Towards a theory of resource-bounded belief revision

Renata Wassermann
ILLC - Plantage Muidergracht, 24
1018-TV Amsterdam, The Netherlands
renata@fwi.uva.nl

1 Introduction

The problem of belief revision, that is, how the beliefs of an agent change in the presence of new information, has been recently addressed by various authors. In most approaches, the agents are considered to be ideal, perfect beings in the sense that they never forget anything, never hold inconsistent beliefs and their beliefs are closed under logical implication. Harman [Har86] presents an analysis of belief revision for non-ideal agents that, even though very informal, will be used as a guideline for a more formal proposal. In this paper we proceed towards a theory of belief revision for resource-bounded agents, where we take in consideration limited memory and capacity of inference.

There are several different ways in which a process of belief revision may take place. It may be necessary to revise one’s beliefs due to new information acquired by the agent from the outside world. But the agent may also reason on beliefs that she already had and come to new conclusions or even find contradictions. Most of the literature following the AGM theory [Gär88, Gär92] only deals with the problem of revision when new information comes from outside. Moreover, only with the case where the agent accepts the new information. The AGM theory does not deal with the case where the information is not accepted. This can be justified by the fact that the considered agent is taken to be ideal, that is, she does not have inconsistent beliefs and her beliefs are closed under logical implication, so that if the agent does not accept the information, nothing changes.

In the case where the agent is not ideal, even if she refuses to accept some new information, during the process of deciding whether or not to accept new information the agent may find new conclusions and inconsistencies.

In Galliers’ proposal for a theory of autonomous belief revision [Gal92], the agents can decide whether to accept the information or not. Galliers defines the result of a revision operation as a set of belief sets, including the original (non-revised) belief set and ordered according to some criteria. The maximal sets are then preferred.

The agent can get new information either from the “outside world”, when she sees or hears something, or from “inside”, when she infers something from her previous beliefs. Since we want to allow for an agent’s beliefs to be inconsistent, even when the agent infers something from her old beliefs, she may refuse to accept the new conclusion. Our claim is that the process of dealing with new information coming from “outside” is not very different from dealing with new beliefs derived from the existing ones, in the sense that they have to be checked before being merged into the agent’s beliefs.

We will first propose a structure to represent an agent’s belief states that distinguishes different status of beliefs according to whether they are explicitly represented or not, whether they are currently active and whether they are fully accepted or provisional.

Then we will define some operations that can be performed on this structure in order to incorporate new beliefs, get rid of beliefs that, because of lack of
evidence or simply lack of space, cannot be held anymore, and change the status
of beliefs. We will show that this set of operations is functionally complete in
relation to all possible movements in the structure.

The well-known AGM theory for belief revision [Gär88] and Nebel's base re-
vision theory [Neb89] can be seen as special cases of our theory and we will show
how they can be embedded in our framework.

For reasons of space, the proofs of lemmas and propositions are omitted. The
interested reader can find them in [Was].

2 Belief states

Several proposals to represent knowledge and formalize reasoning make a distinction
between explicit and implicit beliefs. The way this distinction is made varies
a lot from author to author. The paper by Fagin and Halpern [FH88] for example
presents three different logics for representing beliefs, including a Logic of General
Awareness where the set of an agent's implicit beliefs is logically closed, while the
explicit beliefs are those elements of this set which the agent is aware of. Harman
[Har86] calls explicit those beliefs that in some way are explicitly represented.
His definition of implicit beliefs is rather vague, they can be those that can be
inferred from the explicit, where inference is different from logic implication, but
they may also not be inferable from the explicit beliefs.

Since we are dealing with resource-bound agents, it does not make sense
to state that an agent can infer anything that logically follows from her beliefs.
How much one can infer depends highly on the available resources, like time and
memory.

Following Harman, we will assume that there are beliefs that are explicitly rep-
resented, from which other beliefs can be inferred. Departing from him, though,
we will consider the set of the agent's implicit beliefs to be the set of beliefs
that can be inferred from the explicit beliefs, taking the resource-boundness into
account. We will not restrict ourselves to a particular notion of inference but
consider an inference function \( i \) that will depend on the agent being modeled.

