ENTRENCHMENT KINEMATICS 101†

Abhaya C. Nayak, Norman Y. Foo, Maurice Pagnucco
Knowledge Systems Group, Dept. of Computer Science,
University of Sydney, Madsen Building, F09
Sydney, NSW 2006, Australia

and

Abdul Sattar
School of Computing and Information Technology,
Griffith University, Brisbane, QLD 4111, Australia

ABSTRACT

The AGM framework of belief dynamics fails to account for iterated belief change. This paper extends the AGM framework in order to solve this problem. As in the AGM, an epistemic input is represented as a single sentence. Unlike the AGM, however, a belief state is represented as an epistemic entrenchment relation. Three state transition functions – namely, entrenchment expansion, entrenchment contraction and entrenchment revision – are constructed that preserve the basic appeal of the AGM system.

1. Introduction

The ability to incorporate new information into a belief state while preserving consistency is an important part of intelligent behavior. This problem is faced in such areas as reason maintenance, reasoning about action, belief change in static (or dynamic) worlds and nonmonotonic logics. Many approaches to these problems are influenced by the AGM framework for belief change[1,2]. However, most of them, like the AGM system, fail to provide a reasoned account of iterated belief change. The purpose of this paper is to give an account of iterated belief change that is close to the AGM system in both framework and spirit.

Unlike the AGM system which identifies a belief state with the set of beliefs accepted in that state, we represent a belief state as an epistemic entrenchment relation (denoted by $\preceq$), from which the current beliefs can be extracted. We construct three entrenchment change operations, namely $\oplus$ (expansion), $\ominus$ (contraction) and $\odot$ (revision) that map $\preceq$ to new entrenchment relations given a sentential input. The account is parallel to the AGM system in which there are three belief change operations $+$, $-$ and $*$ that map a theory to a new theory given a sentential epistemic input.

However, the analogy between the AGM account of belief change and ours fails on one count. In the AGM, an exogenous pragmatic factor (selection mechanism), such as epistemic entrenchment, is used in the actual construction of $-$ or $*$; but in the account we offer, no such exogenous variable is hidden. This difference has an important consequence. Since the AGM system cannot update the exogenous factor in question, it is left to the agent to invent a new selection mechanism every time she

† To appear in C. Zhang, J. Debenham and D. Lukose (Eds.), *Proceedings of the Seventh Australian Conference on Artficial Intelligence*, Armidale, Australia, November 1994.

changes her mind. In other words, a principled account of iterated belief change is lacking in the AGM system. We make up for this inadequacy in the AGM system by offering a reasoned account of the change in the selection mechanism itself.

Figure 1 helps in visualizing the basic picture of the account that we offer. Epistemic states (entrenchment relations) are represented by ovals while the solid ar-

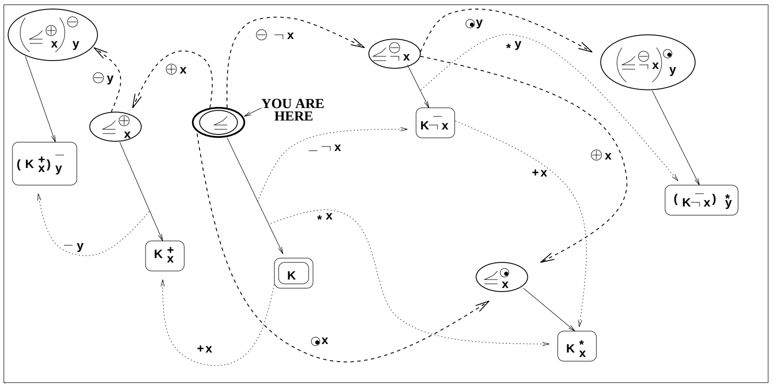

Figure 1: Belief change via change in entrenchment

rows proceeding from these ovals to rectangles (theories) signify the propositions that are accepted in those states. The arrows in broken lines proceeding from these states to other states represent various state transition functions (entrenchment change operations). The dotted arrows stand for the AGM theory change operations. (Note for future reference that not all of them proceed from theories; some of them use both a theory and an entrenchment relation as arguments.) Of particular interest is the way the theory K^*_x is accessed by various routes.

