IS BELIEF REVISION HARDER THAN YOU THOUGHT?

Marianne Winslett*
Stanford University, Computer Science Dept.
Stanford CA 94305

Abstract. Suppose one wishes to construct, use, and maintain a knowledge base (KB) of beliefs about the real world, even though the facts about that world are only partially known. In the AI domain, this problem arises when an agent has a base set of beliefs that reflect partial knowledge about the world, and then tries to incorporate new, possibly contradictory knowledge into this set of beliefs. We choose to represent such a KB as a logical theory, and view the models of the theory as representing possible states of the world that are consistent with the agent's beliefs.

How can new information be incorporated into the KB? For example, given the new information that "b or c is true," how can one get rid of all outdated information about b and c, add the new information, and yet in the process not disturb any other information in the KB? The burden may be placed on the user or other omniscient authority to determine exactly what to add and remove from the KB. But what's really needed is a way to specify the desired change intensionally, by stating some well-formed formula that the state of the world is now known to satisfy and letting the KB algorithms automatically accomplish that change.

This paper explores a technique for updating KBs containing incomplete extensional information. Our approach embeds the incomplete KB and the incoming information in the language of mathematical logic. We present semantics and algorithms for our operators, and discuss the computational complexity of the algorithms. We show that the incorporation of new information is difficult even without the problems associated with justification of prior conclusions and inferences and identification of outdated inference rules and axioms.

1. Introduction

This section informally describes our view of the knowledge base component of an intelligent reasoning agent, and reviews the phases of belief revision required when the agent makes new observations about the world, concluding with an outline of the remainder of the paper.

We envision the KB for an agent as containing, among other items, a set of *extensional beliefs* whose contents are not generated via inference from other data but rather by direct observation or input, and whose justification is therefore simply the fact of direct input. One can imagine these primitive beliefs as stemming from processed sensory input, such as a scene analyzer that uses a line and shadow drawing as its primitive depiction.

In addition, the agent's knowledge base may contain derived or *intensional* beliefs. Intensional beliefs are derived from extensional and intensional beliefs via data-independent logical *inference rules* such as modus ponens, and data-dependent *axioms*, such as functional dependencies. For example, in the blocks world, a typical axiom might say that any block is either on the table or on another block. An extensional belief may also have an intensional justification; for example, in the blocks world an agent may directly observe that block A is on the table and also be able to deduce that fact from the observation that block A is not on top of any other block.

Belief revision comes in a number of guises. First, new extensional information may lead to a change in extensional beliefs. These new beliefs may in turn trigger changes in intensional and previously extensional beliefs, and in axioms. For example, the placement of a new block so as to block vision of the rest of a scene may cause uncertainty about the placement of the objects hidden by the new block, or may cause the beliefs about hidden objects to become intensional (i.e., justified solely by frame axioms) rather than extensional, depending upon the inferences employed by the agent. This process of belief justification, or reconsideration of intensional beliefs, has been studied extensively (see e.g. [Doyle 79]).

Another variety of belief revision occurs when new information implies that the current set of axioms is now incorrect (typically, when the new theory has no

* AT&T Bell Laboratories Scholar

models). There is no established technique for revamping axiom sets, and this is an open area of research [Fagin 83, 84].

What is less well recognized is that the first stage of belief revision, incorporating new extensional beliefs into the pre-existing set of extensional beliefs, is itself quite difficult if either the new or old beliefs involve incomplete information. In this paper we focus entirely on this problem of adding new extensional knowledge to the set of extensional beliefs, and do not consider the activities of subsequent stages of belief revision such as making new inferences, checking belief justifications, or revoking old axioms. In addition, we do not consider here the problems associated with the presence of intensional beliefs and interaction of intensional and extensional beliefs when new extensional information is presented to the agent; the interested reader is referred to [Winslett 86b] for a discussion of these topics. Finally, we do not consider the problem of extracting information from the KB; for that purpose, we suggest operations such as Levesque's **ASK** [Levesque 84].