We will call active beliefs the beliefs that are currently available for use. These
may be recently acquired beliefs, intermediate conclusions in an argument, beliefs
related to the current topic, etc. Every belief has first to become active in order
to be accepted, rejected or revised. Not all of one's beliefs are active at once, the
"amount" of beliefs that can be active is very restricted.

At this point it might be useful to introduce a small example to illustrate the
difference between explicit and active beliefs. Consider a Prolog program. What
we call explicit beliefs is the program itself, that is, the facts and rules that are
explicitly given. The inference function is the immediate consequence operator
and the set of implicit beliefs is the fix-point of this operator. The active beliefs
depend on the queries made.

For the program:

\[
\begin{align*}
p & : - q, r. \\
q. \\
r & : - s. \\
s. \\
\end{align*}
\]

we have:

- Explicit beliefs: \( \{ p : - q, r; q : - s; s \} \)
- Implicit beliefs: \( \{ p : - q, r; q : - s; s \} \cup \{ r; p \} \)
For the query \( r \), the set of active beliefs is first only the query, then \( \{ r ; r : s \} \), then \( \{ r ; r : s ; s \} \), that is, the set contains open queries as well as the clauses of the program that are used to solve them. Note that the three sets of explicit, implicit and active beliefs are different from each other.

Any new belief, either coming from the “outside” or from the “inside”, has to survive inquiry before being incorporated to the existent beliefs. The depth of the inquiry is determined by the agent, by her interest on the subject.

Harman points two different theories about belief dependencies and what could cause a belief to be given up or incorporated. The first, the so-called coherence theory, claims that an agent does not keep track of all the belief dependencies. One does not have to remember the origin of one’s beliefs, they are accepted as long as they are coherent with the rest of the agent’s beliefs.

On the other hand, according to the foundation theory, something can be believed as long as there is a valid justification associated to it. This means an agent should keep track of all the dependencies between her beliefs, which is an unrealistic requirement in the case of real resource-based agents.

There are several theories that predict that agents behave partly according to the foundation theory and partly according to the coherence theory, that is, theories that try to combine aspects of both approaches.

In [Har86], Harman distinguishes fully accepted beliefs from what he calls working hypotheses, the first being those working hypotheses that managed to survive inquiry. Following Harman, a working hypothesis has to survive the best attempts to refute it in order to be fully accepted by an agent. We will call these beliefs _provisional beliefs._

When a belief is fully accepted, one does not need to keep track of its justifications, since by accepting it, the agent is considering that her reasons to believe it are good enough. Unless something occurs that throws some doubts about it, the agent will not have a reason to continue inquiring. On the other hand, provisional beliefs must have their justifications associated, since in a way the agent is still trying to decide whether to fully accept it or not. The set of the provisional beliefs is a subset of the set of active beliefs.

We call the set of provisional beliefs \( \text{prov} \) and its complement in relation to the active beliefs \( \text{acc} \) (from “accepted”).

Beliefs that are active, either in \( \text{prov} \) or in \( \text{acc} \), can be forgotten or stored as explicit (but inactive) beliefs. Since both sets are very limited in size, there must be a mechanism that, in cases of overflow, selects which beliefs will be forgotten or stored. It may be interesting to have the active beliefs ordered by interest, so that things with very low interest will be forgotten first. This ordering could also incorporate recency, beliefs that were recently recalled are more interesting than those that were not used lately.

We call \( E \) the set of explicitly represented beliefs, and \( A \) the set of beliefs that are active. The set \( A \) is partitioned in the two subsets \( \text{prov} \) (provisional beliefs) and \( \text{acc} \) (accepted beliefs).

