A Unified Model of Qualitative Belief Change: A Dynamical Systems Perspective*

Craig Boutilier
Department of Computer Science
University of British Columbia
Vancouver, British Columbia
CANADA, V6T 1Z4
email: *cebly@cs.ubc.ca*
URL: *http://www.cs.ubc.ca/spider/cebly/craig.html*

Abstract

Belief revision and belief update have been proposed as two types of belief change serving different purposes, revision intended to capture changes in belief state reflecting new information about a static world, and update intended to capture changes of belief in response to a changing world. We argue that routine belief change involves elements of both and present a model of *generalized update* that allows updates in response to external changes to inform an agent about its prior beliefs. This model of update combines aspects of revision and update, providing a more realistic characterization of belief change. We show that, under certain assumptions, the original update postulates are satisfied. We also demonstrate that plain revision and plain update are special cases of our model. We also draw parallels to models of stochastic dynamical systems, and use this to develop a model that deals with iterated update and noisy observations in (qualitative settings) that is analogous to Bayesian updating in a quantitative setting.

*Some parts of this report appeared in preliminary form in "Generalized Update: Belief Change in Dynamic Settings," *Proc. of Fourteenth International Joint Conf. on Artificial Intelligence (IJCAI-95)*, Montreal, pp.1550–1556 (1995).

1 Introduction

An underlying premise in much work addressing the design of intelligent agents or programs is that such agents should (either implicitly or explicitly) hold beliefs about the true state of the world. Typically, these beliefs are incomplete, for there is much an agent will not know about its environment. In realistic settings one must also expect an agent's beliefs to be incorrect from time to time. If an agent is in a position to make observations and detect such errors, a mechanism is required whereby the agent can change its beliefs to incorporate new information. Finally, an agent that finds itself in a dynamic, evolving environment (including evolution brought about by its own actions) will be required to change its beliefs about the environment as the environment evolves.

Theories of belief change have received considerable attention in recent years in the AI community, as well as other areas such as philosophy and database systems. One crucial distinction that has come to light in this work is that between *belief revision* and *belief update*. The distinction can be best understood as one pertaining to the source of incorrect beliefs. On the one hand, an agent's beliefs about the world may simply be mistaken or incomplete, for instance, in the case where it adopts some default belief. If an agent observes that this belief is mistaken, it must take steps to correct the misconception. Such a process is know as belief revision, of which the theory of Alchourrón, Gärdenfors and Makinson [2, 15] is the best-known characterization. On the other hand, an agent's beliefs, while correct at one time, may have become inaccurate due to changes in the world. As events occur and other agents act, or as the agent itself takes actions, certain facts become true and others false. An agent observing such processes or their results must take steps to ensure its state of belief reflects these changes. This process is known as belief update, as proposed by Winslett [33] and Katsuno and Mendelzon [21].

On the surface, formalizations of revision and update are quite similar: in both cases, the objective is to define a function that, given the agent's belief state and an "observed" proposition, returns a new belief state. However, conceptually these two processes have been treated distinctly, and the axioms and semantic models proposed to capture revision and update are, for the most part, incompatible—that is, we cannot treat update as a form of revision, nor can we treat revision as a form of update. The properties of these processes are, we shall argue, fundamentally different.

One difficulty with the separation of revision and update is the fact that routine belief change, that is the change of an agent's belief state in response to some observation, typically involves elements of both. We will support below the claim that a given observation often calls for belief change that reflects a response to changes in the world as well as incorrect or incomplete prior beliefs. In this paper, we describe a semantic model for belief change that unifies the two types of belief change. In particular, we generalize classical belief update to incorporate aspects of belief revision. The aim of this model is twofold. First, we provide a unifying and natural semantics for both revision and update that highlights the orthogonal roles both have to play in routine belief change. Second, we attempt to provide a more compelling account of belief update to deal with observations of changes in the world that provide information about the prior world state (i.e., about the agent's prior beliefs). This second objective is a response to difficulties with the classical view of update, which we outline below.

The result of this union is a more robust and realistic notion of update in which observations of change can inform an agent's prior beliefs and expectations. Such observations are pervasive; consider the following example. A warehouse control agent believes it is snowing on Route 1 after yesterday's weather forecast, and expects the arrival of a number of trucks to be delayed. Now suppose a certain truck arrives, causing the agent to update its beliefs; furthermore, contrary to its expectations, the truck arrives on time. There are two possible explanations: either the truck was able to speed through the snow or it did not snow after all. If the latter explanation is more plausible, current update theories cannot arrive at the desired update in a natural way. The observation of the change in the world's state (arrival of the truck) indicates that the agent's prior beliefs (e.g., that it is snowing) were wrong. The update should not simply involve changes that reflect the evolution of the world, but should place these changes in the context of the corrected or *revised* prior beliefs. The agent should revise its beliefs to capture the fact that it is did not snow and adjust its expectations regarding the arrival of other trucks accordingly. Routine belief changes often involve aspects of revision (correcting or augmenting one's beliefs) and update (allowing beliefs about the world to "evolve").

The general model we present to capture such considerations takes as a starting point the notion of *ranked* or structured belief sets. By ranking situations according to their degree of plausibility, we obtain a natural way of assessing degrees of belief and a very natural semantics for belief revision. Such models have been used extensively for revision [20, 15, 6]. To this we add the notion of a *transition* or evolution from one world state to another. As proposed by Katsuno and Mendelzon (KM), updates reflect changes in the world, and transitions can be used to model such changes. However, in contrast to the KM model and following our earlier work [8], we assume that the relative plausibility of transitions (and hence possible updates) is not something that is judged directly; rather we assume that *events* or *actions* provide the impetus for change. The plausibility of a transition is a function of: (a) the plausibility of possible causing events; and (b) the likelihood of that event having the specified outcome. In this way, we can model events or actions that have defeasible effects (which can be judged as more or less likely).

Finally, in response to an observation, an agent attempts to *explain* the observation by postulating conditions under which that observation is expected. An explanation consists of three components: initial conditions, an event (or action), and an outcome of that event. The key aspect of our model is the ranking of such explanations — an explanation is more or less plausible depending on the plausibility of the initial conditions, the plausibility of the event *given* that starting point, and the plausibility of the event's outcome. The belief change that results provides the essence of the *generalized update* (GU) operator: an agent believes the consequences of the most plausible explanations of the observation.

Unlike other theories of update, our model allows an agent to trade off the likelihood of possible events, outcomes and prior beliefs in coming up with plausible explanations of an observation. Of course, by allowing prior beliefs to be "changed" during update, we are essentially folding belief revision into the update process (as we elaborate below). We thus generalize the KM update model to work on structured (rather than flat) belief sets. Furthermore, the information required to generate such explanations is very natural and readily available. We are much more willing to judge the relative plausibility of events and their outcomes than the plausibility of transitions directly. The resulting change in belief, consisting of the consequences of the explanation, is very intuitive.

In Section 2 we present the AGM theory of revision and the KM theory of update, focusing primarily on the semantic models that have been proposed. In our presentation, we adopt the qualitative probabilistic model of [31, 18, 19]. In Section 3 we present our model of generalized update, with an emphasis on semantics, and contrast it with the "flat" KM model. We describe two examples to illustrate the key features of the model.

In Section 4 we analyze the GU operator in detail. We describe the formal relationship between revision, update and GU. We show that under certain assumptions GU satisfies the KM postulates, though we argue that these assumptions are not appropriate in many settings (thus calling into question the generality of the KM postulates). In addition we show that both "flat" KM update and AGM revision are special cases of GU. In particular, the connection formally verifies the intuition that AGM revision is due to changes in belief about a static world, while update reflects belief change about an evolving world.

In Section 5, we briefly discuss the importance of iterated revision in this model, and emphasize connections between GU and Bayesian update in stochastic dynamical systems. We also discuss the role of observations and weight of evidence, and present a model (as well as several alternative suggestions) for dealing with "noisy" observations in belief revision. This is one area of belief revision and belief update that has received virtually no attention.

There have been attempts to provide general semantics for belief change operators (e.g., [12]); but often these models are such that under *certain assumptions* the change is a revision and under others it is an update. In Section 6 we compare some of these related models to GU. We conclude in Section 7 with some directions for future research. Proofs of the main results are found in the appendix.

2 Classical Belief Revision and Belief Update

In this section we review, in turn, the AGM and KM theories of belief change. We present both the syntactic postulates and the semantic models that characterize these theories and describe briefly the κ-calculus of [31, 18, 19], which provides an alternative model for the ordering relationships used by both theories.

Throughout, we assume that an agent has a deductively closed *belief set* K, a set of sentences drawn from some logical language reflecting the agent's beliefs about the current state of the world. For ease of presentation, we assume a logically finite, classical propositional language, denoted $\mathbf{L}_{CPL}$, and consequence operation *Cn*.[1] The belief set K will often be generated by some finite knowledge base *KB* (i.e., $K = Cn(KB)$). The identically true and false propositions are denoted $\top$ and $\bot$, respectively. Given a set of possible worlds (or valuations over $\mathbf{L}_{CPL}$) W and $A \in \mathbf{L}_{CPL}$, we denote by $\|A\|$ the set of A-worlds, the elements of W satisfying A. The worlds satisfying all sentences in a set K is denoted $\|K\|$.

2.1 Belief Revision

Given a belief set K, an agent will often obtain information A not present in K. In this case, K must be *revised* to incorporate A. If A is consistent with K, one expects A to simply be added to K: we call $K^+_A = Cn(K \cup \{A\})$ the *expansion* of K by A. More problematic is the case when $K \vdash \neg A$; certain beliefs must be given up before A is adopted. The *AGM theory* provides a set of guidelines, in the form of the following postulates, governing this process. We use K^*_A to denote the *revision* of K by A.

(R1) K^*_A is a belief set (i.e., deductively closed).

(R2) $A \in K^*_A$.

(R3) $K^*_A \subseteq K^+_A$.

[1] Languages with a denumerable set of atomic variables, or first order languages pose no special difficulties (e.g., see work on first-order conditional logics).

(R4) If $\neg A \notin K$ then $K_A^+ \subseteq K_A^*$.

(R5) $K_A^* = Cn(\perp)$ iff $\vdash \neg A$.

(R6) If $\models A \equiv B$ then $K_A^* = K_B^*$.

(R7) $K_{A\wedge B}^* \subseteq (K_A^*)_B^+$.

(R8) If $\neg B \notin K_A^*$ then $(K_A^*)_B^+ \subseteq K_{A\wedge B}^*$.

Unfortunately, while the postulates constrain possible revisions, they do not dictate the precise beliefs that should be retracted when A is observed. An alternative model of revision, based on the notion of *epistemic entrenchment* [15], has a more constructive nature. Given a belief set K, we can characterize the revision of K by ordering beliefs according to our willingness to give them up. If one of two beliefs must be retracted in order to accommodate some new fact, the less entrenched belief will be relinquished, while the more entrenched persists.

Semantically, an entrenchment relation (hence a revision function) can be modeled using an ordering on possible worlds reflecting their relative plausibility [20, 6]. However, rather than use a qualitative ranking relation, we adopt the presentation of [31, 18] and rank all possible worlds using a κ-*ranking*. Such a ranking $\kappa : W \rightarrow \mathbf{N}$ assigns to each world a natural number reflecting its plausibility or degree of believability. If $\kappa(w) < \kappa(v)$ then w is more plausible than v or "more consistent" with the agent's beliefs. We insist that $\kappa^{-1}(0) \neq \emptyset$, so that maximally plausible worlds are assigned rank 0. These maximally plausible worlds are exactly those consistent with the agent's beliefs; that is, the epistemically possible worlds according to K are those deemed most plausible in κ (see [31] for further details). We sometimes assume κ is a partial function, and loosely write $\kappa(w) = \infty$ to mean $\kappa(w)$ is not defined (i.e., w is not in the domain of κ, or w is *impossible*).