We consider extensional beliefs that have incomplete information in the form of disjunctions or *null values*, which are attribute values that are known to lie in a certain domain but whose value is currently unknown; see e.g. [Imielinski 84], [Reiter 84]. Levesque [84] considered the problem of updating such KBs with his TELL operation; however, TELL could only eliminate models from the set of models for the theory, not change the internal contents of those models. In other words, one could only TELL the KB new information that was consistent with what was already known. This is an important and vital function, but an agent also needs to be able to make changes in the belief set that contradict current beliefs. For example, the agent should be able to change the belief that block A is on block B if, for example, the agent observes a robot arm removing A from B.

In recent work on circuit diagnosis and updating propositional formulas, DeKleer [85] and Reiter [85] take a logic theory describing the correct behavior of a circuit, and consider the problem of making minimal changes in that theory in order to make it consistent with a formula describing the circuit's observed behavior. (Weber [86] does not focus on circuit diagnosis, but takes a similar approach with a slightly different choice of semantics.) One cannot expect to find a polynomial–time algorithm for diagnosis, as the changes needed in the theory—the diagnosis—are themselves the output of the diagnostic process, and the determination of whether any changes are needed at all in a propositional theory—i.e., whether the circuit is functioning correctly—cannot in general be done in polynomial time. However, in Section 3 we present a polynomial–time approach that may be used when only the new "diagnosed" theory is of interest, rather than the changes made to the old theory.

In the database realm, the problem of incorporating new information was considered by Abiteboul and Grahne [Abiteboul 85], who investigate the problem of simple updates on several varieties of relations containing null values and simple auxiliary constraints. They do not frame their investigation in the paradigm of mathematical logic, however, making their work less applicable to AI needs.

In the remainder of this paper, we set forth a facility for incorporating new information into KBs of extensional information. Section 2 introduces extensional belief theories, a formalization of such KBs. Section 3 sets forth a language for adding new information to these theories, and gives syntax, semantics, and a polynomial–time algorithm for a method of adding new information. In [Winslett 86b], the algorithm is proven correct in the sense that the alternative worlds produced under this algorithm are the same as those produced by updating each alternative world individually.

2. Extensional Belief Theories

The language $\mathcal{L}$ for *extensional belief theories,* our formalization of extensional belief sets, contains the usual first–order symbols, including an infinite set of variables and constants, logical connectives, the equality predicate, and quantifiers. $\mathcal{L}$ also includes the truth values T and F, and an infinite set of Skolem constants ϵ, ϵ_1, ϵ_2, ... Skolem constants are the logical formulation of null values and represent existentially quantified variables; we assume that the reader is acquainted with their use. $\mathcal{L}$ does not contain any functions other than constants and Skolem constants, and all constants are standard names.

$\mathcal{L}$ also contains a set of purely extensional predicates (e.g., OnTopOf(), Red()), that is, predicates for which the agent's belief justification is always the fact of direct observation, rather than inference from other beliefs. (In [Winslett 86b] we also consider predicates with intensional aspects.) In addition, $\mathcal{L}$ includes one extra *history* predicate H_R for each extensional predicate R. History predicates are for internal use of the algorithm implementations only; the agent is unaware of their existence.

Unlike the language used by Levesque [84], $\mathcal{L}$ does not include a **K** operator to refer explicitly to the KB's knowledge about the world. The main use of the **K** operator in the TELL operation is to add bits of a closed–

world assumption to the belief set, and we have devised a separate technique for performing that function and for maintaining the closed–world assumption when new information arrives, as described in [Winslett 86].

Definition. A theory $\mathcal{T}$ over $\mathcal{L}$ is an *extensional belief theory* iff

(1) for each pair of constants c_1, c_2 in $\mathcal{L}$, $\mathcal{T}$ contains the *unique name axiom* $c_1 \neq c_2$;

(2) the remainder of $\mathcal{T}$ (its *body*) is any finite set of well–formed formulas of $\mathcal{L}$ that do not contain variables.

