Generalized Update: Belief Change in Dynamic Settings

Craig Boutiller
Department of Computer Science
University of British Columbia
Vancouver, BC V6T 1Z4, CANADA
cebly@cs ubc ca
hftp /Avwwcs ubc ca/spider/cebly/craig html

Abstract

Belief revision and belief update have been proposed as two types of belief change serving different purposes Belief revision is intended to capture changes of an agent's belief state reflecting new information about a static world Belief update is intended to capture changes of belief in response to a changing world We argue that both belief revision and belief update are too restrictive, routine belief change involves elements of both We present a model for *generalized update* that allows updates in response to external changes to inform the 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, in a way that formally verifies the intuition that revision is suitable for "static" belief change

1 Introduction

An underlying premise in much work addressing the design of intelligent agents or programs is that such agents should 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

Theories of belief change have received considerable attention in recent years in the AI community 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 Alchourron, Gardenfors and Makinson (1985, 1988) 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, 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 (1988) and Kalsuno and Mendelzon (1991)

In this paper, we describe a semantic model for belief change that generalizes belief update to incorporate aspects of belief revision The aim of this model is twofold (a) to provide a unifying semantics for both revision and update that highlights the orthogonal roles both have to play in routine belief change and (b) 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

There have been attempts to provide general semantics for belief change operators (e g , (Friedman and Halpern 1994)) but often these models are such that *under certain assumptions* the change is a revision and under others it is an update We argue that routine belief change should involve both update and revision and develop a model that incorporates aspects of both but we show that revision and update as currently conceived, are special cases of our general operator

The result of this union is a more robust and realistic notion of update in which observations of change can inform and 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 (Grove 1988, Gardenfors 1988, Boutilier 1994c) To this we add the notion of a *transition* or evolution from one world state to another As proposed by Katsuno and Meodelzon (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 (Bountiher 1994a), 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 an initial condition, 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 condition, 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

In Section 2 we present the AGM theory of revision and the KM theory of update, emphasizing the semantic models that have been proposed and adopting the qualitative probabilistic model of (1987, 1992) 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 describe the formal relationship between revision, update and GU We show that under certain assumptions GU satisfies 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

2 Classical Belief Revision and Belief Update

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 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) $K^*_A \subseteq K^+_A$

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

(R5) $K^*_A = Cn(\bot)$ iff $\models \neg A$

(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* (Gärdenfors 1988) 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 (Grove 1988, Boutilier 1994c) However rather than use a qualitative ranking relation, we adopt the presentation of (Spohn 1987, Goldszmidt and Pearl 1992) and rank all possible worlds using a *κ-ranking* Such a ranking κ $W \to \mathbf{N}$ 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 (Spohn 1987) 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 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*

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

The ranking κ also induces a revision function to revise by A an agent adopts the most plausible 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 (Grove 1988, Boutilier 1994c, Goldszmidt and Pearl 1992)[1]

The ranking function k can naturally be interpreted as charactenzing 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 4 is observed

As an example, consider the ranking shown in Figure 1 (a), which reflects the epistemic state of someone who believes herbook 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)-worid (K = 1) also satisfies *Inside{G)* she strongly believes she left her book and glasses IN the same place

22 Belief Update

Katsuno and Mendelzon (1991) have proposed a general charactenzation 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 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 *KBoA* denotes the new belief set incorporating this change The *KM postulates* governing admissible update operators are

(U1) $KB \circ A \models 4$

(U2) If $KB \models A$ then $KB \circ A$ is equivalent to KB

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

(U4) If $\models A \equiv B, KB_1 \equiv KB_2$ then $KB_1 \circ A \equiv KB_2 \circ B$

(U5) $(KB \circ A) \wedge B \models KB \circ (A \wedge B)$

(U6) If $KB \circ A \models B$ and $KB \circ B \models A$ then $KB \circ A \equiv KB \circ B$

(U7) If KB is complete then $(KB \circ A) \wedge (KB \circ B) \models KB \circ (A \vee B)$

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

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 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 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,

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

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

If the orderings $\leq_w$ are *total* preorders (so that all elements are comparable), then update operators are characterized by (U1)–(U9) (see (Katsuno and Mendelzon 1991, Boutilier 1994a))

(U9) If KB is complete, $(KB \circ A) \not\models \neg B$ and $(KB \circ A) \models C$ then $(KB \circ (A \wedge B)) \models C$

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

As a concrete example, suppose that someone observes that the grass IN from of her bouse *is* weL Prior to the observation she believed that she left her book outside on the patio and that the grass and book were dry (seeKBinFigure 1(b)) 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 *wer G,* she would have accepted *this* explanation (and its consequences, such as her glasses being dry if they are with her book)