Rather than modeling an agent's epistemic state with a "flat" unstructured belief set K, we use a κ-ranking to capture objective beliefs K as well as entrenchment information that determines how an agent will revise K. An *epistemic state* κ induces the *(objective) belief set*

$$K = \{A \in \mathbf{L}_{CPL} : \kappa^{-1}(0) \subseteq \|A\|\}$$

In other words, the set of most plausible worlds (those such that $\kappa(w) = 0$) determine the agent's beliefs. The ranking κ also induces a revision function: to revise by A an agent adopts the most plausible

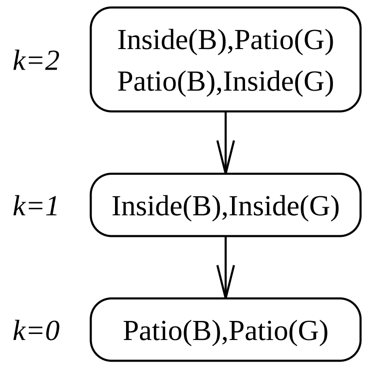


Figure 1: A Revision Model

A-worlds as epistemically possible. Thus, using $\min(A, \kappa)$ to denote this set, we have

$$K_A^* = \{B \in \mathbf{L}_{CPL} : \min(A, \kappa) \subseteq \|B\|\}$$

If $\|A\| \cap W = \emptyset$, we set $\min(A, \kappa) = \emptyset$ and $K_A^* = \mathbf{L}_{CPL}$ (the inconsistent belief set). It is normally assumed that $\|A\| \cap W \neq \emptyset$ for every satisfiable A — thus every proposition is accorded some degree of plausibility. It is well-known that this type of model induces the class of revision functions sanctioned by the AGM postulates [20, 6, 18].[2]

The ranking function κ can naturally be interpreted as characterizing the degree to which an agent is willing to accept certain alternative states of affairs as epistemically possible. As such it seems to be appropriate for modeling changes in belief about an unchanging world. The most plausible A-worlds in our assessment of the *current* state of affairs are adopted when A is observed.

As an example, consider the ranking shown in Figure 1, which reflects the epistemic state of someone who believes her book and glasses are on the patio. If she were to learn that in fact her book is inside, she would also believe her glasses are inside, for the most plausible *Inside*(B)-world ($\kappa = 1$) also satisfies *Inside*(G). This model captures that fact that she strongly believes she left her book and glasses in the same place; that is, the belief *Patio*$(B) \equiv$ *Patio*(G) is more entrenched than either of the beliefs *Patio*(B) or *Patio*(G).

[2]We refer to [4, 6, 12, 19] for a discussion of languages with which one can express properties of belief sets and revision functions. These languages can be used to express belief, degrees of entrenchment and plausibility, conditional belief, and so on.

We can also view κ-rankings as assigning degrees of plausibility to propositions; we define

$$\kappa(A) = \min_{w \models A} \{\kappa(w)\}$$

This can be interpreted as the degree to which proposition A is accepted as plausible (where $\kappa(A) = 0$ means A is maximally plausible, or consistent with the agent's beliefs). We will also have occasion to use the notion of *conditional plausibility*; we define

$$\kappa(B|A) = \kappa(A \wedge B) \perp \kappa(A)$$

Intuitively, this denotes the degree to which B would be considered plausible if A were believed.

These notions are strongly reminiscent of standard concepts from probability theory. In fact, a κ-ranking can be interpreted as a semi-qualitative probability distribution. Using the ε-*semantics* of Adams [1], Goldszmidt and Pearl [17] show how one can interpret the (unconditional and conditional) κ values of propositions as "order of magnitude" probabilities. Under this interpretation, one is able to define analogs of various probabilistic operations, including conditionalization (see Section 5). We do not delve into the details of "κ-arithmetic" here, nor the details of the precise relationship of these ranking functions to probability distributions. We refer to [31, 17, 19] for details. We do note, however, that addition, multiplication and division of probabilities correspond to the minimum, addition and subtraction operations, respectively, for κ-rankings. Thus the definition of $\kappa(B|A)$ above can be seen as a direct counterpart of the usual definition of conditional probability.

Much of the semantics we define below could be reinterpreted in a purely qualitative framework for belief revision (and belief update) in which a simple ordering relation $\leq$ is used to rank possible worlds. However, as will become evident in Sections 4 and 5, much of what we do relies on the expressive power afforded by a quantitative ranking. In particular, our semantics will require that one be able to combine plausibilities that are specified using several distinct rankings. With quantitative κ-rankings, this is straightforward, whereas qualitative rankings do not permit this unless explicit "calibration" information is provided. We elaborate on this in Sections 4 and 5.

2.2 Belief Update

Katsuno and Mendelzon [21] have proposed a general characterization of belief update that seems appropriate when an agent wishes to change its beliefs to reflect changes in, or evolution of, the world. The *KM theory* is also captured by a set of postulates and an equivalent semantic model. Following

[21], we describe update in terms of a knowledge base *KB* rather than a deductively closed belief set K.

If some new fact A is observed in response to some (unspecified) change in the world (i.e., some action or event occurrence), then the formula $KB \diamond A$ denotes the new belief set incorporating this change. The *KM postulates* governing admissible update operators are:

(U1) $KB \diamond A \vdash A$

(U2) If $KB \vdash A$ then $KB \diamond A \;\equiv\; KB$

(U3) If *KB* and A are satisfiable, then $KB \diamond A$ is satisfiable

(U4) If $\vdash A \equiv B$, $KB_1 \equiv KB_2$ then $KB_1 \diamond A \equiv KB_2 \diamond B$

(U5) $(KB \diamond A) \wedge B \vdash KB \diamond (A \wedge B)$

(U6) If $KB \diamond A \vdash B$ and $KB \diamond B \vdash A$ then $KB \diamond A \equiv KB \diamond B$

(U7) If *KB* is complete then $(KB \diamond A) \wedge (KB \diamond B) \vdash KB \diamond (A \vee B)$

(U8) $(KB_1 \vee KB_2) \diamond A \equiv (KB_1 \diamond A) \vee (KB_2 \diamond A)$

The equivalent semantic model of KM sheds more light on the intuitions underlying update. $||KB||$ represents the set of possibilities we are prepared to accept as the actual state of affairs. Since observation A is the result of some change in the actual world, we ought to consider, for each possibility $w \in ||KB||$, the most plausible way (or ways) in which w might have changed in order to make A true. That is, we want to consider the most plausible *evolution* of world w into a world satisfying the observation A. To capture this intuition, Katsuno and Mendelzon propose a family of preorders $\{\leq_w : w \in W\}$, where each $\leq_w$ is a reflexive, transitive relation over W. We interpret each such relation as follows: if $u \leq_w v$ then u is at least as plausible a change relative to w as is v; that is, situation w would more readily evolve into u than it would into v.

Finally, a *faithfulness condition* is imposed: for every world w, the preorder $\leq_w$ has w as a minimum element; that is, $w <_w v$ for all $v \neq w$. Naturally, the most plausible candidate changes in w that result in A are those worlds v satisfying A that are minimal in the relation $\leq_w$. The set of such minimal A-worlds for each relation $\leq_w$, and each $w \in ||KB||$, intuitively capture the situations we ought to accept as possible when updating *KB* with A. In other words,

$$||KB \diamond A|| = \bigcup_{w \in ||KB||} \{\min(A, \leq_w)\}$$

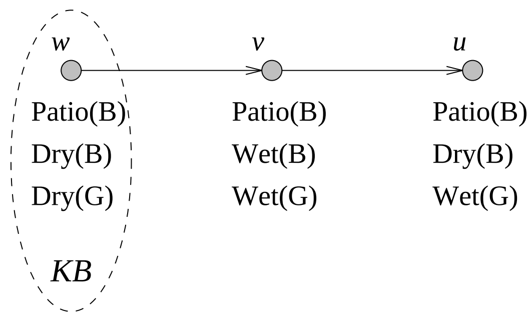


Figure 2: An Update Model

where $\min(A, \leq_w)$ is the set of minimal elements in $\|A\|$ (w.r.t. $\leq_w$).

Update operators determined by any family of preorders $\{\leq_w : w \in W\}$ satisfy the KM postulates. The converse also holds: any KM operator can be represented by such a semantic model. Moreover, if the orderings $\leq_w$ are *total* preorders (so that all elements are comparable), then update operators are characterized by (U1)–(U9) (see [21, 8]):

(U9) If *KB* is complete, $(\mathit{KB} \diamond A) \not\vdash \neg B$ and $(\mathit{KB} \diamond A) \vdash C$ then $(\mathit{KB} \diamond (A \wedge B)) \vdash C$

We assume for the most part that we are dealing with such total update operators (but we discuss this further in Section 4). It should be clear how this (total) model can be recast in terms of κ-rankings: we simply associate a ranking κ_w with each world w (such that $\kappa_w^{-1}(0) = \{w\}$) and use $\min(A, \kappa_w)$ to update by A. Note that the use of κ-rankings requires that the orderings be total.

As a concrete example, suppose that someone observes that the grass in front of her house is wet. Prior to the observation, she believed that she left her book outside on the patio and that the grass and book were dry (see *KB* in Figure 2). As shown in the figure, the most plausible evolution of the epistemically possible world w, given the wet grass, is v; hence she believes her book got wet too. This may be due to the fact that the most likely cause of wet grass is rain, which dampens things on the patio as well. A less plausible transition (world u) is caused by the sprinkler being activated. However, had she observed *Dry(B)* in addition to *Wet(G)*, she would have accepted this less plausible sprinkler explanation—that the sprinkler had been turned on—and any of its additional consequences, such as her glasses being dry if they are with her book.

2.3 An Event-Based Semantics for Update

One of the difficulties with the KM update semantics is the interpretation of the orderings $\leq_w$. This semantics supposes that it is "natural" to directly rank possible evolutions of a world w. In [8] we argue that evolutions or changes in the world should not be ranked directly. We suppose that *events* or actions provide the impetus for change, and the plausibility of a given evolution is determined by the plausibility of the event that caused the change. This approach is motivated by the observation that users can often more readily assess the relative plausibility of an event (in a given context) and the effects of that event, as opposed to directly assessing the plausibility of an evolution.

Apart from providing a more intuitive semantic foundation for belief update, this event-based model is more general than the KM model, and can be used to show that some of the KM postulates are too restrictive to be viewed as a general characterization of the process of belief update [8]. In order to unify update and revision, rather than generalizing the KM update semantics directly, we will base our unifying model on the event-based semantics of [8]. We briefly review the basic elements of this semantics.[3]

We assume a set of events E. An *event* e maps each world into another world, and can be viewed as a function (perhaps partial), $e : W \to W$. The world $e(w)$ is the *outcome* of event e at world w. Events such as these are therefore deterministic.[4]

Since an agent making an observation will often not know *a priori* what event caused the observed fact to hold, we assume that each world has associated with it an *event ordering* $\mu(w)$ that describes the plausibility of various event occurrences at that world. Formally, $\mu : W \to (E \to \mathbf{N})$; we write κ_w to denote the ranking $\mu(w)$. Intuitively, $\kappa_w(e)$ captures the plausibility of the occurrence of event e at world w. Again, we assume κ_w is a partial function over E, with $\kappa_w(e) = \infty$ taken to mean that e cannot occur at w. For each w, we require that $\kappa_w(e) = 0$ for some event e (perhaps several), so that at least one event is considered most plausible. We take the set of events and the ranking functions κ_w to constitute an *event model EM*.

With an event ordering in hand, one can easily rank the possible evolutions of a world w according to the relative plausibility of the events that could cause that evolution. In particular, we can define an *outcome* ranking λ_w for world w over the set W, where $\lambda_w(v)$ denotes the degree of plausibility

[3] Our presentation will rely on the use of κ-rankings (which impose total preorders), whereas the semantics in [8] is purely qualitative (and permits preorders as plausibility relations). In the few places where this influences results, we will make the distinction clear.

[4] In [8], nondeterministic actions are captured by allowing set-valued outcomes. In Section 3 below, we will want to generalize this further to allow these nondeterministic outcomes to be ranked according to plausibility.

associated with the transition of world w to world v. This can be defined as

$$\lambda_w(v) = \min_{e \in E}\{\kappa_w(e) : e(w) = v\}$$

In other words, the evolution of w into v is exactly as plausible as the most plausible event that causes w to evolve into v.

In the case of a deterministic event model, we can define a belief update operator as follows:

$$\|KB \diamond_{EM} A\| = \bigcup_{w \in \|KB\|} \{\min(A, \lambda_w)\}$$

In other words, we simply use the ranking λ_w as we would the plausible change ordering $\leq_w$ in the KM model. One distinction is that for any observation A, one can use the event model to generate an *explanation* for that observation. In other words, one can determine the event-condition pairs, for any condition consistent with *KB*, such that event e is the most plausible cause of the observation A. To revisit the example above, the ordering $\leq_w$ in Figure 2 may be induced by the event ranking where $\kappa_w(\textit{rain}) = 0$ and $\kappa_w(\textit{sprinkler}) = 1$. Not only is the belief *Wet(B)* a consequence of observation *Wet(G)*, but the explanation "It rained" is also forthcoming. We refer to [8] for further details.