Though the models of $\mathcal{T}$ look like possible states of the world consistent with the agent's extensional observations, not everything in a model of $\mathcal{T}$ is an instantiated consequence of extensional beliefs of the agent, due to the presence of history predicates. For this reason we define an *alternative world* of $\mathcal{T}$ as a model of $\mathcal{T}$ minus the equality and history predicates.

3. An Operation For Specifying New Beliefs

We now present an operation for specifying new extensional beliefs or *observations*, based on the language $\mathcal{L}$.

3.1. Observation Syntax

Let ϕ and ω be formulas over $\mathcal{L}$ without history predicates or variables. Then an *observation* takes the form IF ϕ THEN ω.

Examples. Suppose $\mathcal{L}$ contains two predicates, Red() and OnTopOf(), and B, C, and Table are constants in $\mathcal{L}$, presumably denoting blocks and tables. Then the following are observations, with their approximate intended semantics (presented formally in the next section) offered in italics:

IF T THEN $\neg$Red(B) $\vee$ (OnTopOf(B, ϵ) $\wedge$ ($\epsilon \neq$ Table)). *Change alternative worlds so that either B isn't red or it's on top of something other than the table.*

IF $\neg$Red(B) THEN F. *Eliminate all alternative worlds where B is not red.*

IF OnTopOf(B,ϵ) $\vee$ OnTopOf(ϵ,B) THEN $\neg$Red(ϵ) $\wedge$ ($\epsilon \neq$ C). *In each alternative world where B is on top of or below something, change the alternative world so that that something isn't red and isn't C.*

3.2. Semantics

We define the semantics of an observation applied to an extensional belief theory $\mathcal{T}$ by its desired effect on the alternative worlds of $\mathcal{T}$. In particular, the alternative worlds of the new theory must be the same as those obtained by applying the observation separately to each original alternative world. This may be rephrased as follows: Extensional beliefs with incomplete information represent a possibly infinite set of alternative worlds, each different and each one possibly representing the real, unknown world. The correct result of incorporating new information into the KB is that obtained by storing a separate extensional belief theory with complete information for each alternative world and processing the observation in parallel on each separate KB. A necessary and sufficient guarantee of correctness for any more efficient and practical method of observation processing is that it produce the same results as the parallel computation method. Equivalently, we require that the diagram below be commutative: both paths from upper-left-hand corner to lower-right-hand corner must produce the same result.

$$\begin{tikzcd} \mathcal{T} \arrow[r, "\text{has alternative world}"] \arrow[d, "\text{observation}"'] & \mathcal{A} \arrow[d, "\text{observation}"] \\ \mathcal{T}' \arrow[r, "\text{has alternative world}"] & \mathcal{A}' \end{tikzcd}$$

The general criteria guiding our choice of semantics are, first, that an observation cannot directly change the truth valuations of any atomic formulas (*atoms*†) except those that unify†† with atoms of ω. For example, the observation IF T THEN Red(A) cannot change the color of any block but A, and cannot change the truth valuation of formulas such as Green(A). (Of course, after the first stage of belief revision has incorporated the extensional fact that A is red, in a second stage the axioms and rules of inference may be used to retract the belief that A is green and any other outmoded fancies. As noted earlier, in this paper we only consider the first stage of belief revision.)

The second criterion is that the new information in ω is to represent the *most exact and most recent state of extensional knowledge obtainable* about the atoms in ω, and is to *override all previous extensional information* about the atoms of ω. These criteria have a syntactic component: one should not necessarily expect two observations with logically equivalent ω s to produce the same results. For example, if the agent observes IF T

† In this discussion, atoms may contain Skolem constants.

†† In this formulation, two atoms *a* and *b* unify if there exists a substitution of constants and/or Skolem constants for the Skolem constants of *a* and *b* under which *a* and *b* are syntactically identical.

