Beyond Foundational Reasoning

Simon Dixon and Norman Foo
Knowledge Systems Group
Basser Department of Computer Science
University of Sydney, NSW 2006
Australia

e-mail: simon@cs.su.oz.au, norman@cs.su.oz.au

ABSTRACT

Implementations of logical reasoning systems are invariably based on the foundational theory of justification. The AGM logic is based on the coherence theory and the principle of minimal change, but this presents difficulties in its implementation. We show that both the foundational and coherence styles of reasoning are necessary for practical reasoning and that both can be achieved in the AGM logic by choosing particular epistemic entrenchment orderings. We also describe the limitations which must be imposed on the logic in order to implement the system.

1. Introduction

In the theory of justification, two schools of thought have emerged. These are the *foundational* and *coherence* theories of justification[1]. Both styles present formalisms for determining when an agent is justified in holding a particular belief, and are based on examining the relationships between the various beliefs in the current belief set. They differ in their views of what constitutes a valid justification.

The foundationalist school claims that for a proposition to be believed there must be a set of propositions which justify it and which are also all currently believed. These justifying propositions must also have their own set of supporting propositions, which must be justified also, regressing back to a set of self-justifying *foundational* beliefs. In contrast to this, the coherence theory has no such requirement for a belief to be justified, but instead accepts those propositions which improve the overall coherence of the belief set.

A similar pair of methodologies is found in the belief revision literature[2,3]. Here the distinction is not so much in what causes a belief to be accepted initially, but in what causes the belief to persist or else be removed from the belief set. When a justification is no longer valid, the foundational theory requires that the belief it justified be removed from the belief set, whereas the coherence theory allows it to remain, unless it is incoherent with the rest of the new belief set. The underlying principle in the coherence theory of belief revision is minimal change; that is, when adding new beliefs or retracting old beliefs, the aim is to modify the belief set as little as possible without sacrificing the consistency and coherence of the set. The result of this is that a belief which has been justified at some time can still be held to be true, even after some or all of its justifying propositions have been removed from the belief set.

The aim of this paper is firstly to show that both styles of reasoning are necessary, and secondly to demonstrate how the AGM logic with epistemic entrenchment can be used to model both of these methods of reasoning. Various difficulties are encountered when moving from the AGM theory, in which the belief sets are logically closed, and hence usually infinite, to an implementation where the belief sets are necessarily only finite representations of the infinite theories. We discuss a way of overcoming these difficulties, by using a partial closure operation, and we show the limitations of this approach.

Sections 2 and 3 of the paper describe the foundational and coherence theories respectively, and give examples of both styles of reasoning. In section 4, the AGM logic of belief revision is presented, and it is shown that epistemic entrenchment can be used to achieve both foundational and coherence style behaviour in this logic. The 5th section describes an implementation of this approach, and how the limitations of using finite bases rather than infinite theories affect the AGM logic. It is also shown that there are cases in which the finite base implementation can be extended to recover the coherence style behaviour. The paper concludes with a description of current research on this system.

2. The Foundational Theory

The foundational theory requires each proposition to have a set of supporting beliefs as a justification, so that if one of these supporting propositions is not believed, then the proposition itself is not believed, unless there is an alternative justification. The supporting justifications must also have their own justifications, which in turn are also justified, and hence chains of justifications are formed which link together the members of the belief set. Two limitations are placed on these chains of beliefs. Firstly, they must be finite; that is, they must end in a set of beliefs which are self-justifying. These beliefs are called the *foundational beliefs*. Secondly, a cyclic chain of beliefs is not accepted as a valid justification, so that no belief can form part of its own justification. The only exception to this is the foundational beliefs, which are totally justified by themselves, and have no need for justification from other beliefs.

The foundational theory also has ramifications on the dynamic behaviour of a system based on its principles. Foundationalism is impartial to history: the fact that a belief was justified in the past does not aid its credibility in the present; it must have a current justification for it to be believed. As soon as a belief is removed from the belief set, all of the beliefs which were justified by it must also be removed immediately. Sosa[4] likens this behaviour to a pyramid where each belief is represented by a brick, which supports the bricks above it and is supported by the bricks beneath it. When one brick is removed, all of the bricks built above it topple down and are no longer part of the pyramid, while the bricks beneath it and to either side remain unaffected.

There is no dispute that the ability to reason foundationally is necessary for any dynamic reasoning system. In fact, most (if not all) of the implementations of automated reasoning systems are based exclusively on the foundational method. All of the Truth (or Reason) Maintenance Systems[5,6] have this factor in common. We now show that coherence-based reasoning is also necessary to reason correctly in many domains.