3 Generalized 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 ofbeliefcan influence updated beliefs Even simple tasks such as modeling *information gathering actions* are beyond the scope of KM update Consider, for example, Moore's (1985) litmus test the contents of a beaker are unknown and one dips bonus paper into it to determine if it is an acid or a

---

[1] We refer to (Boutilier 1994c, Friedman and Halpern 1994) for a discussion of languages with which one can express properties of belief sets and revision functions

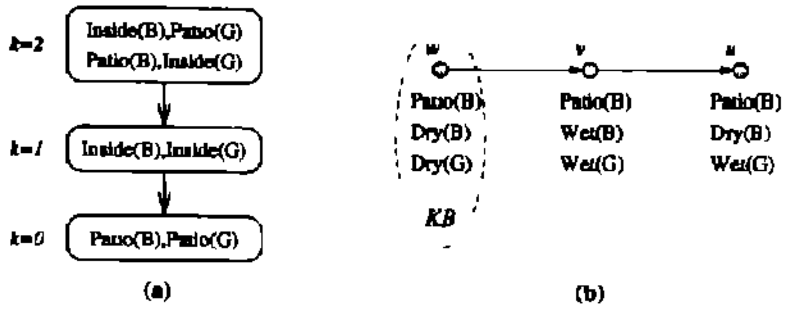

Figure 1 (a) A Revision Model and (b) An Update Model

base The prior state of belief consists of two possible worlds (*acid* and *base)* and 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)[2]

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 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 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

Rather than generalizing the KM update semantics directly we adopt the approach of (Boutilier 1994a), where we argued 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 The motivation for this approach is 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 We extend this idea further by supposing that events are nondeterministic and that their possible outcomes can also be ranked 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

We assume a set of events E An *event* maps each world into a partial κ-ranking over worlds, $e \quad W \rightarrow (W \rightarrow \mathbf{N})$ 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, thus $\kappa_{w,e}$ only ranks the possible outcomes of e We call this evolution of w into v a *transition* which we write $w \xrightarrow{e} v$ 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

Since an agent making an observation will often not know *a priori* what event caused an observation, 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, $\mu \quad W \rightarrow (E \rightarrow \mathbf{N})$, 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

We now describe *generalized update*

Def 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 world (κ) what is likely to happen or not (μ), and the effects of those occurrences (E) Such models contain the information necessary to update K in response to an observation A, we denote the resulting belief set K_A°

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 $w \xrightarrow{e} v$ depends on the plausibility of w, the likelihood that e occurred, and the likelihood of outcome v given w, e In other words [3]

$$\kappa(w \xrightarrow{e} v) = \kappa_{w,e}(v) + \kappa_w(e) + \kappa(w)$$

With this in hand, an updated ranking κ° can be given by

$$\kappa^\circ(v) = \min_{w \in W\, e \in E} \{\kappa_{w,e}(v) + \kappa_w(e) + \kappa(w)\}$$

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

---

[2] 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 (Boutilier 1994a)

[3] We note that this formula is the qualitative analog of the probabilistic equation $Pr(w \xrightarrow{e} v) = Pr(v|w,e) \quad Pr(e|w) \quad Pr(w)$ We refer to (Spohn 1987 Goldszmidt and Pearl 1992) for details on the relationship between qualitative and quantitative probabilities

contained in the model M) We note that the agent's actual beliefs are determined by the minimal worlds in K° (1 e , those v such that $\kappa^\circ(v) = 0$)

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

$$Tr(A) = \{w \xrightarrow{e} v \quad v \models A \text{ and } \kappa(w \xrightarrow{e} v) \neq \infty\}$$

The *most plausible A transitions,* denoted *min (Tr(A)),* are those possible A-transitions with the minimal k 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 these most plausible transition

$$result(A) = \{v \quad w \xrightarrow{e} v \in \min(Tr(A))\}$$

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

$$K^\circ_A = \{B \quad result(A) \subseteq \|B\|\}$$

In other words, an agent updating by observation *A* believes what is true at the states that result from the most plausible A-transitions We also have the following

Prop 1 $result(A) = \min(A, \kappa^\circ)$, *or (equiv)* $K^\circ_A = \{B$ mm$(A, \kappa^\circ) \subseteq \|B\|\}$

This conforms to our intuitions about the updanng process the direct update of A' by A, K°_A determines the same belief set as the process of first updating one's entire epistemic slate *K* to get κ°., and theD performing belief *revision* of κ° by the observation *A* Loosely, we might say $(K^\circ)^*_A = K^\circ_A$

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* An *explanation* for *A* (given model M) is any triple (w,e,i/) such that $w \xrightarrow{e} v \in Tr(A)$ (which implies $\kappa(w \xrightarrow{e} v) < \infty$) Thus it is possible that e occurred at *w,* leading to *v* and resulting in *A* The *most plausible explanations* for *A* are those explanations with minimal K-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 GU can be interpreted as an abductive process Note, however that Proposition 1 means we not generate explanations explicitly