THEN T, this is different from observing IF T THEN Red(A)$\vee\neg$Red(A); one observation reports no change in the redness of A, and the other explicitly points out that whether A is red is now unknown. This syntactic property of the semantics may seem unusual at first, though a short acquaintance should prove sufficient to demystify it.[†]

Under these desiderata, the observation IF ϕ THEN ω is not equivalent to IF T THEN $\phi \rightarrow \omega$; the second observation allows the truth valuations of atomic formulas in ϕ to change, while the first does not. For example, if the agent says something like IF *leprechauns exist* THEN *there's one outside my door now,* this statement will not change any truth valuations about leprechauns, merely observe that if they do exist then there's one hanging around now. On the other hand, with IF T THEN *if leprechauns exist, then there's one outside my door,* the agent opens up the possibility of leprechauns even if the agent previously did not believe in them. The ability to make changes in the theory dependent upon an unknown bit of information, without affecting the truth or falsity of that information, is crucial.

An intuitive motivation is in order for our method of handling Skolem constants. The essential idea is that if the agent only had more information, the agent would not be making an observation containing Skolem constants, but rather an ordinary observation without Skolem constants. Under this assumption, the correct way to handle an observation O with Skolem constants is to consider all the possible Skolem-constant-free observations represented by O and execute each of those in parallel, collecting the models so produced in one large set. Then the result of the observation the agent would have had, had more information been available, is guaranteed to be in that set.

For a formal definition of semantics that meets the criteria outlined in this section, let O be an observation and let $\mathcal{M}$ be a model of an extensional belief theory $\mathcal{T}$. Let σ be a substitution of constants for exactly the Skolem constants of ϕ, ω, and $\mathcal{T}$, such that $\mathcal{M}$ is a model of $(\mathcal{T})_\sigma$, that is, a model of the theory resulting from applying the substitution σ to each formula in $\mathcal{T}$. [††] Then for each pair $\mathcal{M}$ and σ, $\mathcal{S}$ is the set of models produced by applying O to $\mathcal{M}$ as follows: If $(\phi)_\sigma$ is false in $\mathcal{M}$, then $\mathcal{S}$ contains one model, $\mathcal{M}$. Otherwise, $\mathcal{S}$ contains exactly every model $\mathcal{M}^*$ such that

(1) $\mathcal{M}^*$ agrees with $\mathcal{M}$ on the truth values of all Skolem-constant-free atoms except possibly those in $(\omega)_\sigma$; and

(2) $(\omega)_\sigma$ is true in $\mathcal{M}^*$, and its truth valuation does not violate the unique name axioms of $\mathcal{L}$.

Example. If the agent observes IF T THEN Red(A) $\vee$ Green(A), then three models are created from each model $\mathcal{M}$ of $\mathcal{T}$: one where Red(A) $\wedge$ Green(A) is true, one where $\neg$Red(A) $\wedge$ Green(A) is true, and one where Red(A) $\wedge\neg$Green(A) is true—regardless of whether A was red or green in $\mathcal{M}$ originally.

The remarks at the beginning of this section on correctness of observation processing may be summed up in the following definition:

Definition. The incorporation of an observation IF ϕ THEN ω into an extensional belief theory $\mathcal{T}$, producing a new theory $\mathcal{T}'$, is *correct and complete* iff $\mathcal{T}'$ is an extensional belief theory and the alternative worlds of $\mathcal{T}'$ are exactly those alternative worlds represented by the union of the models in the $\mathcal{S}$ sets.

Please note that the observation IF ϕ THEN ω does not set up a new axiom regarding ϕ and ω; rather, the new information is subject to revision at any time, as is all extensional data.

3.3. An Algorithm for Incorporating Observations into the KB

The semantics presented in the previous section describe the effect of an observation on the *models* of a theory; the semantics give no hints whatsoever on how to translate that effect into changes in the extensional belief theory. An algorithm for incorporating observations into the KB cannot proceed by generating models from the theory and updating them directly, because it may require exponential time to generate even one model (since satisfiability testing is involved) and there may be an exponential number of non-isomorphic models. Any algorithm for observations must find a more efficient way of implementing these semantics.