One can show that the event-based semantics for update generalizes the KM model. Under particular assumptions, the classes of update operators determined by each semantics coincide, though some of the necessary requirements on event models may, in certain cases, be unnecessarily restrictive [8]. We defer discussion of this issue until we examine our generalization of this event-based semantics below. A final advantage of this model is that it lends itself readily to the generalizations required to deal with nondeterministic events with outcomes of varying plausibilities, as well as the incorporation of belief revision into the picture to provide a unifying semantics of belief change in dynamical systems.

3 Generalized Update

In this section we first describe some of the difficulties with the KM theory of update, as well as the event-based semantics described above, when it comes to dealing with the routine belief change of an agent embedded in a dynamical system. We then present the *generalized update* model, and illustrate the basic intuitions by means of two examples. We defer a formal analysis of its properties until the following section.

3.1 Difficulties with KM Update

One difficulty with the KM theory of update is that it does not allow an observation to force revision of an agent's beliefs about the state of the world prior to the observation. This is a crucial drawback, for even though one may not care about outdated beliefs directly, information gained about one's prior state of belief can influence updated beliefs.

Even simple tasks such as modeling *information gathering actions* are beyond the scope of KM update. Consider, for example, Moore's [26] litmus test: the contents of a beaker are unknown and one dips litmus paper into it to determine if it is an acid or a base. The prior state of belief is captured by two possible worlds—in both of these worlds, the litmus papers is some neutral color (say, *yellow*), and in one the proposition *acid* holds, while in the other *base* is true. The color of the paper after the test action *should rule out one of the possibilities*. Unfortunately, the KM theory does not allow this to take place; the semantics of update requires that both prior possibilities be updated to reflect the observed color (e.g., *blue*). One is forced to accept that, if the contents were acidic (in which case it should turn red), some extraordinary change occurred (the test failed, the contents of the beaker were switched, etc.). Note that one cannot escape the dilemma by supposing there is *no* such transition, for postulate (U3) ensures that updating *acid* by *blue* is consistent [8].

We can relax the KM update model to allow certain *KB*-worlds to be ruled out if the observation is not reachable through any "reasonable" transition from that world. This would dictate the addition of machinery to give a meaningful interpretation to the term "reasonable." But we must go further. It may be that an observation "conflicts" with *all KB*-worlds. To continue the example, imagine the contents of the beaker are not unknown, but are believed to be acidic. If the test result is *blue*, the KM model requires the agent to postulate some (very unusual) transition from a world where the beaker contains an acid to a world where the paper is blue. Of course, the right thing to do is simply admit that the beaker did not, in fact, contain an acid—the agent should *revise* its beliefs about the contents of the beaker. In order to do this, we must extend the model of update to deal with *structured* or ranked belief sets so that we have some guidance for the revision of our beliefs. In general, belief change will involve certain aspects of both revision and update.

3.2 Generalized Update Semantics

Rather than generalizing the KM update semantics directly, we adopt the event-based approach described in Section 2.3. As above, we assume a set of events E. However, we allow these events to be nondeterministic, and each possible outcome of an event is ranked according to its plausibility. For example, an attempt to pick up a block will likely result in a world where the block is held, but occasionally will fail, leaving the agent empty-handed.

Definition 3.1 An *event* e maps each world into a (partial) κ-ranking over worlds, $e : W \rightarrow (W \rightarrow \mathrm{N})$. We use $\kappa_{w,e}$ to denote the ranking $e(w)$.

Intuitively, $\kappa_{w,e}(v)$ describes the plausibility that world v results when event e occurs at world w. We say v is a *possible outcome* of e at w iff $\kappa_{w,e}(v)$ is defined (i.e., if $\kappa_{w,e}(v) \neq \infty$). We call this evolution of w into v, under the specified event e, a *transition*, which we write $w \stackrel{e}{\rightarrow} v$. We note that since $\kappa_{w,e}$ is a κ-ranking, we must have $\kappa_{w,e}(v) = 0$ for some v; that is, some outcome of event e must be most plausible. We occasionally assume the existence of the *null event* n, such that $\kappa_{w,n}(w) = 0$ and $\kappa_{w,n}(v) = \infty$ if $w \neq v$. The null event ensures (with certainty) that the world does not change.

As in the original event-based semantics, we will assume each world has an event ordering associated with it that describes the plausibility of various event occurrences at that world.

Definition 3.2 An *event ordering* μ maps each world into a (partial) κ-ranking over the set of events E, $\mu : W \rightarrow (E \rightarrow \mathrm{N})$. We write κ_w to denote the ranking $\mu(w)$.

To reiterate, $\kappa_w(e)$ captures the plausibility of the occurrence of event e at world w. Again, we assume κ_w is a partial function over E, with $\kappa_w(e) = \infty$ taken to mean that e cannot occur at w. We also note again that $\kappa_w(e) = 0$ for some e (i.e., some event is most plausible).

Finally, we assume that an agent's epistemic state, its beliefs about the current state of the world, are reflected in a straightforward κ-ranking κ over W. The plausibility accorded to world w is just $\kappa(w)$. These three components are put together to form a *generalized update model*.

Definition 3.3 A *generalized update model* has the form $M = \langle W, \kappa, E, \mu \rangle$, where W is a set of worlds, κ is a κ-ranking over W (the agent's epistemic state), E is a set of events (mappings $\kappa_{w,e}$ over W), and μ is an event ordering (a set of mappings κ_w over E). We assume that K is the belief set induced by κ.

In summary, an agent must have information about the nature of the current state of world (κ), what is likely to happen or not (μ), and the effects of those event occurrences (E). Such models contain the information necessary to update K in response to an observation A; we denote the resulting belief set $K^{\diamond}_{A}$. We now describe the update process.

To begin, we suppose that one "tick of the clock" has passed and that the agent must update its ranking κ to reflect the possible occurrence of certain events, without the benefit of observation. Intuitively, the *posterior* plausibility of a world v depends on the plausibility of the transitions that lead to v. The plausibility of a transition $w \stackrel{e}{\rightarrow} v$ depends on the plausibility of w, the likelihood that e occurred, and the likelihood of outcome v given w, e. In other words:[5]

$$\kappa(w \stackrel{e}{\rightarrow} v) = \kappa_{w,e}(v) + \kappa_w(e) + \kappa(w) \tag{1}$$

With this in hand, an updated ranking $\kappa^\diamond$ can be given by

$$\kappa^\diamond(v) = \min_{w \in W, e \in E} \{\kappa_{w,e}(v) + \kappa_w(e) + \kappa(w)\} = \min_{w \in W, e \in E} \{\kappa(w \stackrel{e}{\rightarrow} v)\} \tag{2}$$

This epistemic state essentially captures the notion that the world has evolved one "step" but that the agent has no information about the nature of this transition (other than that contained in the model M). We note that the agent's actual beliefs are determined by the minimal worlds in $\kappa^\diamond$ (i.e., those v such that $\kappa^\diamond(v) = 0$). We sometimes refer to $\kappa^\diamond$ as the anticipated or *predicted* updated ranking.

As with KM update, updates usually occur in response to some observation, with the assumption that something occurred to cause this observation. After observing A an agent should adjust its beliefs by considering that only the most plausible transitions leading to A actually occurred. The set of *possible A-transitions* is:

$$Tr(A) = \{w \stackrel{e}{\rightarrow} v : v \models A \text{ and } \kappa(w \stackrel{e}{\rightarrow} v) \neq \infty\}$$

The *most plausible A-transitions*, denoted $\min(Tr(A))$, are those possible A-transitions with the minimal κ-ranking. Given that A has actually been observed, an agent should assume that one of these transitions describes the actual course of events. The worlds judged to be epistemically possible are those that result from the most plausible of these transitions:

$$result(A) = \{v : w \stackrel{e}{\rightarrow} v \in \min(Tr(A))\}$$

Definition 3.4 Let K be the belief set determined by update model M. The *generalized update of K by A* (w.r.t M) is

$$K_A^\diamond = \{B : result(A) \subseteq \|B\|\}$$

[5]We note that this formula is the qualitative analog of the probabilistic equation $Pr(w \stackrel{e}{\rightarrow} v) = Pr(v|w,e) \cdot Pr(e|w) \cdot Pr(w)$ as described in Section 2.1.

In other words, an agent updating by observation A believes what is true at the states that result from the most plausible A-transitions.[6]

This model views generalized update by A as the process of determining the most plausible ways in which A may have been brought about. It is not hard to see that the very same belief change operator is dictated by the process of first determining the predicted updated ranking $\kappa^\diamond$ followed by (standard AGM) revision by A with respect to $\kappa^\diamond$.

Proposition 3.1 *result*$(A) = \min(A, \kappa^\diamond)$*; or, equivalently,* $K_A^\diamond = \{B : \min(A, \kappa^\diamond) \subseteq \|B\|\}$

This conforms to our intuitions about the updating process: the direct update of K by A, $K_A^\diamond$, determines the same belief set as the process of first updating one's entire epistemic state κ to get $\kappa^\diamond$, and then performing belief *revision* of $\kappa^\diamond$ by the observation A. Loosely, we might say $(K^\diamond)_A^* = K_A^\diamond$.

This notion of update naturally gives rise to the notion of an explanation for observation A. We can view updating by A as a process of postulating the most likely explanations for A and adopting the consequences of these explanations as our new beliefs. Unlike update of unstructured belief sets, explanations must consider (and trade-off) plausible initial conditions, events and event outcomes that lead to A.

Definition 3.5 An *explanation* for A (given model M) is any triple $\langle w, e, v\rangle$ such that $w \stackrel{e}{\rightarrow} v \in \mathit{Tr}(A)$ (which implies $\kappa(w \stackrel{e}{\rightarrow} v) < \infty$).

An explanation thus takes the form "It is *possible* that e occurred at w, leading to v and resulting in A." Of course, many of these explanation can be highly implausible.

Definition 3.6 The triple $\langle w, e, v\rangle$ is a *most plausible explanation* for A iff $w \stackrel{e}{\rightarrow} v \in \min(\mathit{Tr}(A))$.

In other words, the most plausible explanations are those explanations with minimal κ-ranking.

If A is *explainable* (i.e., if the set of explanations is not empty), then the most plausible explanations correspond to the most plausible A-transitions: thus generalized update can be interpreted as an abductive process. Given observation A, we can determine our updated belief set by first finding the most plausible explanations for A, and then adopting the "consequences" of these explanations are our new belief set. It is simply the form of the explanation—"something was true, something occurred, and it had this outcome leading to observation A"— that is more complex than in many other forms of abduction. Note, however, that Proposition 1 means we are not required to generate explanations explicitly in order to produce the updated belief set.

[6]Note that the exact form of $\diamond$ depends on the entire generalized update model M. To keep notation simple, we do not subscript the operator; the update model defining $\diamond$ should always be clear from context.

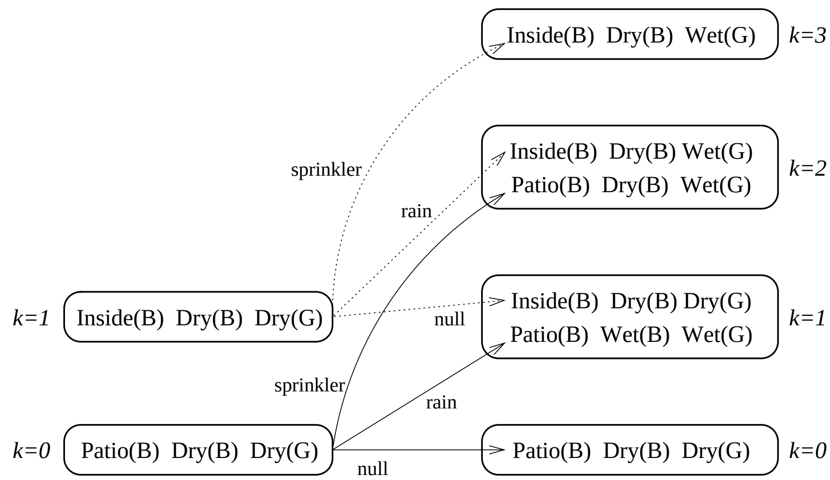


Figure 3: Generalized Update with Multiple Events

3.3 Examples

Before considering the formal properties of this model, we illustrate its nature with two examples. To keep the treatment simple, in the first example we use only deterministic events, while in the second we assume a single nondeterministic event (the agent's action).

