RENATA WASSERMANN

ON STRUCTURED BELIEF BASES

RENATA WASSERMANN

ON STRUCTURED BELIEF BASES

ABSTRACT: Most existing approaches to belief revision describe the behaviour of a highly idealized rational agent. In operations of belief change for more realistic agents, usually only a small part of an agent's beliefs is accessed at one time. This should be taken into account if we are looking for cognitively more appropriate operations. Furthermore, it makes implementation more feasible.

In this paper we show how extra structure of belief bases can be used for implementing local change as defined in [Hansson and Wassermann, 1999], where only the relevant part of an agent's beliefs is considered. Some ideas for adding structure to belief bases are presented. Structured belief bases can be seen as graphs, where each node is a belief and two nodes are adjacent if and only if they are related. We also define some notions of relatedness.

1 INTRODUCTION

Intelligent agents must have some way of revising their beliefs. Even though the AGM paradigm [Alchourrón *et al.*, 1985] provides a very elegant and powerful framework for reasoning about how a rational agent should change his beliefs when confronted with new information, it tells us very little about how that agent could really perform such belief changes. Moreover, the rational agent described is a highly idealized one, a perfect reasoner with infinite memory, logical ability, no inconsistent beliefs and no time constraints.

In the AGM paradigm, the belief state of an agent is represented by a *belief set*, a set of logical formulas that is closed under logical consequence. The use of belief sets has received several criticisms, the two most important being that they are usually to big to be (computationally) dealt with, and that they are not expressive enough, not allowing for example for distinctions between beliefs with an independent standing and merely derived ones.

One alternative for the use of belief sets if to represent the belief state of an agent by means of a set that is not closed under logical consequence, called *belief bases*. This alternative has been extensively studied and AGM-like operations have been defined for belief bases. However the term "belief base" has been used by two different communities with different meanings. Most authors interested in implementing AGM belief revision [Nebel, 1990; Williams, 1994; Dixon, 1994] use belief bases to represent the belief set of an agent. A belief base in these approaches is a finite axiomatization of the theory (belief set) representing the agent's commitments. The fact that a formula of the belief set is an element of the belief base does not distiguish it from other formulas that are not part of the belief base but follow logically from it. There are just two possibilities for a formula: either it follows from the belief base, and so the agent is commited to it, or it does not. Authors following this approach usually try to avoid redundancies in the belief base. In databases, for example, this is a desirable property, since it makes alterations of the data easier. On the other hand, adding new information becomes much more complex, since one has to check for redundancies.

The approach that we follow in this paper is in line with another use of the term "belief base", defended for example by [Fuhrmann, 1991; Rott, 1996; Hansson, 1998]. According to this line, a belief base is an arbitrary set of formulas representing those beliefs of an agent that have independent standing. Some authors [Fagin and Halpern, 1988] call these the *explicit beliefs*. Beliefs that follow from the ones in the belief base but are not part of it have a different status, being merely derived beliefs. The set of formulas to which the agent is commited, also called *implicit beliefs*, is formed by taking the logical closure of the agent's belief base. A model of the belief state of a resource-bounded agent is presented in [Wassermann, 1999], which includes explicit, implicit and active beliefs. The set of active beliefs is the set of formulas that is relevant for use at a certain moment. In [Hansson and Wassermann, 1999] we present one way of retrieving the active beliefs from the belief base. This method and its shortcomings are detailed in section 2.

While in the first notion of belief base the stress is on keeping a non-redundant base, and thus requiring a non-trivial operation of expansion, in the second notion what matters is the justification of an agent for believing the formulas of the base. In this case, adding new information to the belief base via expansion is a trivial operation. The complexity arises when one has to give up beliefs (contraction) or use the beliefs for reasoning, since they may well be inconsistent.[1]

When trying to model more realistic agents, one should take into account some limitations of the agents. We may hold inconsistent beliefs without even realizing it, and this does not prevent us from being able to reason about other topics. Realistic agents should not try to keep their whole set of beliefs consistent, but only solve inconsistencies when they become relevant for a line of reasoning.

Even if real agents (natural or artificial) had the ability to calculate all the consequences of their beliefs, they would not have any interest in doing so. Some consequences are useful, but an agent should not waste resouces calculating that $p \vee q$, $p \wedge p$, $p \wedge p \vee q$ and $p \wedge p \wedge p \vee q$ all follow from his

[1] In [Rott, 1996], two ways of using belief bases are distinguished. Either the belief base is a finite axiomatization of a belief set, which can be obtained by applying Cn on the base (the horizontal perspective), or the belief base is an arbitrary set of formulas, possibly inconsistent, and the commitments of the agent are obtained by applying a more sophisticated inference operator on the base (the vertical perspective). We follow in this paper the vertical perspective.

belief in p. In approaches to belief revision that use belief bases [Fuhrmann, 1991; Hansson, 1994; Nebel, 1990] instead of logically closed belief sets, the operations defined usually depend on finding maximal subsets of the belief base not implying some formula α, or minimal subsets of the base implying α.[2]Exploring all subsets of a reasonably sized belief base is a very expensive operation.

One way to attack the problem is to try to reduce the size of the set to be explored. Intuitively, not all of an agent's beliefs are relevant for deciding what to do with new information. There should be a way of isolating the subset of a belief base that contains the relevant beliefs for an operation of belief change. In [Hansson and Wassermann, 1999], this approach is explored. A notion of *compartment* is presented for retrieving the relevant part of a belief base and then local operations of belief change are defined that act only on the relevant part of the beliefs. The problem is that the way the compartments are defined uses the notion of minimal subsets of the base implying some formula, that is, finding a compartment is as expensive as performing a traditional operation of belief change. However, the representation results obtained in [Hansson and Wassermann, 1999] for the local versions of the belief change operations are very general and do not depend on the particular way the compartment is defined, but only on certain properties of the resulting compartments.

In this paper we present more efficient ways of retrieving the relevant part of a belief base. The retrieved set can then be "plugged" into the constructions of local operations. Even though the local operations also rely on finding minimal or maximal subsets implying or not implying certain formulas, the set to be explored is reduced to a manageable size.