2. The AGM Paradigm

In the AGM approach, the object language (the language in which the beliefs of an agent are represented) is a propositional language $\mathcal{L}$ closed under the usual connectives $\neg$, $\rightarrow$, $\leftrightarrow$, $\wedge$ and $\vee$. Two eminent members of $\mathcal{L}$ are $\top$ (Truth) and $\bot$ (Falsity). The background logic (the logic of the agent) is represented by its consequence operation Cn which satisfies the following conditions: For all sets of sentences Γ, Γ' and all sentences x and y,

Inclusion:	$\Gamma \subseteq Cn(\Gamma)$
Iteration:	$Cn(Cn(\Gamma)) = Cn(\Gamma)$
Monotonicity:	$Cn(\Gamma) \subseteq Cn(\Gamma')$ whenever $\Gamma \subseteq \Gamma'$
Supraclassicality:	$x \in Cn(\Gamma)$ if Γ classically implies x
Deduction:	$y \in Cn(\Gamma \cup \{x\})$ iff $(x \rightarrow y) \in Cn(\Gamma)$
Compactness:	If $x \in Cn(\Gamma)$ then $x \in Cn(\Gamma')$ for some finite $\Gamma' \subseteq \Gamma$.

We often write $\Gamma \vdash x$ for $x \in Cn(\Gamma)$.

Any set of sentences K (possibly with subscripts or superscripts) closed under Cn is called a belief set (or theory) and represents a possible belief state. All members of K are the sentences that the agent has as beliefs. By $\mathbf{K}$ we denote the set of all possible belief sets (in $\mathcal{L}$). Thus $\mathbf{K}$ may be interpreted as the conceptual framework of the agent. We use $K_\perp$ to denote the absurd belief set $\mathcal{L}$.

At least three forms of epistemic change are recognized in the AGM framework. These changes are represented by three operations $+, -, * : \mathbf{K} \times \mathcal{L} \rightarrow \mathbf{K}$ that, for every belief set, returns a belief set given a sentential epistemic input. Since members of $\mathbf{K}$ are interpreted as belief states, these operations may be viewed as state transition functions. The operation $+$ (expansion) is the simplest in the sense that K_x^+ is simply defined to be $Cn(K \cup \{x\})$. However this operation is inconsistency injecting in so far as x is inconsistent with K. This shortcoming of the operation $+$ is overcome by the operation $*$ (revision). Revision of a theory K by x is seen as the expansion by x of a theory K' which in turn is the outcome of retracting $\neg x$ from K (Levi Identity). Belief retraction is carried out by an operation, $-$, called the belief contraction operation. Thus the Levi Identity can be simply stated as: $K_x^* = (K_{\neg x}^-)_x^+$.

In the AGM framework, a set of eight revision (contraction) postulates are put forth as putative constraints that a revision (contraction) operation should satisfy. For the sake of brevity we refrain from reproducing or discussing them here; but they can be found, for instance, in Gärdenfors[2] (§3.3, §3.4). We call any theory revision (contraction) operation satisfying these eight revision (contraction) postulates an "AGM Rational" theory revision (contraction) operation.

It is well known that as long as the contraction operation, –, is AGM rational, the revision operation $*$ defined from it via the Levi Identity will also be AGM rational. Furthermore, for every AGM rational theory revision operation $*$, there exists an AGM rational theory contraction operation $-$ from which it can be derived by the Levi Identity. Unfortunately, given a fixed theory K, there are a number of AGM rational revision (contraction) operations associated with it. Obviously, some selection mechanism is required to choose one rather than the other of these operations. An epistemic entrenchment relation $\preceq$ is such a mechanism.