The set \( I \) of the implicit beliefs is a superset of \( E \), determined by a function \( i \) that gives us what can be inferred from a set in one step:

\[
I = E \cup i(E) \cup i(i(E)) \cup ... = i^*(E)
\]

The properties of the inference function \( i \) depend on the agent under consideration, but there are some properties that we usually would like the function to satisfy, like reflexivity, that is, \( X \subseteq i(X) \). We take the set of implicit beliefs to be the transitive closure under the inference function (and thus, we have
Intuitively it seems that implicit beliefs can also be active. In a sense, all of an agent's active beliefs have to be explicitly represented, even if only temporarily. But they do not need to be represented in the same way as the explicit beliefs are. Recall the example of a Prolog program. The explicit beliefs are the facts and rules that constitute the program. The implicit beliefs are those facts that can be inferred from the program, while the active beliefs are those facts and rules that have been derived or used at a certain point (acc) and the queries (prov).

Figure 1: Structure of an agent's beliefs

Our view of the structure of an agent's beliefs (figure 1) differs from the ones found in the literature, like [Kon86] that does not distinguish active from explicit beliefs, [FH88] that start from the set of implicit beliefs and eliminate things to get to the set of the explicit beliefs and [Har86] that considers that implicit beliefs can be derived from the explicit ones in a different way than by inference.

We also drop the requirement that belief sets have to be consistent, since we believe that agents can have inconsistent beliefs without believing everything (we can keep some beliefs that we even know to be inconsistent).

A provisional belief may already be believed (explicitly or implicitly) without the agent noticing it. We would like to further refine the set $A$ to mirror the difference between beliefs that the agent is “aware” of believing, in the sense that she believes she believes it and beliefs that she is not aware of.

As can be seen in figure 2, the set of provisional beliefs (prov) includes beliefs that were not previously believed as well as beliefs that are already (implicitly or explicitly) believed but of which the agent is not aware.

The set $E_A$ is a subset of $E \cap A$ and contains those beliefs that, besides being explicit and active, are also actively believed to be believed. Analogously, the set $I_A$, the set of active accepted beliefs, is a subset of $I \cap A$. The set $E_A$ is implicitly defined as $I_A \cap E$.

A belief state $\beta$ can be completely determined by the following parameters:

- The set $E$ of explicit fully accepted beliefs;
- The inference function $i$, that takes as argument a set of beliefs and gives as result the set of beliefs that can be inferred by the agent from the argument set in one step;
- The set $A$ of active beliefs, partitioned in the sets prov of provisional beliefs and acc of active accepted beliefs and
Figure 2: The active beliefs

- The set $I_A$ of active and implicit beliefs that the agent is aware of.

From now on, we will refer to a belief state as $\beta = (E, i, A, I_A)$.

3 Moving from one set to the other

After having defined how the beliefs are structured, the next step is to study how the sets can change, that is, how beliefs can get in or out from a set.

We will define an operation $(\cup^*)$ similar to Gärdenfors’ expansion [Gär88], but that takes the resource boundness into account. If $X$ is a set with a maximum size $m$ and $p$ is an element we want to add to $X$, we define:

$$X \cup^* \{p\} = X' \cup \{p\}, \text{ where } X' \subseteq X \text{ and } |X'| < m$$

The intended meaning of this operation is that it is a simple union as long as the set is not “full”. When the set is already at its maximum size, something has to be dismissed. If the set $X$ is ordered, we can have that the minimal elements of the set are the first to be dismissed.

In our structure, even though the set $E$ has a limited size, this size may be big enough not to be considered. The important restriction problem is the size of the set $A$. Elements dismissed from $A$ that do not belong to $E$ or $I$ simply disappear, are forgotten. The elements of $A \cap I$ that are dismissed from $A$ remain in $E$ or $I$.

There are several ways in which a belief state $\beta = (E, i, A, I_A)$ can be modified (the numbers refer to the regions in figure 3):

- Observation $(+_o)$ - The agent receives new information from “outside”, either via observation or communication. The new information goes first to the set $prov$ of provisional beliefs (regions 3, 5 or 7). The new information state $\beta'$ is given by:

$$\beta' = \beta +_o \{\varphi\} = (E, i, A', I_A'),$$

where $A' = A \cup^* \{\varphi\}$ and $I_A' = I_A \cap A'$. 