The paper proceeds as follows: in the next section the basic definitions and results of [Hansson and Wassermann, 1999] are summarized. In section 3, a computer system is presented which has a very efficient mechanism for retrieving relevant beliefs from a belief base with extra structure. The ideas of the system inspired some considerations on more general ways of structuring the base and on formalizing the notion of retrieving the most relevant beliefs for a certain input. These considerations are presented in section 4. In section 5, some examples of relatedness relations are presented that can be used for structuring the belief base. Computational aspects are examined in section 6. Some related work is discussed in section 7, and conclusions are presented in section 8.

In the rest of the paper we will be working within a propositional framework, with a language $\mathcal{L}$ containing the basic boolean connectives. We will

[2]Alternative approaches based on epistemic entrenchments (E-bases [Rott, 1991]; "ensconcements" [Williams, 1994]) rely on a well behaved pre-order of the beliefs in the base to construct operations of belief change. Given the pre-order, the operations become less complex, but constructing and maintaining this pre-order can be computationally very costly.

assume that a belief base is a finite set of formulas of $\mathcal{L}$.

It is possible to ignore differences in the syntactical form of formulas, although we do not do it here, by assuming (i) a logical approach: when we say "p occurs in φ" we actually mean "for all ψ such that $\vdash \varphi \leftrightarrow \psi$, p occurs in ψ"; or (ii) a computational approach: the formulas are all represented in some normal form, and equivalent formulas have equal representation.

We will use the terms " belief base" and "database" indiscriminately to refer to a set of formulas of $\mathcal{L}$.

2 LOCAL CHANGE

In [Hansson and Wassermann, 1999], we presented a framework for performing local reasoning. The key idea is that not all of an agent's beliefs are relevant for an operation of belief change. As an example, suppose that I wake up and see that the sun is shining in Amsterdam. This contradicts my previous beliefs about the weather, like the one that it always rains in Holland, and leads to revision. However, the revision process does not have to take my beliefs about mathematics into account. These beliefs are completely irrelevant for the change taking place and should not play any role in it.

How do we determine which of the beliefs from a belief base are relevant for a certain operation? We are interested in the concept of a *compartment* around a formula or a set of formulas. In [Hansson and Wassermann, 1999], given two sets of formulas A and B, we define the relevant formulas in B for A as those that contribute for proving or disproving any formula of A. We formalize this using the notion of *kernel sets* - minimal subsets implying a given sentence:

DEFINITION 1. [Hansson, 1994] Let C be an inference operation on $\mathcal{L}$. Then the kernel operation $\perp\!\!\!\perp_C$ is the operation such that for all subsets B and elements α of $\mathcal{L}$, $X \in B \perp\!\!\!\perp_C \alpha$ if and only if:

1. $X \subseteq B$
2. $\alpha \in C(X)$
3. for all Y, if $Y \subset X$ then $\alpha \notin C(Y)$

The elements of $B \perp\!\!\!\perp_C \alpha$ are called α-kernels.

We write $\perp\!\!\!\perp$ as an abbreviation of $\perp\!\!\!\perp_{Cn}$, where Cn is a classical consequence operation.

The compartment of B around a set of sentences A is given by the union of the α and $\neg\alpha$-kernels of B for all elements α of A. We discard the inconsistent kernels, since we do not want an inconsistent kernel like $\{p, \neg p\}$

to be included in the compartment around q only because it implies every sentence in the language.

DEFINITION 2. [Hansson and Wassermann, 1999] The A-compartment of B, where A and B are sets of sentences is defined as:

$c(A,B) = \bigcup_{\alpha\in A}(\bigcup((B \perp\!\!\!\perp \alpha) \cup (B \perp\!\!\!\perp \neg\alpha) \setminus (B \perp\!\!\!\perp \bot))$

We call c a compartmentalization function.

Note that the compartments defined above are usually overlapping and should not be seen as a partition of the belief base into different topics or subjects.

Note also that the definition of compartment around a set as the union of the compartments around each of its elements has some limitations. It may be the case that for some formulas α, β and γ, α is not relevant for β or for γ, but is relevant for the set $\{\beta, \gamma\}$. An example due to Dubois and cited in [Herzig, 1997] lets α be "I take a bath", β be "I use a hairdryer" and γ be "I die".

OBSERVATION 3.

(1) For all sets A and B of sentences, $c(A,B) = c(A, c(A,B))$.

(2) If $A \subseteq A'$ and $B \subseteq B'$, then $c(A,B) \subseteq c(A',B')$.

Even though the way the compartments are defined is purely logical, they depend on the belief base (and hence, additions or subtractions change the compartments) and on the set around which the compartment is built, which can be seen as giving the context in which the relevant formulas have to be retrieved.

Given the notion of compartment around a set A, we can define the localization of a given inference operation C to A as:

DEFINITION 4. [Hansson and Wassermann, 1999] Let C be an inference operation on $\mathcal{L}$ and let c be a compartmentalization function as in definition 2. Then for any set A, the A-localization of C is the inference operation C_A such that for all sets B of sentences: $C_A(B) = C(c(A,B))$.

A set B is A-*locally consistent* if and only if $\bot \notin C_A(B)$

THEOREM 5. *[Hansson and Wassermann, 1999] Let be C_A the A-localization of an inference operation C. Then:*

1. *If C satisfies monotony, then so does C_A.*
2. *If C satisfies monotony and compactness, then so does C_A.*

We also showed how local versions of the operations of contraction, revision, consolidation and semi-revision can be constructed and characterized by means of postulates. I reproduce below one of the results obtained in [Hansson and Wassermann, 1999], namely the characterization of local kernel contraction, which can be used for deriving the others.

The idea behind kernel contraction is that, if we remove from the belief base at least one element of each α-kernel (minimal subset of the base that implies α), we obtain a belief base that does not imply α. In the case of the local operation, we are interested in obtaining a belief base that does not A-locally imply α, where A is a set of sentences. We consider then only those kernel sets that A-locally imply α and use an *incision function* to select the formulas from the kernel sets to be removed. An incision function is a function defined from sets of sets of sentences into sets of sentences, selecting at least one sentence from each set of the argument.