The *Observation Algorithm* proposed in this section for incorporating observations into an extensional belief theory $\mathcal{T}$ may be summarized as follows: *For each atom f appearing in $\mathcal{T}$ that unifies with an atom of ω, replace all occurrences of f in $\mathcal{T}$ by a history atom.* [†] *Then add a new formula to $\mathcal{T}$ that defines the correct*

[†] Other possible semantics are considered in [Winslett 86b].

[††] Since Skolem constants do not appear directly in models, the purpose of σ is to associate the Skolem constants in O with specific constants in $\mathcal{M}$, so that the agent can directly refer to objects such as "that block that I earlier observed was red, though I wasn't able to tell exactly which block it was."

[†] These history atoms are not visible externally, i.e., they may not appear in any information requested from or provided by the KB; they are for internal KB use only.

truth valuation of f when ϕ is false, and another formula to give the correct valuation of f when ϕ is true.

Before a more formal presentation of the Observation Algorithm, let us motivate its workings in a series of examples that will illustrate the problems and principles underlying the algorithm. Let the body of $\mathcal{T}$ be $\neg$ OnTopOf(A, B), and the new observation be IF T THEN OnTopOf(A, B).

One's natural instinct is to add $\phi \rightarrow \omega$ to $\mathcal{T}$, because the observation says that ω is to be true in all alternative worlds where ϕ is true now. Unfortunately, ω probably contradicts the rest of $\mathcal{T}$. For example, adding T$\rightarrow$ OnTopOf(A, B) to $\mathcal{T}$ makes $\mathcal{T}$ inconsistent. Evidently ω may contradict parts of $\mathcal{T}$, and those parts must be removed from $\mathcal{T}$; in this case it would suffice to simply remove the formula $\neg$ OnTopOf(A, B).

But suppose that the body of $\mathcal{T}$ contains more complicated formulas, such as Red(A)$\leftrightarrow \neg$ OnTopOf(A, B). One cannot simply excise $\neg$ OnTopOf(A, B) or replace it by a truth value without changing the models for the rest of the atoms of $\mathcal{T}$; but by the semantics for observations, no truth valuation for extensional belief atoms except that of OnTopOf(A,B) can be affected by the requested observation. We conclude that contradictory wffs cannot simply be excised. They may be ferreted out and removed by a process such as that used in [Weber 86]; however, in the worst case such a process will multiply the space required to store the theory by a factor that is exponential in the number of atoms in the observation!

The solution to this problem is to replace all occurrences of OnTopOf(A, B) in $\mathcal{T}$ by another atom. However, the atom used must not be part of the alternative world of the agent, as otherwise the replacement might change that atom's truth valuation. This is where the special history predicates of $\mathcal{L}$ come into play; we can replace each atom of ω by a history atom throughout $\mathcal{T}$, and make only minimal changes in the truth valuations in the alternative worlds of $\mathcal{T}$. In the current case, OnTopOf(A, B) is replaced by $H_{OnTopOf}$(A, B, O), where O is simply a unique ID for the current observation. † For convenience, we will write $H_{OnTopOf}$(A, B, O) as H(OnTopOf(A, B), O), to avoid the subscript. The substitution that replaces every atom f of ω by its history atom $H(f, O)$ is called the *history substitution* and is written σ_H.

Let's now look at a slightly more complicated observation. Suppose that the agent observes O: IF Red(B) THEN OnTopOf(A, B), when $\mathcal{T}$ contains $\neg$ OnTopOf(A, B). As just explained, the first step is to replace this body by $(\neg \text{OnTopOf(A, B)})_{\sigma_H}$, i.e., $\neg H$(OnTopOf(A,B),O). Within a model $\mathcal{M}$ of $\mathcal{T}$, this step interchanges the truth valuations of every atom f in ω and its history atom $H(f, O)$; if ϕ was true in $\mathcal{M}$ initially, then $(\phi)_{\sigma_H}$ is now true in $\mathcal{M}$.