An epistemic entrenchment relation $\preceq$ is defined over the members of $\mathcal{L}$, and is relative to the belief set K. The following are the *standard* conditions that the relation $\preceq$ is required to satisfy: where x and y are two arbitrary sentences of $\mathcal{L}$,

$(SEE1)$	If $x \preceq y$ and $y \preceq z$ then $x \preceq z$	(transitivity)
$(SEE2)$	If $x \vdash y$ then $x \preceq y$	(dominance)
$(SEE3)$	For any x and y, $x \preceq x \wedge y$ or $y \preceq x \wedge y$	(conjunctiveness)
$(SEE4)$	When $K \neq K_\perp$, $x \notin K$ iff $x \preceq y$ for all y	(minimality)
$(SEE5)$	If $y \preceq x$ for all y, then $\vdash x$	(maximality)

The sentence $x \preceq y$ may be interpreted as saying that it is not more difficult to give up x than it is to give up y (or, x is at least as corrigible as y). The motivation behind the constraints (SEE1)-(SEE5) can be found in Gärdenfors[2] (§4.6).

The primary function of $\preceq$ is to construct the theory contraction operation $-_\preceq$:

(SEE–): $y \in K_x^{-\preceq}$ iff $y \in K$ and either $x \prec (x \vee y)$ or $\vdash x$

where $\prec$ is the strict part of the SEE relation $\preceq$. Coupled with the Levi Identity, this leads to the following revision operation $*_\preceq$:

(SEE$*$): $y \in K_x^{*\preceq}$ iff either $(x \rightarrow \neg y) \prec (x \rightarrow y)$ or $\vdash \neg x$

Gärdenfors and Makinson[3] have shown that if $\preceq$ satisfies the constraints (SEE1)–(SEE5), then the corresponding contraction and revision operations are AGM-rational, and that every relation satisfying (SEE1)-(SEE5) can be defined from some AGM rational contraction (revision) operation (also see Gärdenfors[2]).

3. Why Entrenchment Kinematics

In the AGM system, a belief state is represented as a theory. However, as we noted in the last section, in the construction of a unique theory revision (contraction) operation, the AGM relies on the associated entrenchment relation which is not part of the AGM belief state at all. An unfortunate consequence of this is that a decent account of iterated theory change is not available in the AGM system.

Consider, for instance, a theory K, its associated entrenchment relation $\preceq$, and two sequential epistemic inputs x and y. Our task is to compute the resultant belief state $(K^*_x)^*_y$. It is easy to see how the intermediate state K^*_x can be computed. But we cannot compute $(K^*_x)^*_y$ unless we obtain an entrenchment relation associated with K^*_x. Since, $K \neq K^*_x$ in general, it follows from (SEE4) that the appropriate entrenchment relation associated with K^*_x is likely to be different from $\preceq$. We can run parallel arguments to show that the result of certain other sequential belief changes, such as $(K^+_x)^-_y$ or $(K^-_x)^*_y$, cannot be computed within the AGM framework either.

The implication of the above discussion is the following: an account of merely how a belief corpus is to be modified is not sufficient; a modest goal of any belief change program should also include a recipe for updating the selection mechanism that accompanies the belief corpus. However, since K can be extracted from its associated entrenchment relation $\preceq$ (the principal case is taken care of by SEE4; in the trivial case $K = K_\perp$), it is preferable to represent a belief state as an entrenchment relation rather than a theory. If so, then the state transition functions $+, -$ and $*$ now map belief states (entrenchment relations) to theories which are not belief states at all. Thus they do not obey the principle of category matching which has been recognized as a cardinal principle of belief dynamics[4].

In light of the above discussion, the goal of a belief change program can be further simplified: a sufficient goal of a belief change program is to give a recipe of how to modify a selection mechanism, in particular, an epistemic entrenchment relation, to a new one, given an input. Having recognized this goal of belief change, Hans Rott[7] entertained the following proposal for SEE *revision*:

$$x \preceq_p y \text{ iff } p \rightarrow x \preceq p \rightarrow y$$

where $\preceq_p$ is the result of revising $\preceq$ in light of p. But he found this definition wanting because of its anomalous behavior. However, the fact that our account of entrenchment *revision* is a slight generalization of Rott's aborted proposal (see definition 3(c)) shows that Rott was actually on the right track.