DEFINITION 6. [Hansson, 1994] An *incision function* is any function $\sigma : \mathcal{P}(\mathcal{P}(\mathcal{L})) \to \mathcal{P}(\mathcal{L})$ such that for any subset S of $\mathcal{P}(\mathcal{L})$:

1. $\sigma(S) \subseteq \bigcup S$
2. If $\emptyset \neq X \in S$, then $X \cap \sigma(S) \neq \emptyset$

Example: Let $S = \{\{a, b\}, \{c\}, \{d, e, f\}\}$. Then we may have an incision function σ such that $\sigma(S) = \{b, c, e\}$.

DEFINITION 7. [Hansson, 1994] Let C be an inference operation on $\mathcal{L}$ and σ an incision function. The local kernel contraction of B determined by C and σ is the operation $\dot{-}_{C,\sigma}$ such that for all sets of sentences B:
$B \dot{-}_{C,\sigma} \alpha = B \setminus \sigma(B \perp\!\!\!\perp_C \alpha)$

The following theorem characterizes the operation of local kernel contraction and is a generalization of the result obtained in [Hansson, 1994]. The intended interpretation is that C should be a localization C_A of an inference operator, as defined above. However, the formal result is more general and is therefore stated in terms of an operation C of a more general type.

THEOREM 8. *[Hansson and Wassermann, 1999] Let C be an inference operation satisfying monotony and compactness. An operation $\dot{-}$ is an operation of local kernel contraction determined by C and some incision function if and only if for all sets of sentences B:*

- *If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)*
- *$B \dot{-} \alpha \subseteq B$ (inclusion)*
- *If $x \in B \setminus B \dot{-} \alpha$, then there is some $B' \subseteq B$ such that $\alpha \notin C(B')$ and $\alpha \in C(B' \cup \{x\})$ (core-retainment)*
- *If for all subsets B' of B $\alpha \in C(B')$ if and only if $\beta \in C(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)*

The problem with this approach is that the way the compartments are calculated is extremely inefficient. On the whole, the local change operations are computationally as costly as the original versions, although intuitively more appropriate.

However, the representation results proven in [Hansson and Wassermann, 1999] do not rely on the way the compartments are defined, but only on properties of the local inference operation obtained, namely, monotony and compactness. The compartment construction in definition 2 can be seen as an example, as one way of constructing a compartment such that the localization of the consequence operator to the compartment satisfies the requirements. This means that if we find another way of retrieving the relevant beliefs from a base such that the associated local inference operation is also monotonic and compact, the results in [Hansson and Wassermann, 1999] go through.

In the rest of the paper, some ideas for efficient ways of retrieving relevant information are presented.

Note that even if we apply traditional constructions for belief change on the set of relevant beliefs, since this set can be kept much smaller than the belief base as a whole, the operations become more feasible.

3 RABIT

In this section we briefly discuss the system RABIT (**R**easoning **A**bout **B**eliefs **I**n **T**ime). For details, the reader is referred to [Gary, 1993; Gary and Elgot-Drapkin, 1996].

RABIT is intended to simulate commonsense reasoning in an efficient, psychologically based way.

The system consists of four modules, each one representing another sort of memory: LTM (Long-Term Memory), STM (Short-Term Memory), ITM (Intermediate-Term Memory) and RTM (Relevant-Term Memory). The idea is that all the information (the beliefs) is stored in the LTM component, but the reasoning is performed on a small subset of the beliefs in LTM, the STM. This reasoning is based on the step-logic formalism [Drapkin and Perlis, 1986]. The ITM component stores the history of the reasoning process and the RTM works as a kind of context, storing the relevant concepts (in RABIT, any symbol in the language, like predicate names and constants, represent a concept). The idea of working on a small subset of the belief base is psychologically as well as computationally motivated.

RABIT has a very efficient algorithm for retrieving the relevant information from the LTM. The LTM is organized as a bipartite graph where the nodes are either formulas or concepts. Each formula is linked to all the concepts that occur in it. So, (example from [Gary and Elgot-Drapkin, 1996] the formula *penguin(tweety)* is linked to the nodes *penguin* and *tweety*. If one wants to retrieve every formula that mentions Tweety, one only needs to take all nodes adjacent to Tweety in the graph.

A method of marker-passing is used for determining which formulas of the LTM go to STM. The system chooses the relevant concepts to start with (for example, the concepts occurring in a new belief acquired) and "spreads" the activation through the edges of the graph. The activation level decreases every time it passes a new edge. Every belief with an activation level above a certain threshold is copied into the STM. By regulating the different parameters, like the initial activation, the decrease function and the threshold of activation, one can keep the STM small.

The LTM may contain contradictions, that are only solved when they become evident in the STM. Beliefs in the LTM as well as in the STM have their source attached. When a contradiction arises in the STM, a precedence order of the sources determines which belief should be given up from the STM. But nothing is deleted from the LTM. This means that a contradiction may be retrieved over and over again due to the same input. In a more recent version of RABIT, there is an "adaptive behaviour" that increases the distance between the formula that "lost" in the contradiction solving procedure and any node, so that it becomes less and less likely that the contradiction will be retrieved again.

The RABIT architecture is intuitively very appealing, but there are some aspects of it that prevent it from being useful for our purposes:

1. The retrieval mechanism is very efficient, but what happens after the relevant part of the LTM has been retrieved is that the system starts inferring everything it can from the beliefs in the STM, that is, it is not goal-directed.

2. The LTM only grows, new beliefs are added, but nothing is deleted. Besides the fact that this presents obvious disadvantages from the computational point of view, it does not seem very intuitive that an agent, after concluding that a belief is false, continues to hold it for the same reasons. The belief could certainly be reinferred by another line of reasoning but we do not want a system that follows always the same (doomed to failure) line of reasoning.

3. The way in which the LTM is organized is purely syntactical. There are no links between concepts, for example. If we know that Paul is a lecturer, that a lecturer is a member of the staff and that members of the staff are people, there is no "short path" between beliefs about people in general and beliefs about Paul.

The aspects above show that the reasoning part of RABIT is not adequated for our purposes. Nevertheless, the architecture of the program is going to be used in the rest of the paper. RABIT's architecture is based on cognitive models of memory and the fact that these models use the notions of small short term memory and relevance links between beliefs provides an independent motivation for our model of a rational agent's belief state. In the next two sections, we will explore and generalize those aspects of the RABIT architecture that will allow us to construct more efficient operations for belief change.