It is now possible to act on the original algorithmic intuition and add $(\phi)_{\sigma_H} \rightarrow \omega$ to the body of $\mathcal{T}$, establishing correct truth valuations for the ground atomic formulas of ω in models where ϕ was true initially. In the blocks world example, the body of $\mathcal{T}$ now contains the two formulas $\neg H$(OnTopOf(A,B),O) and Red(B)$\rightarrow$ OnTopOf(A,B).

Unfortunately, the fact that if B is *not* red then A is *not* on top of B has been lost! The solution is to also add formulas governing truth valuations for atoms in ω when ϕ is false: Add $\neg(\phi)_{\sigma_H} \rightarrow (f \leftrightarrow H(f, O))$ to $\mathcal{T}$ for each atom f in ω. Then $\mathcal{T}$ contains $\neg H$(OnTopOf(A,B),O), Red(B)$\rightarrow$ OnTopOf(A,B), and $\neg$ Red(B)$\rightarrow$ (OnTopOf(A,B)$\leftrightarrow H$(OnTopOf(A,B),O)). $\mathcal{T}$ now has the desired alternative worlds.

The informal algorithm proposed so far does not work when Skolem constants are present in either the theory or the observation. The basic difficulty is that one must update every atom in the theory that unifies with something in ω, since truth valuations for that atom might possibly be changed by the new observation. For example, suppose the body of $\mathcal{T}$ contains the formula Red(ϵ), and the agent receives the new information IF T THEN $\neg$ Red(B). In other words, the agent knew that some object was red, and has observed that block B is now not red, quite possibly because it has just been painted green. † A moment's thought shows that quite possibly no object is now red (e.g., if the agent has been painting them one by one), and so the formula Red(ϵ), which unifies with Red(B), must be changed in some way; $(\epsilon \neq \text{B}) \rightarrow \text{Red}(\epsilon)$ is the obvious replacement. In the general case, it is necessary to replace all atoms in $\mathcal{T}$ that unify with atoms of ω by history atoms as part of the history substitution step.

Let's examine one final example. Suppose the agent's theory initially contains the wff Red(A) and the new observation takes the form IF T THEN Red(ϵ). The suggested algorithm produces a new theory containing the three formulas H((Red, A), O), T$\rightarrow$ Red(ϵ),

† If the argument O were not present, then a similar substitution in a later observation involving OnTopOf(A, B) would make big changes in the alternative worlds of $\mathcal{T}$ at that time.

† In other words, the observation leaves open the possibility that the underlying state of the world has changed. To say that block B is not the red object ϵ, while retaining the belief that some object is red, the appropriate observation is IF Red(B) THEN F; this new observation says that the state of the world has not changed, but that we now have more information about its state. Although both observations talk about the redness of block B, their semantics are quite different.

and $F \rightarrow (\text{Red}(\epsilon) \leftrightarrow H((\text{Red}, \epsilon), O))$. Unfortunately, this theory has models where Red(A) is false! The problem is that the algorithm does not yet properly take care of the alternative worlds where ϵ is not bound to A; for in those worlds, Red(A) must still be true, regardless of what the new information in the observation may be. The solution is to add $(\epsilon \neq \text{A}) \rightarrow (\text{Red(A)} \leftrightarrow H(\text{Red(A)}, O))$ to $\mathcal{T}$, and in fact this new theory has the desired alternative worlds.

The lessons of the preceding examples may be summarized as an algorithm for executing an observation O given by IF ϕ THEN ω against an extensional belief theory $\mathcal{T}$.

The Observation Algorithm

Step 1. Make history. For each atom f either appearing in ω or appearing in $\mathcal{T}$ and unifying with an atom of ω, replace all occurrences of f by its history atom $H(f, O)$ in the body of $\mathcal{T}$, i.e., replace the body $\mathcal{N}$ of $\mathcal{T}$ by $(\mathcal{N})_{\sigma_H}$. Call the resulting theory $\mathcal{T}'$.