4. Entrenchment Kinematics

In the AGM, an entrenchment relation is seen extraneous to the belief state. But we consider the entrenchment relation to be a belief state in itself, and the set of current beliefs (belief set, theory) as, so to say, its epistemic content. Accordingly, reference to K in the definition of $\preceq$ (as in (SEE4)) is rather inappropriate. Furthermore, according to (SEE5), maximally entrenched beliefs are invariably taken to be logical truths. It seems that this is an inappropriate constraint, especially when the agent in question holds inconsistent beliefs. If an agent does not mind courting inconsistency, there is no reason to assume that such an agent's maximally entrenched beliefs would be logical truths. Accordingly we slightly modify the definition of an epistemic entrenchment relation. From here onwards, by an EE relation we refer to any relation $\preceq$ such that for any arbitrary sentences x, y, z of $\mathcal{L}$,

$(EE1)$	If $x \preceq y$ and $y \preceq z$ then $x \preceq z$	(transitivity)
$(EE2)$	If $x \vdash y$ then $x \preceq y$	(dominance)
$(EE3)$	For any x and y, $x \preceq x \wedge y$ or $y \preceq x \wedge y$	(conjunctiveness)
$(EE4)$	Given that $\perp \prec z$ for some z, if $y \preceq x$ for all y, then $\vdash x$	(maximality).

Competing proposals toward the weakening of SEE relations can be found in Rott[4,8].

We will denote by $\mathcal{R}$ the set of all possible EE relations over the language $\mathcal{L}$. Thus, whereas in the AGM system the conceptual framework is represented by $\mathbf{K}$, in our framework it is represented by $\mathcal{R}$. Two eminent members of $\mathcal{R}$ are $\preceq_\perp$ and $\preceq_\top$. The former is obviously the absurd EE relation: $x \preceq_\perp y$ for all x and y. The latter is the one associated with the empty theory $Cn(\emptyset)$, defined as: $x \preceq_\top y$ iff either $\not\vdash x$ or $\vdash y$. Thus, while $\preceq_\perp$ is a flat entrenchment ordering, $\preceq_\top$ is a two-level entrenchment ordering that assigns the minimal entrenchment to all non-tautologies and the maximal entrenchment to all the tautologies. By the epistemic content $EC(\preceq)$ of any EE relation $\preceq$ we refer to the theory that is extracted from $\preceq$ as defined below:

Definition 1 *(Epistemic Content)*

$$EC(\preceq) = \begin{cases} \{x \mid \perp \prec x\} & \textit{if } \perp \textit{ is not } \preceq\textit{-maximal} \\ K_\perp & \textit{otherwise} \end{cases}$$

is the set of sentences considered accepted in the belief state $\preceq$.

In accordance with this modification in the definition of an epistemic entrenchment relation, we modify the constructions **SEE**$-$ and and **SEE**$*$ to **EE**$-$ and and **EE**$*$ respectively:

(EE$-$): $y \in K_x^{-\preceq}$ iff either $\vdash y$ or both $y \in K$ and either $x \prec (x \vee y)$ or $\vdash x$

(EE$*$): $y \in K_x^{*\preceq}$ iff either $(x \rightarrow \neg y) \prec (x \rightarrow y)$ or $\vdash \neg x$ or $x \vdash y$.