4 STRUCTURING THE BELIEF BASE

As we have seen, if we have an efficient method of retrieving relevant beliefs from a belief base, such that the inference relation obtained by localizing some inference operator C is monotonic and compact, we can use it together with the representation results in [Hansson and Wassermann, 1999] and get efficient, well defined operations of belief change.

In this section we show how to use extra structure of belief bases in order to make the retrieval of the relevant beliefs more efficient.

We begin by assuming that a relatedness relation between formulas of the language is given, with the intended meaning as follows:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ are directly related.

For the moment we leave open what we mean by "directly related". It may have a psychological interpretation given by statements or concepts that individuals associate, or a semantic interpretation - φ and ψ are assertions about the same (or related) topic. In section 5 we give some examples of relatedness relation.

Some desirable properties of $\mathcal{R}$ are:

1. reflexivity - $\mathcal{R}(\varphi, \varphi)$
2. symmetry - $\mathcal{R}(\varphi, \psi) \Leftrightarrow \mathcal{R}(\psi, \varphi)$
3. negation invariance - $\mathcal{R}(\varphi, \neg\varphi)$

Transitivity is not desirable, for two reasons: first, we want to be able to model agents with limited resources, who may be unable to calculate the transitive closure of the relatedness function. Second, we want to be able to talk about degrees of relatedness.

Given a relatedness relation, we can represent a belief base as a (possibly disconnected) graph where each node is a formula and there is an edge between φ and ψ if and only if $\mathcal{R}(\varphi, \psi)$. This graph representation gives us immediately a notion of degrees of relatedness: the shorter the path between two formulas of the base is, the closer related they are. Another notion made clear is that of connectedness: the connected components partition the graph into unrelated "topics" or "subjects". Sentences in the same connected component are somehow related, even if far apart.

Formally:

DEFINITION 9. Let B be a belief base and $\mathcal{R}$ be a relation between formulas. A $\mathcal{R}$-path between two formulas φ and ψ in a belief base B is a sequence $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ of formulas such that:

1. $\varphi_0 = \varphi$ and $\varphi_n = \psi$
2. $\{\varphi_1, ..., \varphi_{n-1}\} \subseteq B$
3. $\mathcal{R}(\varphi, \varphi_1)$, $\mathcal{R}(\varphi_1, \varphi_2)$, ..., and $\mathcal{R}(\varphi_n, \psi)$.

If it is clear from the context to which relation we refer we will talk simply about a path in B.

We represent the fact that P is a path between φ and ψ by $\varphi \overset{P}{\leadsto} \psi$.

The length of a path $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ is $l(P) = n$

Note that the extremities of a path in B are not necessarily elements of B.

DEFINITION 10. Let B be a belief base and $\mathcal{R}$ a relation between formulas of the language. We say that two formulas φ and ψ are related in B by $\mathcal{R}$ if and only if there is a path P in B such that $\varphi \overset{P}{\leadsto} \psi$.

Given two formulas φ and ψ and a belief base B, we can use the length of the shortest path between them in B as the degree of unrelatedness of the formulas. If the formulas are not related in B, the degree of unrelatedness is set to infinity. Formulas with a shorter path between them in B are closer related in B.

DEFINITION 11. Let B be a belief base, $\mathcal{R}$ a relation between formulas of the language and φ and ψ formulas. The unrelatedness degree of φ and ψ in B is given by:

$$u(\varphi, \psi) = \begin{cases} 0 & \text{if } \varphi = \psi \text{ and } \varphi \in B \\ \min\{l(P) | \varphi \overset{P}{\leadsto} \psi, P in\ B\} & \text{if } \varphi \text{ and } \psi \text{ are related in B by } \mathcal{R} \\ \infty & \text{otherwise} \end{cases}$$

OBSERVATION 12. A distance function on a set U is a function $d : U \times U \to \mathbb{R}$ such that: (i) $d(x, y) \geq 0$; (ii) $d(x, y) = 0$ iff $x = y$; (iii) $d(x, y) = d(y, x)$; and (iv) $d(x, y) \leq d(x, z) + d(z, y)$. The unrelatedness degree u restricted to the elements of B is a distance function.

In figure 1 we see an example of a belief base structured by a relation $\mathcal{R}$ defined by:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share an atom

This is just an example of a relatedness relation. This relation is clearly too simplistic to capture a cognitive notion of relevance. Nevertheless, it has some interesting properties which make it a good starting point for studying relevance. The intended interpretation of $\mathcal{R}(\varphi, \psi)$ is: "Given a formula φ, take every formula ψ in which we believe and that involves the atoms in φ".

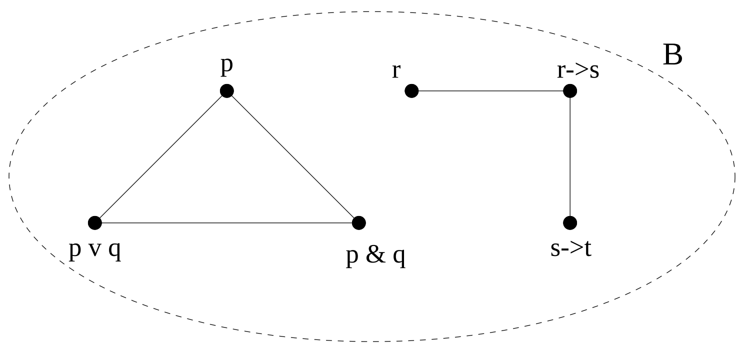

Figure 1. Structured belief base

The edges in the figure represent the relation $\mathcal{R}$ between the elements of the belief base. The graph representing the structured base has two disconnected components. The formulas $p \wedge q$ and $r \rightarrow s$ are not related in B. The formula $q \wedge r$ is related to all formulas in the belief base, since it is possible to find a path between it and all formulas in the base. Remember that the initial and final nodes of the path do not have to be elements of the base. This means that if one adds the formula $q \wedge r$ to the base, all the other elements become related to each other in B. This can be interpreted in the following way: the two disconnected components of the original base represent beliefs about unrelated subjects. As soon as one introduces a belief mentioning the subjects of both components, all beliefs in the two componnents become related to each other. There is an implicit assumption here that if one choses to add a belief $p \wedge q$ instead of two beliefs p and q, then there is some relation between p and q.