3. The Coherence Theory

A number of different notions of coherence have been proposed, and these are discussed at length by Thagard[7]. *Deductive coherence* depends on logical consistency and entailment relationships between the members of the belief set. This is the style of coherence found in the AGM logic of belief. A second type of coherence is *probabilistic coherence*, which depends on propositions having probability assignments which are consistent with the axioms of probability. This requires extralogical information in the form of probability assignments for each proposition. Thirdly, there is *semantic coherence*, which exists if propositions have similar meanings, and again requires extralogical input. Finally, *explanatory coherence* depends on the existence of some type of explanatory relationship between pairs of propositions. These explanations are also expressed outside of the logical framework of the set of beliefs. In this paper we call deductive coherence weak coherence, since it is based only on the logical relationships within the belief set, whereas the other styles of coherence all rely on extralogical information and hence describe a stronger notion of coherence.

The main principle of the coherence theory applied to belief dynamics is minimal change. This is especially true of weak coherence, where a belief will not be removed from a belief set unless it is explicitly required to be removed or it would cause the system to be inconsistent if it were to remain. The AGM logic is based on weak coherence, but it is possible to achieve a stronger notion of coherence in the system by adding an epistemic entrenchment ordering on the beliefs, which will be discussed in section 4.

Since the coherence theory has not featured as greatly as the foundational theory in literature on dynamic reasoning, one may question its necessity. The coherence theory is often presented as the way in which humans reason, and while this does not in itself justify the need for coherence-based reasoning, it does give us reason to examine the issues carefully. There is, in fact, at least one large class of examples which require coherence-based reasoning. This is modelling actions which have consequences which persist beyond the duration of the action, or which are irreversible.

For example, we may model a simple electrical circuit containing a variable voltage source, a lamp and a fuse in series, and include a rule that states:

$$Exceeds(Current, Fuse\_Rating) \rightarrow Blow(Fuse)$$

Clearly, as soon as the fuse blows, the current will stop flowing, and hence the antecedent of the rule is no longer true. Nevertheless, the fuse is still blown. Now if we were to reason foundationally using this rule, then as soon as the current stopped flowing we would believe that the fuse stopped being blown. This may be a trivial example, but the same reasoning applies to any irreversible action, and to any action with persistent consequences. For these actions, a minimal change logic is required. This does not in any way lessen the need for foundational reasoning, as it is not difficult to revise the above example by replacing the fuse with a circuit-breaker, which breaks the circuit when the current exceeds a limit, but then restores it again after a small time delay. In this case the circuit is broken when the current is too high, but when the current stops, the break is restored. In section 4, we will consider these two examples in greater detail.

Other styles of non-foundational reasoning are discussed by Cornman[8], but these are shown to be of no practical import. This leaves us with the need for a system which can reason with two different types of rules: foundational and coherence-based rules. We now show that the AGM logic with epistemic entrenchment is such a system, and discuss the practical issues involved in implementing the system.

4. The AGM Logic and Epistemic Entrenchment

The AGM logic[9] models the current state of belief by a belief set, denoted K, which is a logically closed set of first order formulae, and an ordering on these formulae, called epistemic entrenchment, which obeys the following axioms (for all formulae A, B, C):

(EE1) If $Ent(A) \leq Ent(B)$ and $Ent(B) \leq Ent(C)$ then $Ent(A) \leq Ent(C)$
(EE2) If $A \vdash B$ then $Ent(A) \leq Ent(B)$
(EE3) For any A and B, $Ent(A) \leq Ent(A \wedge B)$ or $Ent(B) \leq Ent(A \wedge B)$
(EE4) When $K \neq K_{\perp}$, $A \notin K$ iff $Ent(A) \leq Ent(B)$ for all B
(EE5) If $Ent(B) \leq Ent(A)$ for all B then $\vdash A$

Three different operations may be performed on this belief set: expansion, contraction and revision. Expansion by α, denoted K^+_α, involves adding the belief α and its consequences to the belief set K. Contraction, denoted K^-_α, is the removal of α and sufficient other formulae so that the resulting belief set does not contain α. The revision operation, K^*_α, is the incorporation of the new belief α into K such that K remains consistent. That is, any beliefs which contradict α must be removed from K before α is added. Revision may be seen as the composition of a contraction and an expansion operation, as specified by the Levi identity:

$$K^*_\alpha = (K^-_{\neg\alpha})^+_\alpha$$

Contraction and revision cannot be specified set-theoretically, as there are in general many solutions which satisfy the axioms for rational contraction and revision operators. Uniqueness is obtained via the epistemic entrenchment ordering on the beliefs, and one further condition, the (C-) condition:

$$B \in K^-_A \text{ iff } B \in K \text{ and either } \vdash A \text{ or } Ent(A) < Ent(A \vee B)$$

Informally, $Ent(\alpha) < Ent(\beta)$ means that it is easier to give up belief in α than to give up belief in β.