Figure 3 illustrates the prior belief state of an agent who believes her book is on the patio and that both the grass and her book are dry. However, if her book is not on the patio, she believes she has left it inside ($\kappa(\textit{Inside}(B)) = 1$). We omit other less plausible worlds. We assume three events: it might rain, the sprinkler might be turned on, or nothing happens (the null event). She judges $\kappa_w(\textit{null}) = 0$, $\kappa_w(\textit{rain}) = 1$ and $\kappa_w(\textit{sprinkler}) = 2$, so *rain* is more plausible than *sprinkler* (we assume a "global" ordering, suitable for all w). The outcomes of these events are deterministic — in particular, both rain and the sprinkler will make the grass wet, but the book will only get wet if it rains and it is on the patio. Now, if wet grass is observed, our agent will update her beliefs to accept *Wet(G)*. A consequence of this is that she will now believe her book is wet: the most likely explanation is simply that it rained. If *Wet(G)* $\wedge$ *Dry(B)* are both observed (for instance, if she is told the book is safe), there are two most plausible posterior worlds satisfying the observation (i.e., $\kappa(\textit{Wet(G)} \wedge \textit{Dry(B)}) = 2$). This corresponds to the existence of two plausible explanations: either the book is on the patio ($\kappa = 0$) and the sprinkler turned on ($\kappa = 2$); or the book is inside ($\kappa = 1$) and it rained ($\kappa = 1$). The result of this update on the agent's state of belief is such that the agent is no longer sure where the book is. If we had instead set

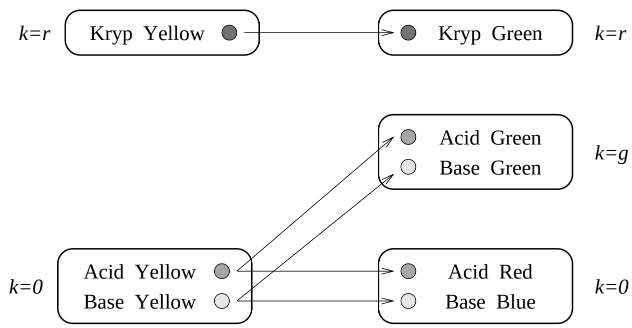


Figure 4: Generalized Update with Multiple Action Outcomes

$\kappa(\textit{sprinkler}) = 3$, observing *Wet(G)*$\wedge$*Dry(B)* would have caused the agent to believe that the book had been inside all along. The sprinkler explanation for the dry book becomes less plausible than having left the book inside. We see then that observing certain changes in the world can cause an agent to revise its beliefs about previous states of affairs. These *revisions* can impact on subsequent predictions and behavior (e.g., if the book is inside then so are her glasses).[7]

A second example is shown in Figure 4. We assume only one possible event (or action), that of dipping litmus paper in a beaker. The beaker is believed to contain either an acid or a base ($\kappa = 0$); little plausibility ($\kappa = r$) is accorded the possibility that it contains some other substance (say, kryptonite). The expected outcome of the test is a color change of the litmus paper: it changes from yellow to red if the substance is an acid, to blue if it is a base, and to green if it is kryptonite. However, the litmus test can fail some small percentage of the time, in which case the paper also turns green. This outcome is also accorded little plausibility ($\kappa = g$). If the paper is dipped, and *red* is observed, the agent will adopt the new belief *acid*. Unlike KM update, generalized update permits observations to rule out possible transitions, or previously epistemically possible worlds. As such, it is an appropriate model for revision and expansion of beliefs due to information-gathering actions. An observed outcome of green presents two competing explanations: either the test failed (the substance is an acid or a base, and we still don't know which) or the beaker contains kryptonite. The most plausible explanation and the updated belief state depend on the relative magnitudes of g and r. The figure suggests that $g < r$, so the a test failure is most plausible and the belief *acid*$\vee$*base* is retained. If test failures are more rare

[7] The world satisfying *Inside*(B), *Dry(B)*, *Wet(G)* at $\kappa = 3$ is shown for illustration. Technically, that world has rank 1 since it occurs below, and the explanation "sprinkler and book inside" will never be adopted, unless further propositions and observations can distinguish the two worlds (e.g., other effects of the events in question).

($r < g$), then this outcome would cause the agent to believe the beaker held kryptonite.

4 Relationship to Revision and Update

We now turn to the question of the relationship of GU to the classical AGM revision and KM update models. The analysis of the update postulates is in many ways similar to that presented in [8], where the simple event-based semantics of Section 2.3 was developed. There we showed that under certain assumptions this event-based operator satisfies the KM postulates, though we argued that these assumptions are not always appropriate. The key difference here is that the abductive approach has been generalized to allow ranked outcomes of events, and more importantly, ranked belief structures. This has surprisingly little impact on the analysis of the KM postulates—the basic models satisfy the same postulates and the same assumptions can be used to ensure satisfaction of addition postulates—with one significant exception: the postulate (U8) becomes (in a certain sense) meaningless under GU. We elaborate on this below.

We first note that our model satisfies a number of the KM postulates.

Proposition 4.1 *If $\diamond$ is the GU operator induced by some GU model then $\diamond$ satisfies postulates (U1), (U4), (U5), (U6), (U7) and (U9).*

We note that GU satisfies the all of the same update postulates as the basic event-model for flat belief states [8].

One key difference between the GU model and the KM model is reflected in (U2), which asserts that *KB* $\diamond A$ is equivalent to *KB* whenever *KB* entails A. This cannot be the case in general, for even if *KB* $\models A$, the most plausible event occurrence may be something that changes another proposition while leaving A true. Observing A may simply mean that the change proceeded as expected.[8] (U2) is appropriate only if we are willing to assume *persistence* of propositions, that changes (are believed to) occur only if evidence for them is observed. While appropriate in some settings, this is not a universal principle suitable for belief change. Nevertheless, we can model it by assuming *inert update models*, in which the null event is the most plausible in any situation.

Definition 4.1 A GU model $M = \langle W, \kappa, E, \mu \rangle$ is *inert* if E contains the null event n and $\kappa_w^{-1}(0) = \{n\}$ for all $w \in W$.

[8] As an example, consider an assembly line monitor that observes a "status OK" signal. This may well already be believed; but the agent should still update its belief state by *changing* its belief about the number of parts that have been produced.

Proposition 4.2 *If $\diamond$ is induced by an inert GU model then $\diamond$ satisfies (U2).*

The second key difference is reflected in the failure of (U3), which asserts that *KB* $\diamond A$ is satisfiable if A is. In our model, this corresponds to every A being explainable no matter what beliefs are held. GU models need not satisfy (U3). Consider the case where no event can result in an A-world (i.e., where $Tr(A) = \emptyset$): the observation of A is then unexplainable, and $K_A^{\diamond} = \mathbf{L}_{CPL}$, the inconsistent belief set. To prevent this, we can simply insist that every satisfiable sentence A is explainable.

Definition 4.2 A GU model $\langle W, \kappa, E, \mu \rangle$ is *complete* iff for any satisfiable $A \in \mathbf{L}_{CPL}$, there are $w, v \in W$, $e \in E$ such that $\kappa(w) < \infty$, $\kappa_w(e) < \infty$, $\kappa_{w,e}(v) < \infty$ and $v \models A$.

Proposition 4.3 *If $\diamond$ is induced by a complete GU model then $\diamond$ satisfies (U3).*

In [8] we criticized (U3) as inappropriate for the update of flat belief sets. For example, if our beliefs corresponded to a single world where *acid* is believed, (U3) forces the observation of *blue* to behave quite poorly (as described above). However, such a maxim is much more reasonable in generalized update. It does not force one to propose wildly implausible transitions from prior epistemically possible states; instead one can revise one's beliefs to account for the observation. In this case, we simply give up the belief *acid*. For this reason, (U3) may be seen as a reasonable postulate for GU, in which case we might take complete GU models to provide the appropriate semantic underpinnings for belief revision and update.

There are a number of systematic ways in which one can enforce the condition of completeness such as requiring the existence of "miraculous" events that can cause anything [8]. In our setting, one quite reasonable condition we might impose is that all worlds have some plausibility (i.e., κ is a total function on W) and that the null event is possible (though not necessarily very plausible) at each of those. The first requirement is usually assumed of epistemic states (e.g., in the belief revision literature), and the second simply ensures that all worlds persist with *some* degree of plausibility. Thus while explanations of A may be implausible they will not be impossible.

Finally, putting Propositions 4.1, 4.2, and 4.3 together we have:

Theorem 4.4 *If $\diamond$ is induced by a complete, inert GU model then $\diamond$ satisfies (U1)–(U7) and (U9).*

We note that the converse of this theorem and the preceding propositions is easy to verify, though not especially interesting. Primarily, we are interested in determining the nature of belief change *given* information about beliefs, events and event orderings, rather than the construction of models that corroborate arbitrary operators satisfying the postulates. We also note that our characterization theorem includes (U9) because of our use of κ-rankings, which totally order events and worlds. One of the main reasons for using such rankings, as discussed in Section 2.1, is that they allow the scales of plausibility used to rank worlds, events and outcomes to be compared and added. In general, the use of qualitative ranking relations does not admit this flexibility unless one is willing to postulate a "metric" by which a combination of preorders can be compared. This is not a difficult task, but is somewhat more cumbersome than the approach provided here. Equivalent results should be obtainable in the more general setting of arbitrary preorders.

There are two special cases of GU that are worth mentioning in passing. First, we note that "plain" KM update of unstructured belief sets is easily captured in our model by the simple restriction of κ to rank worlds only as plausible ($\kappa = 0$) or impossible ($\kappa = \infty$). Second, reasoning about agent-controlled action (and observations) is also possible, as indicated in the litmus example. To do so, we simply view an agent's actions as events: we associate with each action a a κ-ranking $\kappa_{w,a}$ that ranks outcomes of action a at world w. We take the key difference between actions and events (at least, as far as belief change is concerned) to be that actions are within the agent's control so that it has *direct* knowledge of their occurrence. As such, actions need not be ranked according to their plausibility of occurrence, nor do they need to be postulated as part of an explanation. Observations can only be explained by supposing the action had a particular (perhaps unexpected) outcome, or by revising beliefs about the initial conditions, or both.[9]

To complete our analysis of the KM postulates, we turn our attention to (U8). None of our characterization results involve (U8) because it cannot be enforced in a reasonable way in our model. The reason for this is our move to *ranked* models of epistemic states and our ability to explain (even "most plausibly" explain) an observation using initial conditions that conflict with our beliefs—that is, our ability to have observations *revise* our initial beliefs.

(U8) is the only update postulate that relates the update of different initial states of belief, namely KB_1, KB_2 and $KB_1 \vee KB_2$. For a given observation A, the update of KB_1 will generate a set of most plausible explanations $w \stackrel{e}{\to} v$, where $v \models A$. Unfortunately, unless *all* of these explanations are such that $w \in \|KB_1\|$, the properties of $KB_1 \diamond A$ are determined not only by KB_1 (together with the event model), but also by the ranking κ that represents our initial epistemic state. Unless we impose strong and unnatural conditions on the relationships of the rankings κ upon which our updates of KB_1, KB_2 and $KB_1 \vee KB_2$ are based, very little can be said about the relationship asserted by (U8). Indeed, we

[9] Concurrent events and actions require special attention, however, and are beyond the scope of this paper. Our framework is certainly compatible with standard treatments of concurrency.

must allow update to proceed for the many distinct epistemic states κ that determine a given knowledge base *KB*.

We note that for a fixed event model (i.e., a fixed E and μ), we *can* relate the update of KB_1, KB_2 and $KB_1 \vee KB_2$ under the following condition: the updated belief state $K_A^\diamond$ is a function of the initial belief state K (e.g., KB_1, KB_2 or $KB_1 \vee KB_2$) and not of the ranking κ. If this is the case, it is easy to verify that (U8) will be satisfied. However, the *only* circumstance under which this condition will hold is if a categorical preference is given to explanations for A of the form $w \stackrel{e}{\to} v$ where $w \in \|K\|$. But this simply means that one is unwilling to revise ones prior beliefs to account for an observation. In other words, postulate (U8) only makes sense when we are dealing with *flat, unstructured epistemic states*—precisely the types of models whose weakness GU is designed to counteract!

We wrap up by considering how AGM belief revision can be modeled in our framework. The common folklore states that belief revision is a form of belief change suitable when the world is static or unchanging. To verify this intuition, we propose *static update models*.