We now show, given the structure of a belief base, how to retrieve the set of formulas relevant for a given formula α:

DEFINITION 13. The set of formulas of B which are relevant for α with degree i is given by:

$$\Delta^i(\alpha, B) = \{\varphi \in B | u(\alpha, \varphi) = i\} \text{ for } i \geq 0$$

DEFINITION 14.

The set of formulas of B which are relevant for α up to degree n is given by:

$$\Delta^{\leq n}(\alpha, B) = \bigcup_{0 \leq i \leq n} \Delta^i(\alpha, B) \text{ for } n \geq 0$$

We say that $\Delta^{\leq \omega}(\alpha, B) = \bigcup_{i \geq 0} \Delta^i(\alpha, B)$ is the set of relevant formulas for α.

In section 2 we presented a notion of compartment around a sentence. In definition 2 the α-compartment of a belief base B was defined as the set of formulas of B that contribute to proving either α or its negation. These were the formulas of B that were relevant for an operation of belief change. As we maentioned before, for the representation results obtained in [Hansson and Wassermann, 1999] the particular construction of the compartments does not matter, but only the properties of the inference relation obtained. We now show some properties of the inference operation obtained using $\Delta^{\leq n}$ instead of c (definition 2) as a compartmentalization function.

PROPOSITION 15. *Let B be a belief base and α a formula. For every natural number n and any inference operator C, if C is monotonic and compact, then the local inference operations defined as $C^n_\alpha(B) = C(\Delta^{\leq n}(\alpha, B))$ are monotonic and compact.*

Proof: Since all sets considered are finite, compactness follows trivially. For monotony, let B and D be sets of formulas such that $B \subseteq D$. Let $\beta \in C^n_\alpha(B)$, i.e., $\beta \in C(\Delta^{\leq n}(\alpha, B))$. It is easy to see that $\Delta^{\leq n}(\alpha, B) \subseteq \Delta^{\leq n}(\alpha, D)$. Hence, since C is monotonic, $\beta \in C(\Delta^{\leq n}(\alpha, D))$, i.e., $\beta \in C^n_\alpha(D)$. □

This means that for any n, $\Delta^{\leq n}$ can be used as a compartmentalization function to define local operations characterized by the results presented in [Hansson and Wassermann, 1999] and summarized in section 2. In the case of an ideal agent, that is, an agent with no limits in its capacity of retrieval, all of the agent's relevant beliefs are retrieved, that is, $\Delta^{\leq\omega}$ is used as a compartmentalization function. Otherwise, we can limit the size of the set retrieved by the choice of n.

In figure 2, we see an example of a structured belief base $B = \{\alpha, \beta, \gamma, \delta, \varepsilon, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, o, \pi, \rho, \sigma, \phi, \varphi, \chi\}$. The dotted circles represent different levels of relevance for α. We have:

$\Delta^0(\alpha, B) = \{\alpha\}$
$\Delta^1(\alpha, B) = \{\beta, \chi, \delta, \varepsilon\}$
$\Delta^2(\alpha, B) = \{\gamma, \eta, \iota, \varphi, \kappa, \lambda, \mu, \phi\}$
$\Delta^3(\alpha, B) = \{\nu, o, \pi, \theta, \rho, \sigma\}$
$\Delta^{\leq\omega}(\alpha, B) = \Delta^0(\alpha, B) \cup \Delta^1(\alpha, B) \cup \Delta^2(\alpha, B) \cup \Delta^3(\alpha, B) = B$

5 WHERE DOES THE STRUCTURE COME FROM?

In the previous section we assumed that a relation of relatedness between elements of the language was given. In this section we will present some ways of deriving such a relation from the belief base.

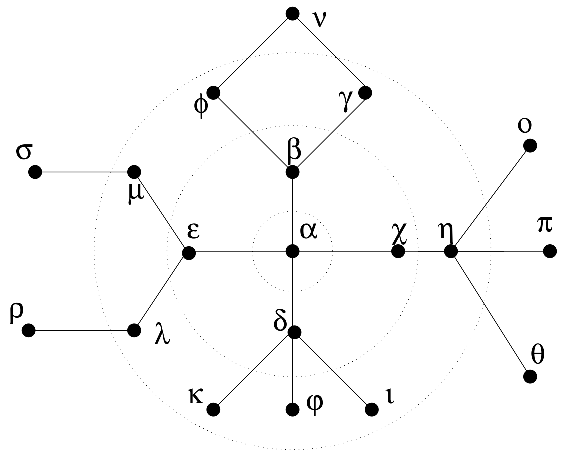

Figure 2. Degrees of relevance

5.1 Database definitions

In many real applications, there is a part of the database dedicated to the definition of concepts which can be used to generate a relatedness relation.

Let B be a database and A be the subset of B where all the definitions are made. We can say that atoms appearing in the same formula of A are related. We can say that two formulas φ and ψ are related if they share an atom or if there are atoms p in φ and q in ψ such that p and q are related.

Formally:

DEFINITION 16. Let A be a set of formulas. We define a relation R_A on atoms by:

$R_A(p, q)$ if and only if either $p = q$ or p and q occur in the same formula of A.

The relation R_A can be extended to a relation on the language:

DEFINITION 17. Let $\mathcal{R}_A$ be a relation on formulas defined as:

$\mathcal{R}_A(\varphi, \psi)$ iff there are atoms p in φ and q in ψ such that $R_A(p, q)$.

It is easy to see that $\mathcal{R}_A$ is symmetric and reflexive and that for every φ, $\mathcal{R}_A(\varphi, \neg\varphi)$.

The transitive closure of the relation $\mathcal{R}_A$ gives an equivalence relation that determines a partition of the belief base. Each member of the partition is completely independent from the others with respect to the definitions in A.

We can also have $A = B$, that is, use the whole belief base for deriving the relatedness relation.