5
Figure 3: Numbered regions

- Inference ($+$) - The new information is inferred from the accepted active beliefs, that is, it comes from the set $i(I_A)$ and goes to region 6 ($I_A$).
  
  $\beta' = \beta + i \{ \varphi \} = (E, i, A', I_A')$,
  
  where $\varphi \in i(I_A)$,
  
  $A' = A \cup \{ \varphi \}$ and
  
  $I_A' = (I_A \cup \{ \varphi \}) \cap A'$.

- Retrieval ($+$) - An explicit belief that is inactive becomes active. It moves from region 2 to 4.
  
  $\beta' = \beta + r \{ \varphi \} = (E, i, A', I_A')$,
  
  where $\varphi \in E$,
  
  $A' = A \cup \{ \varphi \}$ and
  
  $I_A' = (I_A \cup \{ \varphi \}) \cap A'$.

- Acceptance (explicit) ($+$) - A provisional belief that was in region 7 is fully accepted and moves into $E_A$ (region 4), a belief that was in $\text{prov} \cap E$ (region 3) is discovered to be already believed and moves into $E_A$ or a belief that was in $I_A$ (region 6) is considered to be worth becoming explicit.
  
  $\beta' = \beta + a_e \{ \varphi \} = (E', i, A, I_A')$,
  
  where $E' = E \cup \{ \varphi \}$ and
  
  $I_A' = I_A \cup \{ \varphi \}$.

- Acceptance (implicit) ($+$) - A belief that was in $\text{prov}$ (region 5) is seen to be already implicitly believed and moves into $I_A$ (region 6).
  
  $\beta' = \beta + a_i \{ \varphi \} = (E, i, A, I_A \cup \{ \varphi \})$

- Non-acceptance (rejection) ($+$) - A belief that was in $\text{prov}$ (region 7) is not accepted because it did not survived inquiry and is dismissed.
  
  $\beta' = \beta + n \{ \varphi \} = (E, i, A - \{ \varphi \}, I_A)$

- Doubting ($+$) - A belief that was fully accepted is questioned, moving from region 4 to region 7.
  
  $\beta' = \beta + d \{ \varphi \} = (E', i, A, I_A')$. 
The operations defined above describe how beliefs are incorporated to the structure representing an agent's beliefs and how they move from one set to the other within the structure. They do not specify how beliefs get out of the structure, though. Elimination of beliefs are a side effect of the expansion operation $\cup^*$, being thus an "unconscious" operation. However, depending on the agent being modeled it may be useful to have an operation to perform "conscious" forgetting, for example for robots.

We will call active belief revision those revisions that the agent performs consciously, in the sense that the changes are all related to the set of active beliefs. A belief cannot move from region 1 to region 2 in figure 3 without ever being active.

The seven operations defined above can be combined to model more complex operations. For example: when an agent gets new information via observation, the belief will first come to the set prov through the operation $+_{\alpha}$ and then the agent may fully accept it or discover that she already believed it ($+_{\alpha}$).

**Proposition 3.1** The set of operations defined above is complete in relation to all possible changes in a belief state.

The seven operations are sufficient (but not necessary) to describe any active belief revision given our structure. Actually, five of these seven operations are sufficient, since an inference can be simulated by an observation followed by implicit acceptance and in an analogous way a retrieval can be simulated by an observation followed by explicit acceptance.

### 4 Embedding the AGM theory

In this section, we compare our proposal with the rationality postulates for belief changes given in [Gär88] and show that our theory is compatible with the AGM paradigm. We show how the operations of expansion and contraction in the AGM sense can be seen as a special case of applications of the operations defined in section 3. We also show how to embed the revision of belief bases [Neb89] in our framework.