Definition 4.3 An update model $M = \langle W, \kappa, E, \mu \rangle$ is *static* if $E = \{n\}$ where n is the null event n.[10]

Assuming, as is usual in the belief revision literature, that κ is a total function over W (i.e., $\kappa(w) < \infty$ for all $w \in W$), we obtain the following result:

Theorem 4.5 *If $\diamond$ is induced by a static GU model then $\diamond$ satisfies (R1)–(R8).*

Static event models have as the only possible transitions those of the form $w \stackrel{n}{\to} w$ with plausibility $\kappa(w)$. Thus, the informal intuition about belief revision (and the AGM model) can be verified formally: AGM revision is a particular form of GU suitable for a "static" system. (The converse of Theorem 6 is easily verified.)

We note that the assumption of staticness is in fact much stronger than is needed to prove satisfaction of the AGM postulates. Indeed, simple inert models will satisfy this condition as well (as we see in Section 5.1). The reason we consider static models to be the *correct* specialization of GU for modeling revision will become clear once we discuss iterated belief update in Section 5.1.

5 Iterated Belief Update and Observations

The model we have described is strongly related to standard Bayesian models of belief update in stochastic dynamical systems. Roughly, in these models, an agent's epistemic state is captured by a prob-

[10] As above we assume κ is a total function on W.

ability distribution. The probabilities associated with various event occurrences and the outcomes of actions and events can be used in a Bayesian updating process to determine how to update this distribution given a specific observation. The updated distribution characterizes the agent's updated epistemic state, which then may be subject to update given subsequent observations.

While the connections to generalized update are quite strong,[11] there is one key difference: the stochastic dynamical system view shows how to produce an updated *epistemic state*, not just a simple belief state. The GU model, as described above, gives a characterization of the agent's updated *belief state* in response to an observation. The belief state itself does not provide guidance for changes in belief due to subsequent observations.

We address this problem, that of *iterated belief change* in this section. We first elaborate on the issues of iterated change and update of epistemic states. We then provide a short description of the basic partially observable, stochastic dynamical system model. Finally, we show how the intuitions of the quantitative dynamical system model can be captured in several different ways in our qualitative model (some of these being more direct than others).

5.1 Iterated Belief Change

To illustrate the need for a more elaborate specification of generalized update, we consider the raining example of Section 3.3. GU specifies that an agent with an epistemic state κ given by the initial ranking in Figure 3 and who observes *Wet(G)* will possess the belief state $\{$*Patio*(B), *Wet(B)*, *Wet(G)*$\}$. This corresponds to the revision of the updated ranking $\kappa^{\diamond}$ by the observation *Wet(G)*. Unfortunately, while it specifies an updated belief state, GU fails to dictate an *updated epistemic state* (see Figure 5). Thus, the agent has no idea how to incorporate subsequent observations in general. For example, if the agent now observes *Dry(B)*, it requires a legitimate epistemic state (e.g., a ranking) in order to revise or update this new belief state to account for the new observation. With only the basic belief state in its possession, the new observation is simply inconsistent, unless there is some miraculous book drying event.

We note that this problem is strongly related to the problem of *iterated belief revision*: given an epistemic state and an observation by which an agent's beliefs are to be revised, how should this *epistemic state* change (not just the belief state). It has been observed by a number of authors that the

[11] Indeed, the equations defining an updated ranking and the belief state induced by a specific observation can be viewed as qualitative analogs of the quantitative relations used for partially-observable stochastic dynamical systems. This will become apparent below.

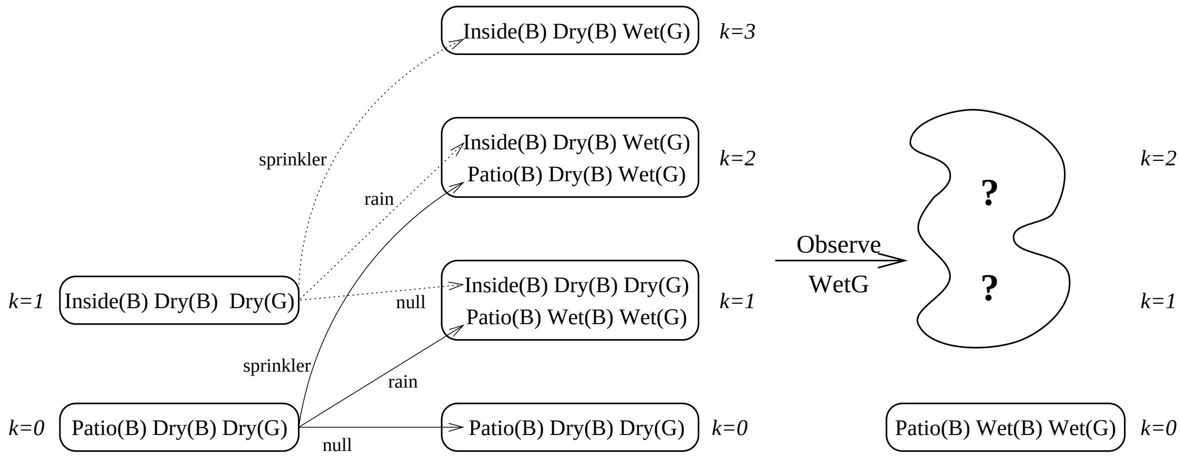


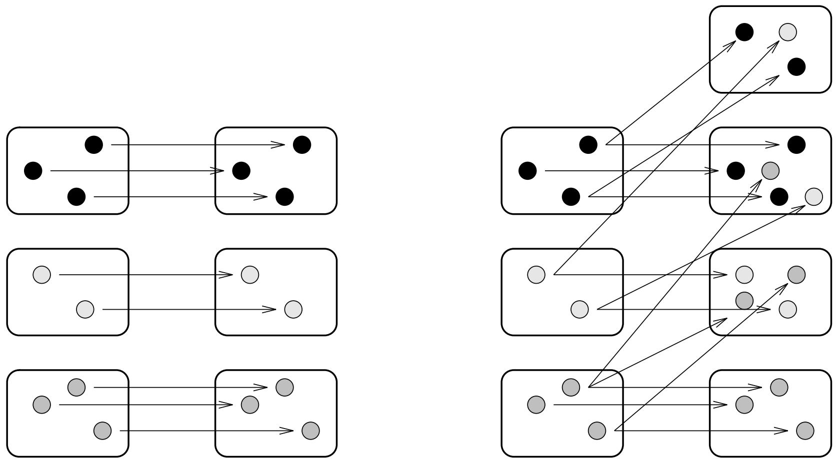
Figure 5: The Problem of Iterated Update

AGM theory of revision has little to say in this respect [23, 27]. As a result, a number of proposal for extending theories of belief revision to dictate revised rankings or epistemic states have been proposed [31, 5, 9, 32, 11]. In fact, it is the "revision component" of GU (as opposed to the "update component") that fails to adequately characterize the required change in epistemic state. This can be seen clearly in Figure 5: the model of the system dynamics allows one to update the entire ranking, while the revision of the updated ranking simply produces a belief set. Thus, a reasonable theory of belief revision, that specifies how one should revise an epistemic state, will automatically dictate how to apply GU to epistemic states.

We will describe several methods of effecting iterated revision (and, in particular, revision of epistemic states) within the GU framework below. In particular, two models of iterated revision within the theory of κ-rankings, as developed by Spohn [31] (and further investigated by Goldszmidt and Pearl [18, 19]), can be applied almost directly to GU. However, the application of these techniques to GU can be better motivated by illustrating the connections of GU to stochastic dynamical systems, which we do in the next section. We note that in a purely qualitative setting (with a relation $\leq$ of relative plausibility replacing a ranking function κ), iterated revision is especially problematic.[12] However, at one of the models developed below can be applied in such a setting.

Before proceeding with our account of iterated GU, we pause to reflect on the new light shed on

[12]For a discussion of some of the difficulties, see [23, 9, 7, 11].



(a) (b)

Figure 6: Iterated Revision with (a) Static and (b) Inert GU Models

Theorem 4.5, the relationship of GU and belief revision, by this view of iterated revision. Theorem 4.5 shows that if our GU-model is static (i.e., if the null event is the only plausible event), then GU will satisfy the eight AGM postulates for belief revision. The intuitions underlying such static models are clear, and are illustrated graphically in Figure 6(a): the update portion of GU leaves the ranking unaltered, thus allowing the revision component to proceed on the original ranking.

However, because of the weakness of the AGM postulates, the conditions of Theorem 4.5 are much stronger than necessary. For example, suppose we have an *inert* model, one where the null event is more plausible than other events (see Figure 6(b)). In such a model, other events can occur, but some "preference" is given to explanations of observations that require no changes in the world. Due to the fact that the AGM theory doesn't impose strong restrictions on the form of updated rankings, we have the following result:

Proposition 5.1 *If $\diamond$ is induced by an inert GU model then $\diamond$ satisfies (R1)–(R8).*

Intuitively, the only requirement of the AGM postulates is that the agent's epistemic state is characterized by some ranking in which worlds consistent with its beliefs form the set of most plausible worlds. In Figure 6(b) we can see that the update portion of GU preserves the agent's belief state, just like a static model. This means that the revision portion of GU will satisfy the AGM postulates with respect to the original belief state. Unfortunately, the inert model, unlike the static model, can shift the relative plausibility of all worlds other than those with rank 0. This means that generalized update by an observation that is not consistent with the agent's original belief set does not generally produce the same belief set that would be obtained by straightforward revision with respect to the original ranking.

Although the AGM postulates are satisfied, we should not consider "inert generalized update" to be the "special case" of GU corresponding to revision. Since the model is not truly static, one can produce truly bizarre results that we would not expect of belief revision. For instance, given an initial belief set K, it is quite possible with an inert model that $K^{\diamond}_{A}$ is very different from $(K^{\diamond}_{\top})^{\diamond}_{A}$. In other words, updating by a null observation followed by A can result in a different belief set than updating immediately by A. This is, of course, not surprising, given that the world can change. It is, however, surprising that it should satisfy the AGM postulates if these postulates are intended to characterize revision in static environments. Again, we emphasize that this is due to the fact that the AGM theory says nothing substantial about iterated revision or the revision of epistemic states.

We note that recent proposals for dealing with iterated revision (e.g., see Boutilier's MC-revision model [9], or the postulates of Darwiche and Pearl [11]) all insist that revision by $\top$ not affect the agent's epistemic state. It is clear that an inert model does not satisfy this requirement. However, static models do. Thus we legitimize the claim that, in fact, static GU models exactly capture the intuitions underlying belief revision, and formally verify the conventional wisdom that belief revision corresponds to belief change about static environments.

5.2 Stochastic Dynamical Systems and GU

We begin with a very brief description of a simple model of a partially observable, stochastic dynamical system.[13] As above, a system can be in a number of possible *states* $\mathcal{S}$.[14] The system dynamics can be characterized by two families of probabilities functions. *Event probabilities* refer to the probability of a particular event occurring at a given state: $\Pr(e|s)$ refers to the probability of e occurring in state s.[15] *Outcome probabilities* capture the probability of a particular state t resulting from the occurrence of an event e at state s, and are denoted $\Pr(t|s,e)$.

An agent's epistemic state is represented by a distribution over states, $\Pr_k(s)$ denoting the agent's degree of belief that the system is in state s at time k. The agent can update its distribution using its

[13] The interested reader can consult [25, 10, 3, 30, 24] for more detailed models, discussions of control, etc.

[14] We will use the term "state" in the context of stochastic dynamical systems, and "world" when discussing GU.

[15] Technically, $\Pr(e|s)$ is not a conditional probability, but a distribution over E as a function of s, $\Pr_s(e)$. We use this notation for its suggestiveness and familiarity. (Similarly for $\Pr(t|s,e)$ defined below.)

model of the dynamics as follows:

$$\Pr_{k+1}(t) = \sum_{e \in E} \sum_{s \in \mathcal{S}} \Pr_k(s) \Pr(e|s) \Pr(t|s,e) \tag{3}$$

This corresponds to the agent's predications about how the system will evolve in one "clock tick."

In order to account for observations the agent might make of the system state, we assume an *observation model*: we have a set of possible observations $\mathcal{O}$ together with a family of distributions $\Pr(o|s)$ representing the probability of making an observation $o \in \mathcal{O}$ when the true system state is s.[16] When an agent makes an observation o at time $k+1$, we can view the change of its epistemic state as a two-stage process: first, it updates its distribution to form $\Pr_{k+1}$ as above; second, it conditions this distribution on the observation o to obtain $\Pr^o_{k+1}$. This second phase can be computed using a simple application of Bayes Rule:

$$\Pr^o_{k+1}(s) = \frac{\Pr(o|s) \Pr_{k+1}(s)}{\sum_t \Pr(o|t) \Pr_{k+1}(t)} \tag{4}$$

We now turn our attention to the relationship between GU and this model of belief update in stochastic dynamical systems. We first note that straightforward predictive update, the updating of a ranking, follows exactly the same "rules" as the update of a probability distribution, the distinction being that qualitative probabilities ("kappas") are used in GU. In particular, Equation 2 is precisely the qualitative analog of Equation 3.