5.2 The syntactic approach

Rodrigues [Rodrigues, 1997] claims that to ask for a primitive relation between propositional variables or a primitive subject matter assignment may be to strong a requirement. He proposes instead to use directly the relation introduced in figure 1, i.e.:

$\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share an atom

He shows that this is actually the smallest relation satisfying the conditions given in [Epstein, 1990] for a relatedness relation.[3]

Epstein [Epstein, 1990] takes the topic or subject matter of a proposition as a primitive and says two propositions are related if they share a subject matter. He considers a set of topics $\mathcal{S}$ and supposes that each propositional variable has a non-empty set of topics associated with it. The set of topics associated with a proposition is simply the union of the sets of topics associated to each variable appearing in it. Epstein shows that one can also take a relatedness relation $\mathcal{R}$ as a primitive and derive a topic assignment from it by taking the subject matter of a formula φ to be $s(\varphi) = \{\{\varphi, \psi\} | \mathcal{R}(\varphi, \psi)\}$. Clearly, the result of applying the definition of relatedness to this topic assignment, i. e., $\mathcal{R}(\varphi, \psi)$ if and only if $s(\varphi) \cap (\psi)$ is non-empty is the relation $\mathcal{R}$.

Another way of deriving the relatedness structure from the given data is to consider the knowledge base as a whole and create a graph that has as nodes formulas and atoms. Each formula is then linked to the atoms that occur in it. This is the approach followed by the RABIT system. The derived structure is that of a (possibly disconnected) bipartite graph. Atoms can be related to each other only via some formula in the database that contains them. In the same way, formulas are only related to other formulas via some atom they share. This has the advantage that it is very easy to insert a new formula in the graph, one only has to link it to the atoms it contains. On the other hand, since there is no link between atoms, even if the atoms p and q are intuitively directly related, the formulas p and q will only be related if there is any formula in the belief base containing both. One could use tricks, like adding some formula *related(p,q)* to the belief base if the agent believes that p and q are related or adding a tautology like $(p \vee \neg p) \vee (q \vee \neg q)$ to the belief base in order to relate p and q.

[3] According to [Epstein, 1990], any relatedness relation R should satisfy:
R1 - $R(\varphi, \varphi)$
R2 - $R(\varphi, \psi)$ iff $R(\neg\varphi, \psi)$
R3 - $R(\varphi, \psi)$ iff $R(\psi, \varphi)$
R4 - $R(\varphi, \gamma \to \psi)$ iff $R(\varphi, \gamma)$ or $R(\varphi, \psi)$
R5 - $R(\varphi, \gamma \wedge \psi)$ iff $R(\varphi, \gamma \to \psi)$.

Epstein's and RABIT's approaches have the property that the relatedness relation is defined on the whole language and is independent from the particular belief base. This means that adding beliefs to the structure consists of adding the input formula α to the set of nodes and adding an edge between α and every node β such that $\mathcal{R}(\alpha, \beta)$. Deleting a belief α from the structure consists of removing α from the set of nodes and removing all the edges leaving from it. After adding or deleting a belief, the result is another structured base. The relation $\mathcal{R}$ does not have to be recalculated.

5.3 The logical approach

We can define a relatedness relation $\mathcal{R}$ that captures a notion closer to the logical compartments in section 2 as:

$\mathcal{R}(\varphi, \psi)$ if and only if there is a set $A \subseteq B$ such that $A \cup \{\varphi\} \not\vdash \bot$ and either $A \not\vdash \psi$ and $A \cup \{\varphi\} \vdash \psi$ or $A \not\vdash \neg\psi$ and $A \cup \{\varphi\} \vdash \neg\psi$

or:

$\mathcal{R}(\varphi, \psi)$ if and only if $\varphi \in c(\psi, B)$

where c is a compartmentalization function as in definition 2.

Note that this relation is not symmetric. The definitions of path, unrelatedness degree and relevant set of formulas (Δ^ω) in section 4 can be maintained. The obvious problem of this approach is that it may be very hard to find the related pairs, but if the relation is given, this problem is avoided. If the language is finite, one way in which the relation can be given is as follows: All sets of formulas can be listed. For each pair of formulas φ and ψ, one can select all sets where φ occurs, check whether they imply ψ or $\neg\psi$ and if so, check their subsets to see whether they also imply ψ or $\neg\psi$. If not, then $\mathcal{R}(\varphi, \psi)$. This is of course very costly, but if the same language is going to be used several times, one can pre-compute this relation and use it for all belief bases.

5.4 The hybrid approach

A fourth alternative is to combine the database or the syntactic approach with the logical one. If the database is such that it contains several (small) independent modules, this can be a good strategy. One can first apply the first notion of relatedness and find a partition of the database. Then one only has to consider the partition members related to the input formula and look for the relevant formulas inside them using the logic formulation of compartments given in section 2. It can be shown that the compartment obtained is the same that would be obtained checking the whole set.

Let B be a belief base and let $\mathcal{R}$ be the relatedness relation introduced in figure 1, i. e.: $\mathcal{R}(\varphi, \psi)$ if and only if φ and ψ share an atom. Since $\mathcal{R}$

is reflexive and symmetric, its transitive closure defines a partition of the elements of B. Let $\{B_1, B_2, ..., B_n\}$ be the elements of the partition. We say that a set B_i is related to a formula α ($\mathcal{R}(\alpha, B_i)$)if and only if $\mathcal{R}(\alpha, \beta)$ for some $\beta \in B_i$. We denote by P_α the set of elements of the partition that are related to α, i.e., $P_\alpha = \{B_i | \mathcal{R}(\alpha, B_i)\}$.

THEOREM 18. *Let B be a belief base, α be a contingent formula and c be the compartmentalization function as in definition 2.Then* $c(\alpha, B) = c(\alpha, \bigcup P_\alpha)$.