Step 2. Define the observation. Add the wff $(\phi)_{\sigma_H} \rightarrow \omega$ to $\mathcal{T}'$.

Step 3. Restrict the observation. First a bit of terminology: if atoms f and g unify with one another under substitutions σ_1 and σ_2, then σ_1 is *more general* than σ_2 if it contains fewer individual substitutions for Skolem constants. For each atom f in $\mathcal{T}'$ such that f is not an equality atom and f unifies with some atom of ω, let Σ be the set of all most general substitutions σ under which f unifies with atoms of w. Add the wff

$$(f \leftrightarrow H(f, O)) \vee ((\phi)_{\sigma_H} \wedge \bigvee_{\sigma \in \Sigma} \sigma)$$

to $\mathcal{T}'$, writing the substitution $\sigma = {}^{\epsilon_1 \ldots \epsilon_n}_{c_1 \ldots c_n}$ as $(\epsilon_1 = c_1) \wedge \cdots \wedge (\epsilon_n = c_n)$. Intuitively, for f an atom that might possibly have its truth valuation changed by observation O, this formula says that the truth valuation of f can change only in a model where ϕ was true originally, and further that in any model so created, f must be unified with an atom of ω. ◇

The models of $\mathcal{T}'$ produced by the Observation Algorithm represent exactly the alternative worlds that O is defined to produce from $\mathcal{T}$:

Theorem 1. Given an extensional belief theory $\mathcal{T}$ and an observation O, the Observation Algorithm correctly and completely performs O. In particular,

(1) the Observation Algorithm produces a legal extensional belief theory $\mathcal{T}'$;

(2) The alternative worlds of $\mathcal{T}'$ are the same as the alternative worlds produced by directly updating the models of $\mathcal{T}$. ◇

The interested reader is referred to [Winslett 86b] for the proof of Theorem 1.

3.4. Cost of the Observation Algorithm

Let k be the number of instances of atoms in the observation O; and let R be the maximum number of distinct atoms of $\mathcal{T}$ over the same extensional predicate. When $\mathcal{T}$ and O contain no Skolem constants, the Observation Algorithm will process O in time $O(k \log(R))$ (the same asymptotic cost as for ordinary database updates) and increase the size of $\mathcal{T}$ by $O(k)$ worst case. This is not to say that an $O(k \log(R))$ implementation of observations is the best choice; rather, it is advisable to devote extra time to heuristics for minimizing the length of the formulas to be added to $\mathcal{T}$. Nonetheless, a worst-case time estimate for the algorithm is informative, as it tells us how much time must be devoted to the algorithm proper. The data structures required for this running time are described elsewhere [Winslett 86].

When Skolem constants appear in $\mathcal{T}$ or in O, the controlling factor in costs is the number of atoms of $\mathcal{T}$ that unify with atoms of O. If n atoms of $\mathcal{T}$ each unify with one atom of O, then $\mathcal{T}$ will grow by $O(n+k)$. In the worst case, *every* atom of $\mathcal{T}$ may unify with *every* atom of O, in which case after a series of m observations, the number of occurrences of atoms in $\mathcal{T}$ may *multiply* by $O(mk)$. To prevent excessive growth in $\mathcal{T}$, we have devised a scheme of delayed evaluation and simplification of expensive observations, by bounding the permissible number of unifications for the atoms of an incoming observation.

4. Other Work

For a discussion of the interaction of extensional and intensional information in belief revision (e.g., how to enforce a class of axioms in the face of incoming information) and other possible semantics for belief revision, the interested reader is referred to [Winslett 86b].

5. Summary and Conclusion

In this paper we represent the set of extensional beliefs of an agent as a logical theory, and view the models of the theory as representing possible states of the world that are consistent with the agent's extensional beliefs. The extensional portion of an agent's knowledge base is envisioned as a set of formulas that are not generated via inference from other data but rather by direct observation; the remainder of the KB contains data-independent logical inference rules, data-dependent axioms, and intensional beliefs that are derived from extensional and other intensional beliefs using the agent's axioms and rules of inference. The agent's extensional beliefs may contain incomplete information in the form of disjunctions or *null values* (attribute values that are known to lie in a certain domain but whose value is currently unknown). We formalize the extensional belief set as an *extensional belief theory*; formulas in the body of an extensional belief theory may be any sentences without universal quantification.