Accounting for observations is not quite so straightforward. The assumption underlying all work in belief revision and update, including the GU model as presented, is that the agent observes propositions directly. Unlike the standard dynamical system model, observations are actually part of the state (i.e., a proposition that is determined by the state). This makes it difficult to deal with fallible or "noisy" observations (though not impossible—see the next section). In order to account for such deterministic observations, we will specialize the dynamical system model by assuming that each state dictates precisely the observation that will be made. In other words, we assume that $\Pr(o|s) = 1$ if $s \models o$ and $\Pr(o|s) = 0$ if $s \not\models o$. In order to make sense of this, we must consider the proposition or variable that the agent is attempting to observe, otherwise there is no reason the agent cannot observe the entire state. Therefore, we make a tacit assumption that an agent explicitly acts to observe the truth value of a particular proposition. This action does not change the world in any way (it is a null event), but tells

[16] In more general models, one can allow the observation to depend on state transitions or on the action taken by the agent. Such complications are not germane to the discussion here.

the agent whether the proposition is true or false.[17] We will relax the assumption that this observation process is infallible in the next section.

With this deterministic observation model in hand, we note that updating by Bayes Rule in Equation 4 simplifies to simple conditioning by the observation o. In other words,

$$\Pr^{o}_{k+1}(s) = \begin{cases} \frac{\Pr_{k+1}(s)}{\sum_t \{\Pr_{k+1}(t) : t \models o\}} & \text{if} \quad s \models o \\ 0 & \text{if} \quad s \not\models o \end{cases} \tag{5}$$

Generalized update can be used to produce a new κ-ranking by conditioning in an entirely analogous way. Recall that $\kappa^{\diamond}$ refers to the updated ranking before revision by the operation (as defined by Equation 2). We define

$$\kappa^{\diamond}_{A}(w) = \begin{cases} \kappa^{\diamond}(w) \perp \kappa^{\diamond}(A) & \text{if} \quad w \models A \\ \infty & \text{if} \quad w \not\models A \end{cases} \tag{6}$$

This process of *conditioning* (in a static context) is described in [19]. Intuitively, the agent's initial ranking is updated, then the observation A is applied to the updated ranking by removing all $\neg A$-worlds (the remaining worlds are normalized by subtracting $\kappa^{\diamond}(A)$ so that a legitimate ranking results). We dub this process *infallible GU* because all observations are treated as being certain. The process of infallible GU naturally lends itself to iteration, since generalized update by an observation A results in an updated ranking. The result of infallible update by the observation *Wet(G)* in the rain example is illustrated in Figure 7. We note that conditioning is an especially simple form of updating an epistemic state in response to an observation. It can be be applied directly to purely qualitative ranking in the obvious way.

5.3 Noisy Observations in GU

The key difficulty with the infallible model of GU, in other words, conditioning directly on observed propositions, is the assumption that propositions in the domain are directly observable without error. For instance, given the infallible update by *Wet(G)* in the previous example, an agent cannot subsequently meaningfully update his epistemic state by the observation *Dry(G)* unless there is an explaining event that immediately causes the grass to dry. In no way will an explanation of *Dry(G)* include the possibility that the earlier observation was incorrect.

[17] Such actions might take the form of Scherl and Levesque's [28] *sense*$_P$ actions. This action returns an observation of the form P or $\neg P$ for a particular proposition P. This is the sense in which states "determine" observations (relative to a *given* sensing action).

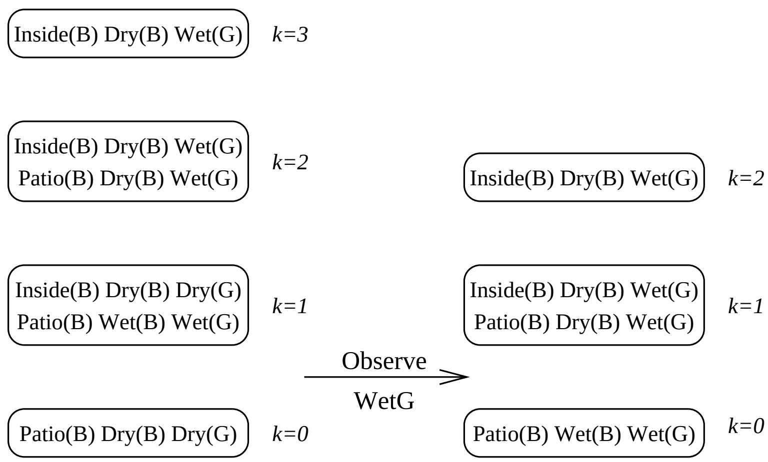


Figure 7: Conditioning and Infallible Generalized Update

One can deal with this difficulty in a rather obvious way: simply add propositions to the domain that refer to the observations that have been or can be made by the agent. The fact that an observation of the proposition *Wet(G)* is fallible can be modeled by relating the conditional degree of belief a proposition describing the observation (say *obsWetG*) to the proposition itself. For instance, an agent's epistemic state might reflect the fact that any state in which *Wet(G)* $\wedge$ *obsWetG* holds is more plausible than the corresponding world satisfying *Wet(G)* $\wedge$ *obsDryG* (e.g., by some amount $\kappa(\mathit{obsDryG}|\mathit{Wet(G)})$). In this sense, while conditioning by the observation *obsWetG* does preclude the fact that you might have made observation *obsDryG*, it does not rule out the possibility *Dry(G)*; it merely makes it less plausible. Of course, in order to fully develop this model of conditioning by observations, we would be forced to deal with time in some way or another. For example, conditioning on *obsWetG* rules out any worlds in which *obsDryG* is true. However, one could *subsequently* make a contradictory observation of dry grass. In order to condition on this new observation, we would require a different (but related) proposition that refers to *obsDryG* at the next point in time.[18] We could, for instance, use time-stamped observational propositions. In order to specify such a model in a convenient way, we would have to develop additional machinery. We do not pursue this here, though it could be developed in a less awkward fashion in the run-based model of Friedman and Halpern [12] (which is described below).

Instead, we will present a second alternative, based directly on standard stochastic observation

[18]A distinct observational proposition is necessary, since conditioning on *obsWetG* removes all worlds where *obsDryG* holds. Subsequent conditioning on *obsDryG* would result in inconsistent beliefs.

models. We assume that there exists a set of observations $\mathcal{O}$ that are related in a nondeterministic way to possible worlds. The *observation model* ρ describes the plausibility of various observations in different world states. Formally, $\rho : W \rightarrow (\mathcal{O} \rightarrow \mathrm{N})$ maps worlds into κ-rankings over the observation set. The ranking $\rho(w)$ describes how likely the agent is to observe a given $o \in \mathcal{O}$ in world w. We write $\kappa(o|w)$ for $\rho(w)(o)$. As usual, for each w, there is at least one o such that $\kappa(o|w) = 0$.

Once again, we do not assume that an observation is dictated solely by the true state of the world, but is influenced by the *observation action* taken by the agent. For instance, if the agent executes an "observe grass" action, the appropriate observation set might be $\{$*obsWetG*$,$ *obsDryG*$\}$, and the observation model might be specified as follows: $\kappa($*obsWetG*$|$*Wet(G)*$) = 0$; $\kappa($*obsDryG*$|$*Wet(G)*$) = 2$; $\kappa($*obsWetG*$|$*Dry(G)*$) = 3$; $\kappa($*obsDryG*$|$*Dry(G)*$) = 0$.[19] To keep the presentation a bit simpler, we suppress any conditioning of observation models on the observational action being taken, especially since this impacts little on the development (and the extension is obvious). We simply assume that an appropriate observation is (actively or passively) obtained by some means.

With this model of "fallible" observations, we can extend GU to a form of *fallible generalized update* much as we did to obtain infallible GU. To do this we will use the analog of Equation 4 directly (recall, that in the infallible/conditioning case, this application of Bayes Rule reduced to the simpler form given by Equation 5). Recall that $\kappa^{\diamond}$ refers to the updated ranking before revision by the operation (as defined by Equation 2). We define, for any observation A:

$$\kappa_A^{\diamond}(w) = \kappa(A|w) + \kappa^{\diamond}(w) \perp \kappa^{\diamond}(A) \qquad (7)$$

We note that the last term reflects the absolute plausibility of making observation A and is defined as:

$$\kappa(A) = \min_{v \in W}\{\kappa(v) + \kappa(A|v)\}$$

Intuitively, this model does precisely what one expects: given an observation A, each world becomes more or less plausible according to the degree to which it gives rise to A. Worlds for which A is expected ($\kappa(A|w) = 0$) keep the same relative rank, while worlds for which A is unexpected ($\kappa(A|w) = k > 0$) become less plausible by the degree k to which A was surprising. Finally, the normalization term $\kappa(A)$ is subtracted to ensure that an appropriate ranking (with minimal elements having rank 0) is obtained. We note that if observations are deterministic, or embedded in the world

[19]This method of specification summarizes the observation model; e.g., $\kappa($*obsWetG*$|$*Wet(G)*$) = 0$ means that $\kappa($*obsWetG*$|w) = 0$ for all $w \models$ *Wet(G)*.

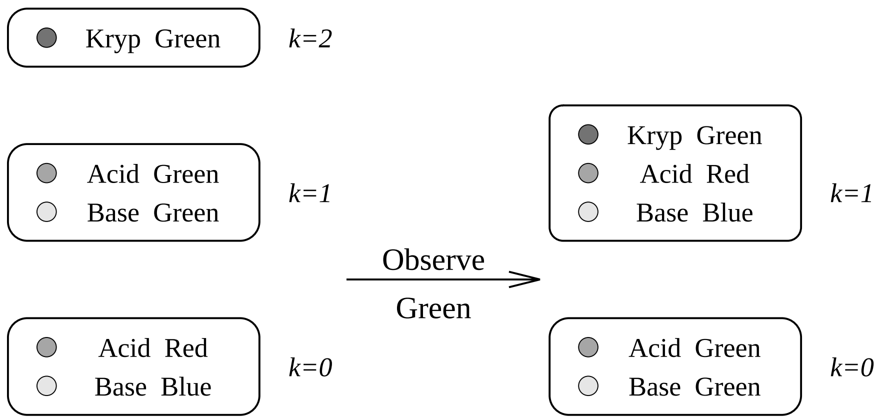


Figure 8: Uncertain Observations and Fallible Generalized Update

state description, then Equation 7 reduces exactly to Equation 6 and is nothing more than conditioning.

To illustrate the process, consider the litmus example of Figure 4. Now imagine that observations of color are imperfect, so that false readings are implausible to degree 2; that is, $\kappa(\textit{obsColor1}|\textit{Color2})$ is 0 if the true color and observed color are the same, and is 2 if they are different. The result of fallible update by the observation *obsGreen* given this observation model is illustrated in Figure 8

We note that this model of noisy observations cannot be used, at least in its full generality, in a purely qualitative setting (though see remarks in Section 7). A key element of the observation model is the ability to specify that certain observations occur with varying degrees of plausibility, and the ability to compare (and combine) the degrees of plausibility accorded to particular worlds by an observation with their prior degrees of plausibility.

Finally, we point out a third way in which noisy observations can be incorporated, through the use of *strength of evidence*. In the original development of κ-functions, Spohn [31] does not emphasize conditioning by observations, but α-*conditioning*, where a proposition A is observed with a certain (integer) strength α. Intuitively, the relative plausibility of all A-worlds is shifted by α (so that if α is positive, A becomes more plausible). This process is dubbed J-*conditioning* in [19] (as it is exactly analogous to Jeffrey conditioning of a probability distribution). In [19] a similar process known as L-*conditioning* is also proposed, whereby A does not become α degrees more plausible, but A is made to plausible to degree α.

Instead of having an explicit observation model, one could directly observe propositions in the domain, and account for the uncertainty of these observations by attaching strengths to them. In this way, instead of direct conditioning by A (and its attendant difficulties), J-conditioning or L-conditioning may be used. We do not pursue this suggestion here; the reader is referred to [31, 19] for details of this type of conditioning. The application of these ideas to GU is straightforward. However, a key question that arises is the following: how does one determine the strength of an observation? The obvious answer is that one must have an observation model. For this reason, we find the model of fallible generalized update to be the most compelling of qualitative belief change in dynamic settings.