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 only one possible evenL

Figure 2(a) illustrates the prior belief state of an agent who believes her book is on the patio *(P)* 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 ($\kappa(In) = 1$) 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 $\kappa_w(null) = 0$, $K_w(rain) = 1$ and $\kappa_w(sprinkler) = 2$, so *ram is* more plausible than *sprinkler* (we assume a "global" ordering, suitable for all *w)* The outcomes of these events are deterministic — in particular both nun 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 *wetG* A consequence of this is that she will now believe her book is wet the most likely explanation is simply that it rained If *wetG* A *dryB* 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 , k*(wetG* A *dryB)* = 2) This corresponds to the existence of two plausible explanations either the book *is* on the patio (K = 0) and the sprinkler turned on (k = 2), or the book is inside (k = 1) and it rained (K = 1) The result is that the agent is no longer sure where the book is If we had instead set *K(sprinkler)* = 3. observing *wetG* A *dryB* would have caused the agent to believe that the book bad 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)[4]

A second example is shown in Figure 2(b) 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 (K = 0), little plausibility *(K* = 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 kryptomte 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 *(K = 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 epistermically possible worlds As such, it is an appropriate model for revision and expansion of beliefs due to information-gathering actions If an outcome of green presents two competing explanations either the test failed (the substance is an acid or a base) or the beaker contains kryptomte The most plausible explanation and the updated belief state depend on the relative magnitudes of *g* and *T* The figure suggests that *g* < r, so the a test failure is most plausible and the belief *acid* V *base* is retained. If test failures are more rare (r < *g),* then this outcome would cause the agent to believe the beaker held kryptonite

4 Relationship to Revision and Update

The analysis of the update postulates is similar to that presented in (Boutilier 1994a) There we described a model of update that used plausible events to explain the occurrence of observations, giving rise to an update operator Only under

---

'The world *In DryB WetG* at *K* - 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 observauons can distinguish the two worlds

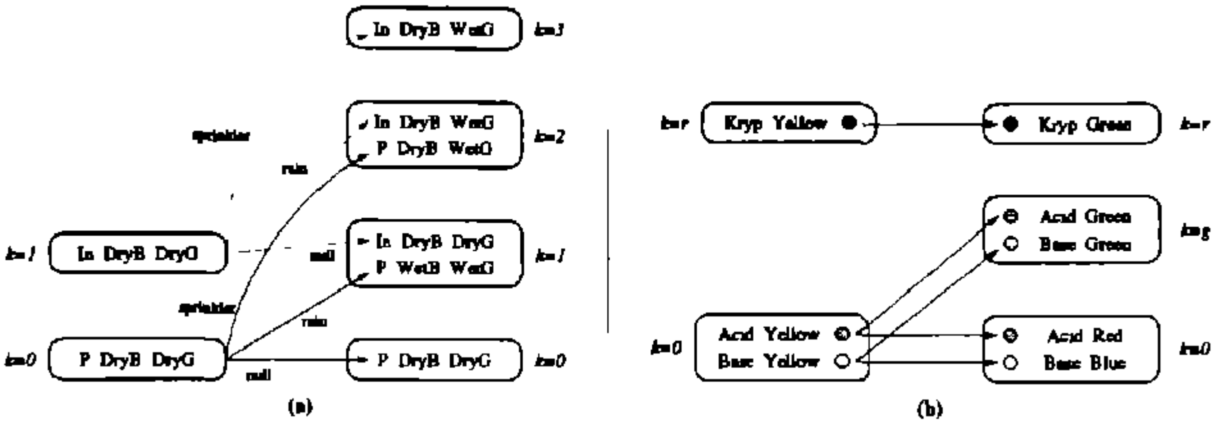

Figure 2 Generalized Update with Multiple (a) Events and (b) Outcomes

certain assumptions docs this operator satisfy the KM postulates, and 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 Surprisingly this has little bearing on the update postulates the same assumptions are required We describe these briefly and refer to (Boutilier 1994a) for further discussion We first note that our model satisfies a number of the KM postulates

Prop 2 If *is the GU operator induced by some GU model then* o *satisfies postulates (Ul) (U4) (V6), (U7)and(U9)*

One key difference between the GU model and the KM model is reflected in (U2) which asserts that *KBo A* is equivalent to *KB* whenever *KB* entails *A* This cannot be the case in general, for even if *KB \= 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 (U2) is appropriate only if we are willing to assume *persistence* of propositions, [hat 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 *centered update models*

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

Prop 3 *If $\diamond$ is induced by a centered GU model then $\diamond$ satisfies (U2)*