Proof: From $\bigcup P_\alpha \subseteq B$ and observation 3 it follows that $c(\alpha, \bigcup P_\alpha) \subseteq c(\alpha, B)$. For the other side of the inclusion, let $\beta \in c(\alpha, B)$. This means that there is $X \subseteq B$ such that X is consistent, inclusion minimal and $\alpha \in Cn(X)$ or $\neg\alpha \in Cn(X)$. Suppose by contradiction that $X \not\subseteq \bigcup P_\alpha$. Then, there is $\gamma \in X$ such that for all $B_i \in P_\alpha$, $\gamma \notin B_i$. From this it follows that α is not related to γ, i.e., there is no $\mathcal{R}$ path from α to γ. Since X is consistent it follows that $\alpha \in Cn(X \setminus \{\gamma\})$ or $\neg\alpha \in Cn(X \setminus \{\gamma\})$, contradicting the minimality of X. Hence, $\beta \in X \subseteq \bigcup P_\alpha$ and from parts 1 and 2 of observation 3 it follows that $c(\alpha, B) \subseteq c(\alpha, \bigcup P_\alpha)$. □

As an example, let $B = \{p, q, p \to q, s, t \vee v, s \to v, w \wedge x, x \to z, z \vee w\}$ and let c be a compartmentalization function as in definition 2. The compartment around $p \vee w$ is given by $c(p \vee w, B) = \{p, w \wedge x\}$. The set B can be clearly divided into three unrelated components: $B_1 = \{p, q, p \to q\}$, $B_2 = \{s, t \vee v, s \to v\}$ and $B_3 = \{w \wedge x, x \to z, z \vee w\}$. The formula $p \vee w$ is only related to two of these components, namely B_1 and B_3. We can then disconsider B_2 and calculate the compartment by: $c(p \vee w, B) = c(p \vee w, B_1 \cup B_3) = \{p, w \wedge x\}$.

6 COMPUTATIONAL ASPECTS

In this section we discuss the complexity of the operation described for retrieving the set of relevant beliefs from a base.

An anytime algorithm is one that whenever it is interrupted has built an approximate solution for a problem, and the longer it runs, the better the approximation gets.

The good thing about the method for retrieving the relevant beliefs is that it is an anytime method, that is, whenever it is interrupted, it has retrieved the most relevant beliefs, and the longer it runs, the closer it gets to retrieving all the relevant beliefs (the maximal connected subgraph). This is a very desirable property for modeling agents that may not have enough time or memory to find all the related beliefs. In the ideal case, if there is no resource limitation, the method succeeds in retrieving a maximal connected subgraph.

Below we present a sketch of an algorithm that takes as input a formula

α and a belief base and returns the set of formulas of the base that are relevant for α. The algorithm can be stopped at anytime, always returning the set of most relevant beliefs for α. The algorithm is a modification of the algorithm BFS for breadth first search in [Cormen *et al.*, 1990].

The belief base is represented by a vector of formulas, each one with a list of pointers to the adjacent nodes in the graph. The nodes adjacent to a formula α are given by Adjacent(α). The complexity of the construction of the list of adjacent nodes depends on the relatedness relation used. For the relation used in figure 1, we can use extra, invisible nodes corresponding to the atoms of the language. Every formula added to the belief base is linked to all atoms appearing on it. Constructing the list has then complexity $O(m \times n)$, where m is the number of occurences of atoms in the formulas involved (the "size" of the belief base) and n is the number of atoms in the language. If the atoms are organized in some kind of lexicographical order, this complexity becomes $O(m \times \log n)$.

Retrieve(α,B,Relevant):
1. If $\alpha \in B$, then mark(α)
2. $\Delta^1(\alpha, B)$:= Adjacent(α)
3. i := 1; stop := false
4. While not stop do
    - 4.1. For all $\beta \in \Delta^i(\alpha, B)$ mark(β)
    - 4.2. i := i+1
    - 4.3 For all $\beta \in \Delta^{i-1}(\alpha, B)$,
      $\Delta^i(\alpha, B) := \Delta^i(\alpha, B) \cup \{\varphi \in$Adjacent$(\beta)$ s.t. not marked$(\varphi)\}$
    - 4.4 If $\Delta^i(\alpha, B) = \emptyset$, then stop := true
5. Relevant := $\{\beta \in B$ s.t. marked$(\beta)\}$

At each step, the algorithm looks at the set retrieved at the previous step and gets all the neighbours that have not been visited yet. When all nodes of a connected component have been visited, it halts. Depending on how the information is encoded, the algorithm runs in linear time.

After retrieving the relevant beliefs, traditional belief change constructions can be applied to it, provided the set is small enough (one can stop the algorithm once the relevant set gets bigger than a certain limit).

7 RELATED WORK

The idea of isolating inconsistencies by means of compartments has been studied by several authors. Most of these did not use the notion of relevance to create the compartments.

Jaśkowski introduced his discursive logics as a way of formalizing "what some participant of a discourse is committed to" [Jaśkowski, 1969]. If participant x asserts α and participant y asserts β, then α and β are true in his logics, but not $\alpha \wedge \beta$, since there may be no participant of the discourse that is committed to both assertions α and β. Assertions associated to different participants of the discourse cannot be mixed.

Lewis [Lewis, 1982] also defended the idea of fragments that cannot be mixed, but the different (possibly overlapping) fragments were part of a single agent's beliefs.

Fagin and Halpern [Fagin and Halpern, 1988] formalized a similar idea. In their Logic of Local Reasoning, an agent may have several "frames of mind". An agent believes α if he beliefs α in some frame of mind. This idea is formalized via an extended Kripke structure, where instead of one accessibility relation between worlds, there is a relation $\mathcal{C}$ between worlds and sets of worlds. If $\mathcal{C}(s) = \{T_1, ..., T_n\}$, then an agent in state s sometimes takes the set of worlds T_1 as possible, sometimes T_2, etc.

Benferhat et al. [Benferhat *et al.*, 1997] use yet another approach to isolate inconsistencies, which is very similar to the approach known in AI as WIDTIO (When In Doubt Throw It Out) [Winslett, 1990]. They define a free (or sound) inference operator that disconsiders all formulas belonging to a minimal inconsistent subset. So, $B \vdash_{free} \alpha$ if and only if $Free(B) \vdash \alpha$, where $Free(B)$ is the result of deleting every minimal inconsistent subset from B.

There are in the literature several attempts to define the concepts of relevance and dependence. In [del Cerro and Herzig, 1996] a set of postulates is presented that any dependence relation should satisfy. They show then how a contraction operation satisfying the Gärdenfors postulates can be obtained from a dependence relation.

Their dependence relations are syntax independent and only logically independent formulas can be related. From this it seems clear that the notion of relevance captured by the postulates is very different from the one we are trying to model. Our notion is independent of the logic.

The idea of using graphs representing dependency relations for finding the relevant part of a database is well-known in the field of logic programming (see, for example, [Kowalski, 1979]). In the case of logic programs, the dependency relation is not symmetric and the graph obtained is a directed one.