6 Related Work

6.1 Trajectory-based Semantics

There are a number of ideas that have directly influenced the generalized update model, including the work cited above on belief revision, belief update, κ-rankings, iterated revision, reasoning about action and dynamical systems. However, there have been few attempts to tie all of these ideas together in a comprehensive, qualitative framework. There is one framework in which belief update and belief revision can both be expressed, namely, the run-based (or trajectory-based) model of Friedman and Halpern [12]. This model is similar to ours, especially in its general outlook on revision and update and its adoption of a dynamical systems perspective.

In the run-based framework, an agent imposes a qualitative plausibility relation (serving the role of a κ-function) not over possible worlds, but on possible *system trajectories* or histories. Each trajectory corresponds to a sequence of *transitions*, or states the world might have passed through. The relative plausibility of a trajectory captures the degree to which an agent thinks that trajectory might be (or has been) realized. When an observation is made at a particular point in time, a normal belief revision process can be applied. The most plausible trajectories after revision by a particular observation make up the agent's new belief state. As such, an agent can have beliefs that extend over time. Friedman and Halpern also show that under certain assumptions belief change in this model satisfies the revision postulates and under certain assumptions satisfies the KM update postulates.

One impediment to the use of a trajectory-based model is determining an initial ranking of entire trajectories. While Friedman and Halpern do not emphasize this point, it is clear that the individual (state-to-state) transitions within a trajectory should correspond to the occurrence of events within the system, and that the (initial) relative plausibility of a trajectory should simply be a function of the (prior) plausibility of its initial state and the plausibilities of the individual event occurrences along the trajectory (together with the plausibility of the corresponding event outcomes). In this respect, it is easy to see how GU can be embedded in the Friedman-Halpern model. The distinctions lie in emphasis: we focus on the source of trajectory plausibilities as a function of individual events, actions, or their outcomes, while Friedman and Halpern emphasize the relative plausibility of entire trajectories.

Viewed in this light we can also see some of the technical assumptions implicit in the GU model. First, our observation and system dynamics models treat the plausibility of observations, events and event/action outcomes as a function of world state; the actual history an agent has passed through does not influence the agent's estimate of action outcome or observation likelihoods. Thus, we are making the *Markovian assumption*: dynamics are independent of history. For this to provide an accurate model of a domain, our states or worlds must contain enough information to render predictions of the future independent of the past. This is generally not very restrictive, and is the assumption underlying almost all work in AI on planning and reasoning about action. However, the trajectory-based model is more general in this respect. Non-Markovian systems are easily representable in the Friedman-Halpern model. Furthermore, we have made an assumption of *stationarity*: the system dynamics and observation model do not change over time. Again, while this assumption is relatively uncontroversial, the Friedman-Halpern model can very easily deal with nonstationary systems.

We note that in subsequent work [14], Friedman and Halpern address the difficulty of associating plausibility with entire trajectories and come to precisely the same conclusions: the Markovian and stationarity assumptions are very natural and provide great leverage in specifying system dynamics (and for their model, easily allow one to determine the plausibility of trajectories). The assumptions that frameworks for qualitative belief change, like GU and the Friedman-Halpern system, show to be effective in modeling belief change are exactly those that have been adopted in the dynamical systems community for years. Finally, we note that Geffner [16] has recently developed a model for reasoning about action using the κ-calculus to represent the defeasibility of action effects. By making the Markov assumption explicit, he uses this model to determine the plausibility of system trajectories.

To illustrate the power of GU (as well as the implications of assumptions such as Markovian dynamics), we consider the *stolen or borrowed car problem*.[20] The setting is as follows: you arrive at a very expensive restaurant, and leave your very expensive car with the valet to be parked. At the end of the evening you pick up your car from the parking lot and begin to drive home. A few miles later, you notice that the odometer reading is extremely high (say 50 miles higher than when you arrived at the restaurant; it's a very expensive car and you are obsessive about the mileage!). How will your belief

[20] The *borrowed car problem* is developed by Friedman and Halpern in their trajectory-based model, but the example is attributed to the author. The related *stolen car* problem is due to Kautz [22].

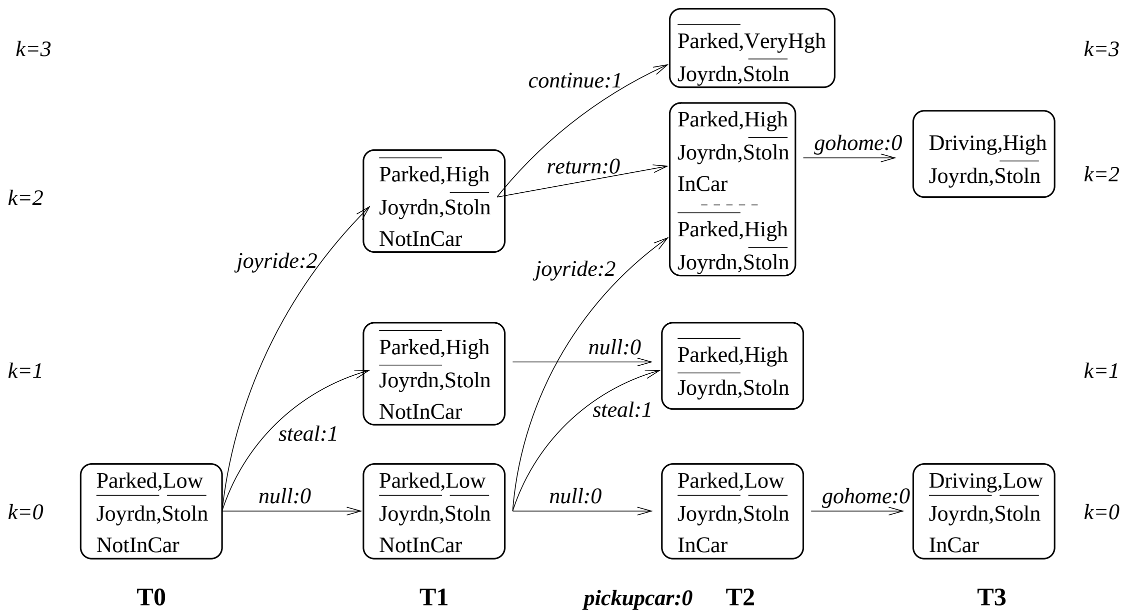


Figure 9: The Borrowed or Stolen Car Problem

state change?

Intuitively, what we are after in this example is the explanation that the valet took your car out for a joyride while you were eating. To see how this might arise, we consider an event model in which there are several possible events, including the null event (where your car stays put), the *steal* event (where your car is stolen from the lot), and the *joyride* event (where your car is borrowed by the valet for some period of time). Furthermore, we might have some estimate of the length of time for which the valet might borrow the car, which could be modeled using a *return* event (the valet returns the car to the lot). One possible event model is illustrated in Figure 9, along with an initial belief state in which you are certain the car is in the lot with low mileage.

We consider several possible observations and updates of epistemic state one could make. Suppose at time 1 you unexpectedly return to the parking lot to pick up the car and find it missing. The most plausible explanation is the *steal* event, and you will believe the car has been *stolen* (as well, the mileage is high). If you make no such observation at time 1 (e.g., any observation actions taken—see Section 5.2 and 5.3—were unrelated to these propositions), then any of the models of GU that deal with iteration (as presented in Section 5) will leave you in the epistemic state (and associated belief state) given by the ranking at time 1.

Suppose you try to pick up the car at time 2 and find the car missing (that is, not parked). The most plausible explanation is that the car was stolen (but note there is no preference as to when the car was stolen, at time 0 or time 1). If the car is still in the lot, and you observe *Parked*, you will assume that nothing strange has happened. Should you also observe that the mileage is *High* at time 2, then you have no choice but to conclude that the car was joyridden and returned.[21]

Suppose, now, that after returning at time 2 to find the car parked (but not yet observing the mileage), you begin to drive home and notice *High* halfway home, at time 3. Your belief at time 2 that nothing unusual happened is now revised: again you must believe that the car was joyridden at time 0 at returned at time 1. Note that this event model asserts that if a car is being joyridden, the most likely event is the return of the car, with a continuation of the ride being unlikely. If you had found the car missing at time 2, and for some reason were convinced that it wasn't stolen, the most plausible explanation would be that it was joyridden at time 1, not at time 0: had it been taken at time 0, it would most likely have been returned by now. Thus, unlike the explanations of a stolen car (where no preference for the time of stealing is given), the joyride explanation comes with a preference for a more recent ride. This is due solely to the fact that our model of the process implicitly provides a duration. Should *continue* be just as likely as *return*, this chronological preference would disappear.

It is worthwhile to point out some of the demands made by the Markov and stationarity assumptions in an example like this. Note that we have added propositions like *joyridden* and *stolen* to the description of the world. In a certain sense, these propositions encode the history of past events. As such, we don't need to keep track of history explicitly, but only the relevant bits of history that we decide to represent in our propositional language. This makes a model based on epistemic states about the "current" state of affairs more attractive than a trajectory-based model: history can be accounted for, but one has the option of distinguishing only relevant historical facts as opposed to considering arbitrary distinctions between entire trajectories.

6.2 Distance-based Semantics

Recently, a model of belief revision based on a distance semantics was proposed [29]. Roughly, we assume the existence of a metric d over W that specifies the "distance" between possible worlds. Intuitively, $d(w, v)$ reflects the degree of difficulty or cost to change from situation w to v. One key assumption is that $d(w, v)$ is minimum exactly when $w = v$. Given a belief set K, and an observation

[21] To keep the diagram simple, we have "folded in" the effect of the "return to pick up car" action with the effects of the other actions (e.g., it causes *InCar* to hold if the car is present).

A, those A-worlds closest (according to d) to some element of $\|K\|$ determine the set K^*_A. Since d is fixed, this model allows for iteration as well.

It is shown in [29] that, under certain assumptions, such a semantics is equivalent to the AGM model of revision (that is, it satisfies the AGM postulates, and can represent any operator that satisfies the postulates). In one sense, this result may seem surprising, for the intuitions underlying the semantics are much like that of update semantics. Indeed, we might view the basic GU semantics as providing a distance metric that can be used in this regard: we can take $d(w, v)$ to be the plausibility of the most plausible transition from w to v. Under the assumption of inertness described in Section 5.1, we can see that $d(w, w)$ (for any w) is minimum. Of course, it was just this inertness assumption that allowed us to show that GU satisfied the AGM postulates (Proposition 5.1). However, as we argued there, this assumption is not the correct way to think about belief revision (as opposed to update). That it satisfies the AGM postulates is a sign of the weakness of these postulates, not of the suitability of inert models as a semantics for revision. As a result, the fact that a distance semantics can be made to satisfy the postulates does not mean revision can be thought of productively in these terms. We do note that distances do play a role in belief update however; we can view the relative plausibility of transitions as a form of distance between worlds.[22]

7 Concluding Remarks

We have provided a model for generalized belief update that extends both the classical update and revision models, combining the crucial aspects of both, and retaining both as special cases. The main feature of GU is its insistence one be allowed to both revise and update one's beliefs about the world in response to an observation. In addition, we have provided an abductive interpretation of update as the process of explaining observations in terms of what was initially true, what event or action may have occurred and the outcome that event may have had. We presented a model for dealing with noisy observations in belief revision, treating a problem that has been virtually ignored in the belief revision and reasoning about action communities. Finally, we have shown the strong connections between the GU model (especially as augmented with observation models) and the well-understood models of stochastic dynamical systems. Indeed, GU can be viewed as a qualitative form of Bayesian update, with the κ-calculus playing the role of probabilistic laws.

In this paper, we have focused exclusively on the semantics of generalized update. Appropriate

[22] Distances of this type are, however, not generally symmetric.

representation languages for the concise expression of events (with defeasible effects), defeasible beliefs and other aspects of the model must still be developed. The specification language proposed by Geffner [16] would seem to be a useful way of representing the system dynamics component of a GU model. Many of the other components of such languages are already in place, based primarily on conditional and dynamic logics, and other languages for actions and defeasible beliefs. However, a number of details regarding compact and natural representation languages and their feasibility are sure to require some subtlety. This undertaking is especially important when the ability to reason about incompletely specified systems is required.

One issue that has remained unexplored to a large extent is that of revising beliefs about system dynamics (event and outcome plausibilities). The GU model supposes that events and outcomes are specified independently of an agent's beliefs, and that the dynamics of the system in question are fixed and known. The same holds true for an agent's observation model. In general, however, one might expect an agent to have beliefs about these entities which are themselves subject to revision. While not inconsistent with our model, a more elaborate treatment requires a language in which (defeasible) *beliefs about* events, outcomes, and so on can be expressed.