Our goal is to define three entrenchment change operations, $\oplus, \ominus$ and $\odot$: $\mathcal{R} \times \mathcal{L} \rightarrow \mathcal{R}$ that are rational in an appropriate sense. Given an EE relation $\preceq$ and a sentence e, we denote by $\preceq_e^\oplus, \preceq_e^\ominus$ and $\preceq_e^\odot$ the EE relations that are returned by $\oplus, \ominus$ and $\odot$ respectively. We will use the following terminology:

Definition 2 *(Rational Entrenchment Change)*

- *An EE expansion operation $\oplus$ is AGM rational iff $EC(\preceq_e^\oplus) = (EC(\preceq))_e^+$*
- *An EE contraction operation $\ominus$ is AGM rational iff $EC(\preceq_e^\ominus) = (EC(\preceq))_e^{-\preceq}$*
- *An EE revision operation $\odot$ is AGM rational iff $EC(\preceq_e^\odot) = (EC(\preceq))_e^{*\preceq}$*
- *The triple $\langle \oplus, \ominus, \odot \rangle$ is Levi connected iff $\preceq_x^\odot = (\preceq_{\neg x}^\ominus)_x^\oplus$.*

It is reasonable to require the entrenchment change operations to be AGM rational in the above sense, and that the Levi connection holds of them. We offer below constructions of such operations keeping in view the desirability of these properties.

Definition 3 *(Entrenchment Revision)*
$x \preceq_e^{\odot} y$ iff any of the following four conditions is satisfied:
(a) $e \vdash \bot$, (b) $e \vdash x$, $e \vdash y$, AND either $\vdash y$ or both $\not\vdash x$ and $x \preceq y$, (c) $e \not\vdash x$, $e \not\vdash y$, and $(e \to x) \preceq (e \to y)$, (d) $e \not\vdash x$ but $e \vdash y$

According to this definition, (a) when the input is contradictory, the resulting belief state is an absurd state. When the input is not contradictory, the following takes place. (b) The relative priority among the consequences of the input is not affected if $\preceq$ is not absurd. (c) For the interpretation of this condition, we need to realize what sentences of the form $\alpha \to \beta$ stand for. The sentence $e \to x$ may be read as the relative information in x given the input e. (Similarly $e \to y$). Given this reading, according to this clause, priority among non-consequences of the input is set to the (original) relative priority between the information they contain relative to the input. It also shows that Rott's proposed definition is a special case of this definition. (d) Consequences of the input are accorded higher priority than the non-consequences of the input. Admittedly this condition is somewhat contentious. However, we might motivate it by pointing out that in the AGM system the input is potentially accorded the highest priority among the contingent sentences since it can override any of the old beliefs.

Theorem 1 (a) *Given that $\preceq$ is an EE relation and e is a sentence, the relation $\preceq_e^{\odot}$ is an EE relation.* (b) *The entrenchment revision operation $\odot$ is AGM rational.*

The first part of this result shows that the operation $\odot$ satisfies the principle of category matching. The second part shows that theory revision performed through this type of entrenchment revision is as good as that done in the AGM style. We show below that entrenchment contraction and entrenchment expansion, when suitably defined, exhibit similar behavior.

Definition 4 *(Entrenchment Contraction)*
$(\preceq_\bot)_e^{\ominus} = \preceq_\top$ if $\not\vdash e$.
$(\preceq)_\top^{\ominus} = \preceq$ for all $\preceq$.
When $\preceq \neq \preceq_\bot$ and $\not\vdash e$, $x \preceq_e^{\ominus} y$ iff any of the following three conditions hold: (a) *$x \preceq \bot$,* (b) *$e \not\prec e \vee x$* (c) *both $e \prec e \vee y$ and $x \preceq y$.*

Interpretation of the two trivial cases is obvious. We give a reading of the principal case. According to the first clause (a), the sentences that are originally minimally entrenched stay minimally entrenched after any contraction. (This is because of the fact that contraction is not the means of acquiring new beliefs.) The second clause (b), read along with condition **C**–, says that the beliefs that are given up because of the contraction become minimally entrenched because of the contraction. The third clause (c) says that the relative priority among the sentences that are not affected by an AGM contraction does not change.