It has been shown[10] that it is possible for an entrenchment to be specified so that the AGM logic may replicate the behaviour of a purely foundational system such as the ATMS[6]. This is done by calculating the *essential supports* of each belief, which are the set of foundational beliefs which, if anyone one of them was removed, would cause the removal of this belief. From the (C-) condition, if α is an essential support of β, then $Ent(\alpha) = Ent(\alpha \vee \beta)$, otherwise $Ent(\alpha) < Ent(\alpha \vee \beta)$. (Note that although the ATMS does not require acyclicity in the chains of justifications, this does not mean that it is not a foundational system. The ATMS does not consider a circular chain of justifications to be sufficient in itself to justify a belief. It still requires that each belief have at least one justification which can be traced back to foundational beliefs.)

We now show that coherence-based reasoning can also be implemented in the AGM logic. Consider a propositional theory with justifications represented by Horn clauses. Suppose we have a rule $\alpha \rightarrow \beta$ in K, and $\alpha \notin K$ and $\beta \notin K$. Then if we implement the rule as a coherence-based rule, we will have $\beta \in (K^+_\alpha)^-_\alpha$ This will be the case, according to the (C-) condition, if and only if, in K^+_α, we have $Ent(\alpha) < Ent(\alpha \vee \beta)$. This entrenchment is a necessary and sufficient condition for $\alpha \rightarrow \beta$ to be treated as a coherence-based rule.

For the example described in section 3, where we have the rule

$$Exceeds(Current, Fuse\_Rating) \rightarrow Blow(Fuse),$$

we note that this is a coherence style rule, and when the belief set is revised by the fact

$$Exceeds(Current, Fuse\_Rating),$$

it is given the entrenchment

$$Ent(Exceeds(Current, Fuse\_Rating)) < Ent(Exceeds(Current, Fuse\_Rating) \vee Blow(Fuse)).$$

This ensures that when the current drops, the reasoner still contains the belief that the fuse is blown.

Alternatively, for the second example, where a circuit-breaker is used instead of a fuse, we encode this as a foundational rule, using the entrenchment

$$Ent(Exceeds(Current, Fuse\_Rating)) = Ent(Exceeds(Current, Fuse\_Rating) \vee Trigger(Circuit-Breaker))$$

5. Implementation of the AGM Logic

In building a system based on the AGM logic, we are immediately faced with some unavoidable problems. The AGM logic relies on the logical closure of belief sets, which means that belief sets are infinite in the general first order case, and even if reduced to the propositional case, belief sets are still exponential in the number of propositions in the language. (For if $\alpha \in K$, then since $\alpha \vdash \alpha \vee \phi$ for all formulae ϕ, we must have $\alpha \vee \phi \in K$.) Therefore belief sets must be represented by a finite base of the full theory. Now if we calculate none of the consequences of the finite base, we lose some of the properties of the AGM logic. For example, if our knowledge base is $B = \{p,\ p \rightarrow q\}$, and we remove the proposition p, without performing any closure on the base, we obtain $B_p^- = \{p \rightarrow q\}$, which corresponds to the rule $p \rightarrow q$ being treated foundationally. In the AGM logic, if we have the entrenchment $Ent(p) < Ent(p \vee q)$, we get $q \in K_p^-$, which corresponds to treating $p \rightarrow q$ as a coherence-based rule. It is true in general that without closure, the coherence style reasoning of the AGM is lost, whereas the foundational nature is retained.

The reason for this is that foundational reasoning requires each belief to have a justification which is currently true. This justification is based on the foundational beliefs, which have no justifications themselves, and therefore they must be members of the finite base. Then any belief based on these foundational beliefs can be derived from the current belief base.

On the other hand, reasoning in the coherence style is dependent on previous contents of the finite bases. That is, if a belief had been justified, and its justification had since been falsified, but there was no reason to remove the belief itself, then under the coherence system, the belief should be retained. Now suppose this belief was not included in the base, but was a consequence of beliefs in the base. Then without closing the base, this belief is not explicitly accepted, and once the justification is removed, it is no longer derivable. Thus beliefs generated from coherence style rules must be mentioned explicitly in the belief base before their justifications are invalidated, or else they will not be retained in the theory.