One final stumbling block to a general treatment of qualitative belief change has to do with purely ranked models in which κ-like ranking information is not available. The κ values of different worlds and events provided us with a direct means of determining (for instance) the relative plausibility of various transitions by permitting the addition of the component κs. Models based on pure ranking information, without quantitative degrees of belief, are often used in belief revision (see, for instance, the semantics for AGM revision given in [20, 6] or the plausibility measures of [13]). To compare the relative plausibility of transitions in such a setting, we must have a way of trading off the relative likelihoods of initial conditions, events, event outcomes, and so forth. Simple ranking information does not allow one to do so; direct judgements of the plausibility of these combinations must be made. A more general qualitative theory would do just that—an example of such a theory is the plausibility measure approach recently proposed in [14]. The κ-approach presented here is a special case of this more general model. However, the essential spirit of such a proposal is identical to that underlying our presentation.

References

[1] Ernest W. Adams. *The Logic of Conditionals.* D.Reidel, Dordrecht, 1975.

[2] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[3] K. J. Astrom. Optimal control of Markov decision processes with incomplete state estimation. *J. Math. Anal. Appl.*, 10:174–205, 1965.

[4] Craig Boutilier. Epistemic entrenchment in autoepistemic logic. *Fundamenta Informaticae*, 17(1–2):5–30, 1992.

[5] Craig Boutilier. Revision sequences and nested conditionals. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 519–525, Chambery, FR, 1993.

[6] Craig Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68:33–85, 1994.

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

[8] Craig Boutilier. Abduction to plausible causes: An event-based model of belief update. *Artificial Intelligence*, 83(1):143–166, 1996.

[9] Craig Boutilier. Iterated revision and minimal revision of conditional beliefs. *Journal of Philosophical Logic*, 25(3):262–305, 1996.

[10] Peter E. Caines. *Linear stochastic systems*. Wiley, New York, 1988.

[11] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. In *Proceedings of the Fifth Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 5–23, Pacific Grove, 1994.

[12] Nir Friedman and Joseph Y. Halpern. A knowledge-based framework for belief change, part II: Revision and update. In *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning*, pages 190–201, Bonn, 1994.

[13] Nir Friedman and Joseph Y. Halpern. Plausibility measures: A user's guide. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 175–184, Montreal, 1995.

[14] Nir Friedman and Joseph Y. Halpern. A qualitative Markov assumption and its implications for belief change. In *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, pages 263–273, Portland, 1996.

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

[16] Hector Geffner. A qualitative model for temporal reasoning with incomplete information. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, pages 1176–1181, Portland, OR, 1996.

[17] Moisés Goldszmidt and Judea Pearl. On the consistency of defeasible databases. *Artificial Intelligence*, 52:121–149, 1991.

[18] Moisés Goldszmidt and Judea Pearl. Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and actions. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, pages 661–672, Cambridge, 1992.

[19] Moisés Goldszmidt and Judea Pearl. Reasoning with qualitative probabilities can be tractable. In *Proceedings of the Eighth Conference on Uncertainty in AI*, pages 112–120, Stanford, 1992.

[20] Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.

[21] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge database and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 387–394, Cambridge, 1991.

[22] Henry A. Kautz. The logic of persistence. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, pages 401–405, Philadelphia, 1986.

[23] Isaac Levi. Iteration of conditionals and the Ramsey test. *Synthese*, 76(1):49–81, 1988.

[24] William S. Lovejoy. A survey of algorithmic methods for partially observed Markov decision processes. *Annals of Operations Research*, 28:47–66, 1991.

[25] D. G. Luenberger. *Introduction to Dynamic Systems: Theory, Models and Applications*. Wiley, New York, 1979.

[26] Robert C. Moore. A formal theory of knowledge and action. In J. Allen, J. Hendler, and A. Tate, editors, *Readings in Planning*, pages 480–519. Morgan-Kaufmann, San Mateo, 1985. 1990.

[27] Hans Rott. Belief change using generalized epistemic entrenchment. *Journal of Logic, Language and Information*, 1(1):45–78, 1992.

[28] Richard B. Scherl and Hector J. Levesque. The frame problem and knowledge-producing actions. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 689–693, Washington, D.C., 1993.

[29] Karl Schlechta, Daniel Lehmann, and Menachem Magidor. Distance semantics for belief revision. In *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pages 137–145, Amsterdam, 1996.

[30] Richard D. Smallwood and Edward J. Sondik. The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 21:1071–1088, 1973.

[31] Wolfgang Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In W.L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change and Statistics*, volume 2, pages 105–134. D. Reidel, Dordrecht, 1987.

[32] Mary-Anne Williams. Transmutations of knowledge systems. In *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning*, pages 619–629, Bonn, 1994.

[33] Marianne Winslett. Reasoning about action using a possible models approach. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 89–93, St. Paul, 1988.

Acknowledgements

Thanks to Nir Friedman, Moisés Goldszmidt, Joe Halpern and David Poole for helpful discussions on this topic. This research was supported by NSERC Research Grant OGP0121843.

A Proofs of Main Results

Proposition 3.1 *result*$(A) = \min(A, \kappa^\diamond)$*; or, equivalently,* $K^\diamond_A = \{B : \min(A, \kappa^\diamond) \subseteq \|B\|\}$

Proof This follows almost immediately from the definitions of *result*(A) and $\kappa^\diamond$. If *Tr*$(A) = \emptyset$ then *result*$(A) = \emptyset$ and $\min(A, \kappa^\diamond) = \emptyset$. Otherwise,

$$\min(A, \kappa^\diamond) = \{v : w \stackrel{e}{\to} v \in \min(\mathit{Tr}(A))\} = \mathit{result}(A)$$

■

Proposition 4.1 *If $\diamond$ is the GU operator induced by some GU model then $\diamond$ satisfies postulates (U1), (U4), (U5), (U6), (U7) and (U9).*

Proof Assume an event model $M = \langle W, \kappa, E, \mu \rangle$ and associated update operator $\diamond$ (for simplicity we drop the subscript). We show in turn that each of these postulates is satisfied.

(U1) By definition, *result*$(A) \subseteq \|A\|$. Immediately we have *KB* $\diamond A \models A$.

(U4) Suppose $\models A \equiv B$. Then *Tr*$(A) =$ *Tr*(B) and *result*$(A) =$ *result*(B). Thus, *KB* $\diamond A \equiv$ *KB* $\diamond B$.

(U5) Let $w \stackrel{e}{\to} v$ be a most plausible A-transition such that $v \models B$. (If no such v exists, $(\mathit{KB} \diamond A) \wedge B$ is unsatisfiable and (U5) holds trivially.) Clearly, $w \stackrel{e}{\to} v$ is a most plausible $A \wedge B$-transition, since a strictly more plausible $A \wedge B$-transition (which is an A-transition) contradicts the fact that $w \stackrel{e}{\to} v$ is most plausible. Thus, *result*$(A) \cap \|B\| \subseteq$ *result*$(A \wedge B)$ and $(\mathit{KB} \diamond A) \wedge B \models \mathit{KB} \diamond (A \wedge B)$.

(U6) Suppose *KB* $\diamond A \models B$ and *KB* $\diamond B \models A$. Then we have *result*$(A) \subseteq \|B\|$ and *result*$(B) \subseteq \|A\|$. If $v \in$ *result*(A), there must exist a most plausible A-transition of the form $w \stackrel{e}{\to} v$ (where $v \models A$). However, since $v \in \|B\|$, $w \stackrel{e}{\to} v$ is a B-transition as well. If $w \stackrel{e}{\to} v \notin \min(\mathit{Tr}(B))$, there must exist a most plausible B-transition $w' \stackrel{e'}{\to} v'$ such that $\kappa(w' \stackrel{e'}{\to} v') < \kappa(w \stackrel{e}{\to} v)$; but since *result*$(B) \subseteq \|A\|$, this contradicts the fact that $w \stackrel{e}{\to} v \in \min(\mathit{Tr}(A))$. Hence, $v \in$ *result*(B) and *result*$(A) \subseteq$ *result*(B). A similar argument shows *result*$(B) \subseteq$ *result*(A); hence *KB* $\diamond A \equiv$ *KB* $\diamond B$.

(U7) Suppose $v \in$ *result*$(A) \cap$ *result*(B). (If there is no such v then $(\mathit{KB} \diamond A) \wedge (\mathit{KB} \diamond B)$ is inconsistent and (U7) holds trivially.) Then there is some e such that $w \stackrel{e}{\to} v$ is a most plausible A-transition and a most plausible B-transition. This ensures that $w \stackrel{e}{\to} v$ is a most plausible $A \vee B$-transition and that $v \in$ *result*$(A \vee B)$. Hence, *result*$(A) \cap$ *result*$(B) \subseteq$ *result*$(A \vee B)$. Therefore, $(\mathit{KB} \diamond A) \wedge (\mathit{KB} \diamond B) \models \mathit{KB} \diamond (A \vee B)$. (Note that this proof does not require that *KB* be complete, in contrast to the conditions put on (U7) in the KM postulates.)

(U9) Let *KB* be complete with $||KB|| = \{w\}$. Suppose *KB* $\diamond A \not\models \neg B$ and *KB* $\diamond A \models C$. Then *result*$(A) \subseteq ||C||$ and *result*$(A) \cap ||B|| \neq \emptyset$. Now if $v \in$ *result*$(A \wedge B)$, then $w \stackrel{e}{\to} v$ must be a most plausible A-transition (since $v \models A$ and any less plausible transition is dominated by $w \stackrel{e'}{\to} v'$ for some $v' \in$ *result*$(A) \cap ||B||$). Hence, *result*$(A \wedge B) \subseteq$ *result*$(A) \cap ||B|| \subseteq ||C||$ and *KB* $\diamond (A \wedge B) \models C$.

■

Proposition 4.2 *If $\diamond$ is induced by an inert GU model then $\diamond$ satisfies (U2).*

Proof Assume that $M = \langle W, \kappa, E, \mu \rangle$ is an inert model, inducing update operator $\diamond$ (for simplicity we drop the subscript). Suppose *KB* $\models A$. Then $\{w \stackrel{n}{\to} w : w \in ||KB||\}$ forms the set of most plausible A-transitions (where n is the null event). Thus *result*$(A) = ||KB||$ and *KB* $\diamond A$ is equivalent to *KB*; (U2) is satisfied. ■

Proposition 4.3 *If $\diamond$ is induced by a complete GU model then $\diamond$ satisfies (U3).*

Proof Assume that $M = \langle W, \kappa, E, \mu \rangle$ is an inert model, inducing update operator $\diamond$ (for simplicity we drop the subscript). Since M is complete, any satisfiable A is explainable (i.e., *Tr*$(A) \neq \emptyset$). So if *KB* is satisfiable, *result*$(A) \neq \emptyset$ and *KB* $\diamond A$ is satisfiable; (U3) is satisfied. ■

Theorem 4.5 *If $\diamond$ is induced by a static GU model then $\diamond$ satisfies (R1)–(R8).*

Proof Let $M = \langle W, \kappa, E, \mu \rangle$ be a static model, inducing update operator $\diamond$ (for simplicity we drop the subscript). Since M is static, we have $\kappa = \kappa^{\diamond}$ (i.e., the predicted updated ranking is identical to the original). By Proposition 3.1,

$$K_A^{\diamond} = \{B : \min(A, \kappa^{\diamond}) \subseteq ||B||\}$$

Thus

$$K_A^{\diamond} = \{B : \min(A, \kappa) \subseteq ||B||\}$$

The standard AGM representation results for κ-rankings ensure the revision postulates are satisfied. (Note: we assume that κ is total; i.e., $\kappa(w) < \infty$ for all w.) ■

Proposition 5.1 *If $\diamond$ is induced by an inert GU model then $\diamond$ satisfies (R1)–(R8).*

Proof Let $M = \langle W, \kappa, E, \mu \rangle$ be an inert model, inducing update operator $\diamond$ (for simplicity we drop the subscript). By Proposition 3.1,

$$K_A^\diamond = \{B : \min(A, \kappa^\diamond) \subseteq \|B\|\}$$

The fact that M is inert ensures that $\min(\kappa^\diamond) = \min(\kappa)$ (i.e., the most plausible worlds–those determining the belief set K–are identical in the original and updated rankings). Since for any κ (and induced K), the updated ranking $\kappa^\diamond$ is a revision model for K, the standard AGM representation results ensure the revision postulates are satisfied. (Note: we assume that κ is total; i.e., $\kappa(w) < \infty$ for all w.) ■