8 CONCLUSIONS AND FURTHER WORK

We have presented a way of introducing more structure to belief bases so that relevant beliefs can be retrieved in an efficient way. The belief bases are represented by a graph, instead of by a set of formulas, where the formulas the agent believes are the nodes and the edges represent a relatedness relation. We have shown some example of relatedness relations and how they can be obtained.

We presented an algorithm to retrieve the most relevant beliefs from a structured belief base and showed that the algorithm could be used instead of the compartmentalization function defined in [Hansson and Wassermann, 1999]. In this way, we obtained efficient and well characterized operations of belief change.

In more realistic agents, the whole belief base may be connected, that is, it may be impossible to isolate a small connected component. In such cases, one can think about more sophisticated notions of connectedness, like finding connected "chunks" in the graph and preferring edges internal to a chunk over the others (see, for example, [van Dongen, 1997]).

The structure of the belief base is used now only to retrieve the most relevant beliefs. Then traditional operations for belief change are applied on the set of beliefs retrieved. Further work includes investigating how to use the structure of the belief base also for implementing more efficient operations of belief change.

Acknowledgments: I would like to thank Chris Albert, Kevin Gary, Hans Rott and Frans Voorbraak for comments on an earlier draft. This research is sponsored by a grant from the Brazilian funding agency CAPES and the Dutch Research Council PIONIER-project Reasoning With Uncertainty.

REFERENCES

[Alchourrón *et al.*, 1985] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change. *Journal of Symbolic Logic*, 50, 1985.

[Benferhat *et al.*, 1997] Salem Benferhat, Didier Dubois, and Henri Prade. Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study. part 1: The flat case. *Studia Logica*, pages 17–45, 1997.

[del Cerro and Herzig, 1996] Luis Fariñas del Cerro and Andreas Herzig. Belief change and dependence. In Yohav Shoham, editor, *Proceedings of TARK 96*, 1996.

[Cormen *et al.*, 1990] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms.* MIT Press, 1990.

[Dixon, 1994] Simon E. Dixon. *Belief Revision: a Computational Approach.* PhD thesis, Basser Department of Computer Science, University of Sidney, 1994.

[van Dongen, 1997] Stijn van Dongen. Graph clustering and information structuring. In E. Kraak and R. Wassermann, editors, *Proceedings of the Accolade 97.* Dutch Graduate School in Logic, 1997.

[Drapkin and Perlis, 1986] Jennifer Drapkin and Don Perlis. Step-logics: An alternative approach to limited reasoning. In Wolfgang Wahlster, editor, *Proceedings of the European Conference on Artificial Intelligence (ECAI).* Wiley, 1986.

[Epstein, 1990] Richard L. Epstein. *The semantic foundations of logic, volume 1: Propositional Logic.* Nijhoff International Philosophy Series. Kluwer Academic Publishers, 1990.

[Fagin and Halpern, 1988] Ronald Fagin and Joseph Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

[Fuhrmann, 1991] André Fuhrmann. Theory contraction through base contraction. *Journal of Philosophical Logic*, 20:175–203, 1991.

[Gary and Elgot-Drapkin, 1996] Kevin Gary and Jennifer Elgot-Drapkin. RABIT-Bridging formal and implementational approaches to commonsense reasoning. Technical report, Arizona State University, 1996. Available at http://www.public.asu.edu/ kgary/.

[Gary, 1993] Kevin Gary. RABIT - A spreading activation approach to real-time commonsense reasoning. Master's thesis, Arizona State University, Tempe, Arizona, 1993. Available at http://www.public.asu.edu/ kgary/.

[Hansson and Wassermann, 1999] Sven Ove Hansson and Renata Wassermann. Local change. In preparation (a preliminary version was presented at the Fourth Symposium on Logical Formalizations of Commonsense Reasoning, London, 1998), 1999.

[Hansson, 1994] Sven Ove Hansson. Kernel contraction. *Journal of Symbolic Logic*, 59:845–859, 1994.

[Hansson, 1998] Sven Ove Hansson. *A Textbook of Belief Dynamics*. Kluwer Academic Press, 1998.

[Herzig, 1997] Andreas Herzig. How to change factual beliefs using laws and independence information. In Dov M. Gabbay, Rudolf Kruse, Andreas Nonnengart, and Hans Juergen Ohlbach, editors, *Qualitative and Quantitative Practical Reasoning - First Int. Joint Conf. on Qual. and Quant. Practical Reasoning; ECSQARU-FAPR'97*, number 1244 in LNCS, pages 311–321. Springer, jun 1997.

[Jaśkowski, 1969] Stanislaw Jaśkowski. Propositional calculus for contradictory deductive systems. *Studia Logica*, XXIV:143–157, 1969. Translation from the Polish original that appeared in Studia Societatis Scientiarum Torunensis, Vol I, No. 5, 1948.

[Kowalski, 1979] Robert Kowalski. *Logic for Problem Solving*. North Holland, 1979.

[Lewis, 1982] David Lewis. Logic for equivocators. *Noûs*, 16:431–441, 1982.

[Nebel, 1990] Bernhard Nebel. *Reasoning and Revision in Hybrid Representation Systems*, volume 422 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, 1990.

[Rodrigues, 1997] Odinaldo T. Rodrigues. *A Methodology for Iterated Information Change*. PhD thesis, Imperial College, University of London, 1997.

[Rott, 1991] Hans Rott. A nonmonotonic conditional logic for belief revision. In André Fuhrmann and Michael Morreau, editors, *The logic of theory change*. Springer Verlag, 1991. LNAI 465.

[Rott, 1996] Hans Rott. Making up one's mind: Foundations, coherence, nonmonotonicity. Habilitationsschrift, University of Konstanz. To appear as *Change, Choice, and Inference*, Oxford University Press, October 1996.

[Wassermann, 1999] Renata Wassermann. Resource-bounded belief revision. *Erkenntnis*, 1999. to appear.

[Williams, 1994] Mary-Anne Williams. On the logic of theory base change. In *Proceedings of the European Workshop on Logics in Artificial Intelligence (JELIA'95)*, 1994.

[Winslett, 1990] Marianne Winslett. *Updating Logic Databases*. Number 9 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1990.