We conclude that we must perform at least partial closure of the belief base. Full closure is intractable, so the aim is to expand the base to a larger base, which contains sufficient consequences of the original base to regain the results of the coherence theory. As a first approximation, consider adding to the belief base any belief which can be derived directly from coherence type rules which have all of their antecedents in the base.

For example, suppose $B(K) = \{\alpha \rightarrow \beta,\ \beta \rightarrow \gamma,\ \alpha\}$, and $\alpha \rightarrow \beta$ is a coherence type rule, and $\beta \rightarrow \gamma$ is foundational. Then in the AGM logic we need the following entrenchments: $Ent(\alpha) < Ent(\alpha \vee \beta)$ and $Ent(\beta) = Ent(\beta \vee \gamma)$, giving $\beta \in K_\alpha^-$. Now applying our first approximation to logical closure, we have $B'(K) = B(K) \cup \{\beta\}$, and $B(K_\alpha^-) = \{\alpha \rightarrow \beta,\ \beta \rightarrow \gamma,\ \beta\}$, in agreement with the AGM axioms.

Now consider a counterexample, where we have the same base, but different types of rules. That is, $\alpha \rightarrow \beta$ is foundational, and $\beta \rightarrow \gamma$ is a coherence type rule. Then the AGM gives $Ent(\alpha) = Ent(\alpha \vee \beta)$ and $Ent(\beta) < Ent(\beta \vee \gamma)$, so that $\gamma \in K_\alpha^-$, and $\beta \notin K_\alpha^-$, whereas our given method adds nothing to the base, so that neither β nor γ are in the contracted base nor derivable from it. So our approximation is insufficient. We need to add to the base any consequents of coherence type rules whose antecedents are provable from the base. This problem is NP-complete even in the propositional case, so in general it is not possible to achieve the behaviour of the AGM logic through partial closure.

Consider one special case, where none of the antecedents of coherence-based rules are justified by other rules. That is, all of the antecedents of the rule must be in the belief base for the justification to be valid. Then the "first approximation" described above gives the desired coherence style rule behaviour, as demonstrated in the first example above.

6. Conclusion

We have shown that both the foundational and the coherence styles of reasoning are necessary for reasoning about the real world, and that it is possible to obtain both types of reasoning using the AGM logic and epistemic entrenchment. The major difficulty in practice is that it is not possible to implement the AGM logic as it stands, because it is based on infinite theories as belief sets, which must be changed to some finite representation so that it can be implemented. In using a finite representation, the foundational nature of rules is retained, but the coherence style rules no longer act in the same way, unless some of the consequences of the finite base are added to the base. In general, this is as difficult as full theory closure, which is semi-decidable in the first order case, and NP-complete in the propositional case. We presented a method of restricting the rules so that partial closure will restore the operation of coherence-based rules.

One point which remains unclear is where the division between the two theories actually occurs. That is, when choosing between the foundational and coherence theories, is this a property of the justification or the fact which is justified by it?

Current research includes the completion of the AGM implementation, and examining other ways of achieving the coherence style reasoning, such as keeping a record of the history of the belief set, by timestamping the beliefs when they are removed by a contraction or revision operation. Using this method, partial closure of the base can be calculated retrospectively, and therefore only when necessary. We are also examining extending the ATMS to allow coherence based reasoning as well as foundational reasoning.

7. References

1. Pappas, G.S. and Swain M. (1978) *Essays on Knowledge and Justification*, pp. 30-35 and 184-308, Cornell University Press.
2. Harman, G. (1986) *Change In View*, Bradford Books, MIT Press, Cambridge, MA.
3. Gärdenfors, P. (1989, first draft) *The Dynamics of Belief Systems: Foundations vs. Coherence Theories*, Department of Philosophy, Lund University.
4. Sosa, E. (1980) *The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge*, Midwest Studies in Philosophy 5, pp. 3-25.
5. Doyle, J. (1979) *A Truth Maintenance System*, Artificial Intelligence 12, pp. 231-272.
6. de Kleer, J. (1986) *An Assumption-Based TMS*, Artificial Intelligence 28, pp. 127-162.
7. Thagard, P. (1989) *Explanatory Coherence*, Behavioural and Brain Sciences 12, pp. 435-502.
8. Cornman, J.W. (1977) *Foundational versus Nonfoundational Theories of Empirical Justification*, in (1), pp. 229-278.
9. Gärdenfors, P. (1988) *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Bradford Books, MIT Press, Cambridge, MA.
10. Dixon, S. and Foo N. (1992) *Encoding the ATMS in AGM Logic (Revised)*, CS Technical Report 441, University of Sydney.