As mentioned before, Gärdenfors defines three kinds of belief changes that can occur: expansion - the addition of a new belief and its consequences, contraction - the retraction of some belief without adding any new belief, and revision - the addition of a belief which may cause the retraction of some old beliefs in order to maintain consistency. We will concentrate on expansion and contraction, since revision can be defined in terms of the other two operations.

In order to define the postulates, he assumes that belief states are modeled by belief sets, that is, sets of sentences closed under logical implication. In order words, if we define the set of logical consequences of a set $X$ by:

$$Cn(X) = \{\varphi | X \vdash \varphi\}$$

a belief set is a set $K$ such that $Cn(K) = K$. There is no distinction between implicit, explicit or active beliefs.

To see how this can be embedded in our structure, we first observe that a belief set in the AGM theory corresponds to a belief state where the sets $I, E, A, E_A$ and $I_A$ happen to be the same and have no size limit, and $i^*$ is $Cn$. Thus, we define a map function $f$ from belief sets into belief states:

$$f(K) = (K, Cn, K, K)$$
**Definition 4.1** Let $\beta = \langle K, Cn, K, K \rangle$ be a belief state and $\varphi_1, \varphi_2, \ldots$ be an enumeration of the formulas in $Cn(K)$. The logical closure of $\beta$ is given by:

$$Cl(\beta) = \beta +_o \{ \varphi_1 \} +_{a_e} \{ \varphi_1 \} +_o \{ \varphi_2 \} +_{a_e} \{ \varphi_2 \} +_o \ldots$$

**Lemma 4.1** Let $\beta = \langle K, Cn, K, K \rangle$ be a belief state. Then $Cl(\beta) = \langle Cn(K), Cn, K, Cn(K) \rangle$.

Given a belief set $K$, [Gär88] defines the expansion of $K$ with $\varphi$ by:

$$K + \varphi = Cn(K \cup \{ \varphi \}) = \{ \psi | K \cup \{ \varphi \} \vdash \psi \}$$

**Definition 4.2** The expansion of a belief state $\beta = \langle K, Cn, K, K \rangle$ by $\varphi$ is given by:

$$\beta +_{AGM} \{ \varphi \} = Cl(\beta +_o \{ \varphi \} +_{a_e} \{ \varphi \})$$

We have then that by mapping the resulting belief set after an AGM expansion into a belief state, we get the same result as predicted by our theory:

**Lemma 4.2** $f(K) +_{AGM} \{ \varphi \} = f(K + \varphi)$

For the contraction of a belief set $K$ in relation to a sentence $\varphi (K - \varphi)$, six basic postulates are given [Gär88]:

1. $K - \varphi$ is a belief set
2. $K - \varphi \subseteq K$
3. if $\varphi \not\in K$, then $K - \varphi = K$
4. if $\not\vdash \varphi$, then $\varphi \not\in K - \varphi$
5. if $\varphi \in K$, then $K \subseteq (K - \varphi) + \varphi$
6. if $\vdash \varphi \leftrightarrow \psi$, then $K - \varphi = K - \psi$

**Definition 4.3** Let $- \cdot$ be an operation satisfying postulates 1-6 and let $\varphi_1, \varphi_2, \ldots$ be an enumeration of the formulas in $K \setminus (K - \varphi)$. The contraction of a belief state $\beta = \langle K, Cn, K, K \rangle$ by $\varphi$ is given by:

$$\beta -_{AGM} \{ \varphi \} = \beta +_d \{ \varphi_1 \} +_{a_e} \{ \varphi_1 \} +_d \{ \varphi_2 \} +_{a_e} \{ \varphi_2 \} +_d \ldots$$

**Lemma 4.3** $f(K) -_{AGM} \{ \varphi \} = f(K - \varphi)$

The revision of a belief set can be defined by combining contraction and expansion, according to the Levi identity [Gär88]. Revising with a belief $\varphi$ is equivalent to contracting by the negation of $\varphi$ and then expanding with $\varphi$:

$$K * \varphi = (K - \neg \varphi) + \varphi$$

This allows use to conclude the following:

**Proposition 4.1** The standard AGM theory of belief revision can be embedded in the framework of belief states equipped with the operations defined in section 3.

