R. [1983]. *The Logic of Decision* (second ed.). University of Chicago Press.

Jin, Y. and M. Thielscher [2007]. Iterated belief revision, revised. *Artificial Intelligence 171*(1), 1–18.

Katsuno, H. and A. Mendelzon [1991]. Propositional knowledge base revision and minimal change. *Artificial Intelligence 52*(3), 263–294.

Konieczny, S. and R. Pino Pérez [2008]. Improvement operators. In G. Brewka and J. Lang (Eds.), *Proceedings of the Eleventh International Conference on the Principles of Knowledge Representation and Reasoning*, Sydney, Australia, pp. 177–186. AAAI Press.

Kyburg, Jr., H. [1961]. *Probability and the Logic of Rational Belief*. Wesleyan University Press.

Kyburg, Jr., H. [1994]. Believing on the basis of evidence. *Computational Intelligence 10*(1), 3–20.

Levesque, H., F. Pirri, and R. Reiter [1998]. Foundations for the situation calculus. *Linköping Electronic Articles in Computer and Information Science 3*(18).

Lindström, S. and W. Rabinowicz [1989]. On probabilistic representation of non-probabilistic belief revision. *Journal of Philosophical Logic 11*(1), 69–101.

Makinson, D. [2011]. Conditional probability in the light of qualitative belief change. *Journal of Philosophical Logic*. To appear.

Nayak, A. [1994]. Iterated belief change based on epistemic entrenchment. *Erkenntnis 41*, 353–390.

Nayak, A. C., M. Pagnucco, and P. Peppas [2003]. Dynamic belief revision operators. *Artificial Intelligence 146*(2), 193–228.

Pearl, J. [1988]. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufman.

Peppas, P. [2007]. Belief revision. In F. van Harmelen, V. Lifschitz, and B. Porter (Eds.), *Handbook of Knowledge Representation*, pp. 317–359. San Diego, USA: Elsevier Science.

Shapiro, S., M. Pagnucco, Y. Lespérance, and H. J. Levesque [2011]. Iterated belief change in the situation calculus. *Artificial Intelligence 175*(1), 165–192.