From time to time an agent will make *observations*, i.e., produce new extensional beliefs. This paper sets forth a language and semantics for observations, and an algorithm for incorporating observations into the agent's KB. This Observation Algorithm is proven correct in the sense that the alternative worlds produced under this algorithm are the same as those produced by processing the observation in each alternative world individually. For beliefs and observations without Skolem constants, the Observation Algorithm has the same asymptotic cost as for an ordinary complete-information database update, but may increase the size of the KB. For observations involving Skolem constants, the increase in size will be severe if many atomic formulas in the KB unify with those in the observation; if desired, a lazy evaluation technique may be used to control expansion. A simulation program has been constructed for a closed-world version of the Observation Algorithm.

In sum, we have shown that the incorporation of new, possibly contradictory extensional information into a set of extensional beliefs is difficult in itself when either the old or new beliefs involve incomplete information, even when considered in isolation from the problems associated with justification of prior conclusions and inferences from extensional data and the identification of outdated and incorrect axioms. We have produced a polynomial-time algorithm for incorporating extensional observations; however, it is not in general possible to process observations efficiently in extensional belief theories if the observations reference intensional beliefs.

6. Acknowledgments

The results cited in this paper owe much to the encouragement of Christos Papadimitriou and the emphasis on rigor of Moshe Vardi. This paper would not have been possible without the AI perspective and suggestions of David C. Wilkins.

7. References

[Abiteboul 85] S. Abiteboul, G. Grahne, "Update Semantics for Incomplete Databases," *Proc. of the VLDB Conf.*, Stockholm, August 1985.

[DeKleer 85] J. DeKleer, B. C. Williams, "Diagnosing Multiple, Possibly Intermittent, Faults, Or, Troubleshooting Without Backtracking," Xerox PARC Tech. Report, 1985.

[Doyle 79] J. Doyle, "A Truth Maintenance System," *Artificial Intelligence* **12**, July 1979.

[Fagin 84] R. Fagin, G. M. Kuper, J. D. Ullman, and M. Y. Vardi, "Updating Logical Databases," *Proc. of the 3rd ACM PODS*, April 1984.

[Fagin 83] R. Fagin, J. D. Ullman, and M. Y. Vardi, "On the Semantics of Updates in Databases," *Proc. of the 2nd ACM PODS*, April 1983.

[Imielinski 84] T. Imielinski and W. Lipski, "Incomplete Information in Relational Databases," *Journal of the ACM* **31**:4, October 1984.

[Levesque 84] H. Levesque, "Foundations of a Functional Approach to Knowledge Representation," *Artificial Intelligence* **23**, July 1984.

[Reiter 84] R. Reiter, "Towards a Logical Reconstruction of Relational Database Theory," in M. Brodie, J. Myopoulos, and J. Schmidt (eds.) *On Conceptual Modelling*, Springer-Verlag, 1984.

[Reiter 85] R. Reiter, "A Theory of Diagnosis from First Principles," University of Toronto Computer Science Dept. Tech. Report No. 86/187, Nov. 1985.

[Weber 86] A. Weber, "Updating Propositional Formulas," *Proc. of the Expert Database Systems Conference*, Charleston SC, April 1986.

[Winslett 86] M. Winslett-Wilkins, "A Model-Theoretic Approach to Updating Logical Databases," Stanford Computer Science Dept. Tech. Report, Jan. 1986. A preliminary version appeared in *Proc. of the 5th ACM PODS*, March 1986.

[Winslett 86b] M. Winslett, "Is Belief Revision Harder Than You Thought?," Stanford University Computer Science Dept. Tech. Report, June 1986.