In order to obtain a contraction operation that satisfies all the AGM postulates, the operations of doubting and non-acceptance have to be repeated until enough sentences have been removed so that $\varphi$ is not a consequence of the first argument of the resulting belief state. Such a definition, as the one in [Gär88], is only valid for ideal agents.

We can define another contraction operation that would simply remove the formula from the sets of the belief state being contracted and take its logical
closure. The first argument of the resulting belief state, \( K' = Cn(K \setminus \{\varphi\}) \) clearly satisfies postulates 1, 2, 3, 5 and 6, given the expansion operation defined above. Since there may be other beliefs in \( K \) from which the agent can infer \( \varphi \), even if \( \varphi \) is not a valid sentence, after the contraction we may have \( \varphi \in K' \) contradicting postulate number 4. Harman [Har86] suggests that a “Get Back Principle” should be followed, that is, “one should not give up a belief one can easily (and rationally) get right back”. According to this principle, one should not contract one’s belief state by a belief \( \varphi \) that can be reinferrred from the other beliefs.

An operation of contraction for resource-bounded agents should ideally lie somewhere in between AGM contraction for ideal agents and the simple subtraction of the belief from the belief state. The AGM contraction presupposes that the agent can keep track of all the dependencies between her beliefs and then, when contracting in relation to one belief, remove also those beliefs which implied it. As discussed in section 2, our agents can keep track of some of the dependencies, namely those related to beliefs in the set \( \text{prov} \), but not all. This will be further explored when we deal with the internal structure of the provisional beliefs.

Also the appropriateness of AGM postulate number 6 for contraction can be discussed, since its validity depends on the agents being modeled and their capacity of detecting logical equivalence.

In the case of resource-bounded agents, there is no such clear distinction between an expansion and a revision, since the agent does not always detect inconsistencies. This means that when the agent gets the information that \( \varphi \) is the case, she may have other beliefs that are inconsistent with \( \varphi \), but keep them together with the new information without noticing it and without starting to believe everything, as would be predicted by the AGM theory.

The use of logically closed sets to represent beliefs has received many criticisms, specially from authors concerned with the computability of the theory of belief change. Besides the fact that belief sets are too large to be represented, they have no distinction between central beliefs and those who were inferred from them. Nebel [Neb89] proposed that instead of always considering a belief set as a whole, one should consider a belief base, a finite set containing the central beliefs, and take its logical closure when necessary.

In our framework, the set \( \text{E} \) of explicit beliefs can play the role of belief base. Nebel distinguishes the explicit and implicit beliefs, but not the active from non-active. Thus, we can define a mapping function from belief bases into belief states as:

\[
g(B) = \langle B, Cn, Cn(B), Cn(B) \rangle
\]

The operations of expansion and contraction of belief bases can be embedded in our framework in a way very similar to the embedding of the AGM theory.

**Proposition 4.2** The theory of belief base revision can be embedded in the framework of belief states.

5 Conclusions and Future Work

We have defined a structure for belief states and a set of operations that describe how belief states can change.

We have shown that these operations are sufficient to describe any change that can occur in the structure and how the standard AGM theory of belief revision...
and the theory of base revision [Neb89] can be seen as particular cases of our theory.

Our theory extends the AGM theory in the sense that it allows us to concentrate on particular subsets of an agent’s beliefs and classifies them according to their status - whether they are explicitly represented, currently active and fully accepted.

The theory also allows us to think of more general forms of belief changes by defining simple operations that work as building blocks to form more complex operations.

As mentioned in section 4, in order to implement a contraction operation for a belief state, we need more structure in the set prov, we have to represent the dependencies between the provisional beliefs. We are currently working on the foundational structure of the provisional beliefs.

Acknowledgments: I would like to thank Frans Voorbraak, Henk Zeevat and Remko Scha for the discussions that led to this paper. This research is sponsored by CAPES - Brazil.

References