The second key difference is reflected in the failure of (U3) (both (U5) and (U8) fail for related reasons), 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 $K_A^{\diamond} = \mathrm{L}_{CPL}$, the inconsistent belief set. To prevent this, we can simply insist that every satisfiable sentence A is explainable

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

Prop 4 *If $\diamond$ is induced by a complete GU model then $\diamond$ satisfies (U3), (U5) and (U8)*

In (Boutilier 1994a) 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*

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 (Boutilier 1994a) In our setting, one quite reasonable condition we might impose is that ail worlds have some plausibility (I e , k is a total function on *W)* and that the null event is possible (not necessarily plausible) at each of those The first requirement is usually assumed of epistemic states, 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 3 and 4 together we have

Thm 5 *If* o *is induced by a complete, centered GU model then* o *satisfies (U1)(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 ordenngs, 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 K-rankings, which totally order events and worlds One of the main reasons for using such rankings 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 m the more general setting however

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 K to rank worlds only as plausible (k = 0) or impossible (k = ∞) 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 /(-ranking $K_{W\,a}$ 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]

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 (he world is static or unchanging To verify this intuition, we propose *static update models*

Def. **An update model** $M = \langle W, \kappa, E, \mu\rangle$ **is** ***static*** **if** $E = \{n\}$ **where** n **is the null event** n[6]

Thm. *6 If 0 is induced by a static GU model then* o *satisfies (R1)-(R8)*

Static event models have as the only possible transitions those of the form u' ->w *w* with plausibility *K(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 eastly verified)

5 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 observanon

In this paper, we have focussed exclusively on the semantics of generalized update Appropriate representation languages for the concise expression of events (with defeasible effects), defeasible beliefs and other aspects of the model must soil be developed However, the many components of such languages are already in place, based primarily on conditional and dynamic logics, and other action languages

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 are static In general, however, one might expect an agent to have beliefs about these entities which are 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

Another crucial Issue Is that *iterated updates* that arise with sequences of events and observations, this introduces several complications One is bow to revise an epistemic state k (rather than a belief set A") in response to an observation, several proposals exist for iterated revision (Spohn 1987, Boutilier 1994b, Williams 1994) but their applicability to this problem remains to be verified A related problem is that the plausibibty of a sequence of transitions need not be a function of the individual transitions, as discussed in (Friedman and Halpern 1994), more sophisticated update *criteria* are required, including judging the plausibibty of sequences of transitions as a whole If such a general semantic picture is titled with a language with which to reason about events, we should be able to recast tbe GU model as a form of belief revision about such "histories " Thus, the general view of explanation as a form of belief revision (Gardenfors 1988, Bouulier and Becher 1994) can be extended to the explanation of observations in dynamic systems

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

References

Alchourron C Gardenfors P and Makinson D 1985 On the logic of theory *change* Partial meet contraction and revisioo functions *J Symbolic Logic* 50-510-530

Bouulier, C 1994a. Abduction to plausible causes An event based model of belief update *Artificial Intelligence* (in press - see also An Event-Based Abductive Model of Update *CSCSI 94* pp 241-248 Banff)

Bouulier C 1994b Iterated revision and minimal revision ofconditional beliefs. / *Philosophical Logic* (in press - sec also Revision Sequences and Nested Conditionals *IJCA1 93* pp.519—525 Chambery)

Bouulier C 1994c Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence* 68 33-85

Bouulier C and Becher V 1994 Abduction as belief revision *Artificial Intelligence* (in press)

Friedman N and Halpern I Y 1994 A knowledge-based frame work for belief change part II Revision and update *KR 94* pp 190-201 Bonn

Gardenfors, P 1988 *Knowledge in Flux Modeling the Dynamics of Epistemic States* MIT Press, Cambridge

Goldszmidt, M and Pearl J 1992 Rank based systems A simple approach to belief revision, belief update and reasoning about evidence and acuons, *KR 92* pp 661-672 Cambridge

Grove A. 1988 Two modellings for theory change *Journal of Philosophical Logic* 17 157-170

Katsuno H and Mendelzon A0 1991 On the difference between updating a knowledge database and revising n_ *KR 91* pp 387-394, Cambridge

Moore, R. C 1985 A formal theory of knowledge and action Allen, J Hendler J and Talc A. eds *Readings in Planning,* pp480-519 Morgan Kaufmann 1990

Spohn W 1987 Ordinal conditional functions A dynamic theory of epistemic states Harper W, Skyrms B eds *Causaaonin Decision, Belief Change and Statistics* pp 105-134 D Reidel

Williams, M-A. 1994 Transmutations of knowledge systems *KR 94* pp 619-629 Bonn

Winslett, M 1988 Reasoning about action using a possible models approach *AAA! 88* pp 89-93 St. Paul

---

Concurrent events and actions require special attention however, and are beyond (he scope of this paper

[6] As above we assume *K* IS a total function on *W*