Theorem 2 (a) *Given that $\preceq$ is an EE relation and e is a sentence, the relation $\preceq_e^{\ominus}$ is an EE relation,* (b) *the entrenchment contraction $\ominus$ is AGM rational.*

Finally, we construct entrenchment expansion in the following manner:

Definition 5 *(Entrenchment Expansion)*

$$\preceq_e^{\oplus} = \begin{cases} \preceq_{\perp} & \text{if } \neg e \in EC(\preceq) \\ \preceq_e^{\odot} & \text{otherwise} \end{cases}$$

This definition essentially says that expansion by an input that conflicts with the current belief state injects inconsistency into it. Otherwise, expansion behaves exactly like revision. This definition accords well with the AGM intuition about theory expansion.

Theorem 3 (a) *Given that $\preceq$ is an EE relation and e is a sentence, the relation $\preceq_e^{\oplus}$ is an EE relation,* (b) *the entrenchment expansion operation $\oplus$ is AGM rational.*

The following result states the connection among entrenchment expansion, contraction and revision.

Theorem 4 *The triple $\langle \oplus, \ominus, \odot \rangle$ is Levi connected.*

This result shows that the entrenchment expansion operation $\oplus$ forges a suitable link between $\ominus$ and $\odot$ the way the theory expansion operation $+$ connects the theory contraction operation $-$ and the theory revision operation $*$ via the Levi Identity in the AGM framework. This shows that, as theory revision factorizes to theory contraction and theory expansion in the AGM, similarly, entrenchment revision can be factorized into entrenchment contraction and entrenchment expansion in our framework.

5. Summary and Discussion

We will now briefly point out how our proposal fares *vis a vis* its competing proposals found in the literature. A more elaborate discussion of these proposals can be found in Nayak[6]. Hansson[5], Rott[8] and Schlechta[9], in their respective accounts of iterated belief change rely on a super selection mechanism: a mechanism that does not change in the course of normal belief change. This leads to a kind of epistemic determinism that is difficult to defend. Spohn[10] and Williams[11] avoid this pit-fall – but only at the cost of making a rather heavy demand on the part of the agent. The measure they impose on a body of beliefs is ordinal, not merely relational; and for them, the epistemic input is more than a mere sentence. Furthermore, they deny the epistemic operations the relatively autonomous status that they enjoy in the AGM framework. The account of iterated belief change offered by Nayak[6] is concerned only with revision and, for the most part, requires the epistemic input to be represented as an EE relation.

The account that we have offered in this paper has none of the above shortcomings. We impose on the body of beliefs a relational measure as in the AGM system. We do not require the epistemic input to be anything other than a sentence. Yet, we offer a recipe of iterated belief change in all its modalities, *viz.* expansion, contraction and revision. We show that, as in the AGM system, the revision operation $\odot$ can be factorized to expansion $\oplus$ and contraction $\ominus$. Furthermore, all appropriate relations among the EE relations and belief sets hold.

Finally, one may wonder whether the constructions of $\oplus$, $\ominus$ and $\odot$ offered in this paper are unique in the sense that no other triple is AGM rational and Levi connected. We leave this question unanswered. However, we have inconclusive reasons to believe that there are other combinations possible with precisely the same features. Since it is beyond the scope of this work to adequately present or discuss these alternatives, we leave this task for some other occasion.

6. References

1. Alchourrón, C. E., P. Gärdenfors, and D. Makinson, "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *Journal of Symbolic Logic,* **50**, pp.510-530, 1985.
2. Gärdenfors, P. *Knowledge in Flux. Modeling the Dynamics of Epistemic States,* Bradford/MIT, Cambridge (Massachussettes), 1988.
3. Gärdenfors, P. and D.Makinson, "Revision of Knowledge Systems using Epistemic Entrenchment," in *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge,* M. Vardi ed., Los Altos, CA: Morgan Kaufmann. pp. 83-95, 1988.
4. Gärdenfors, P. and H. Rott, "Belief Revision", to be published in *Handbook of Logic in AI and Logic Programming*, vol. IV.
5. Hansson, S. O. "A Dyadic Representation of Belief", in Peter Gärdenfors ed. *Belief Revision*, CUP, pp. 89-121, 1992.
6. Nayak, A. C. "Iterated Belief Change Based on Epistemic Entrenchment", forthcoming, *Erkenntnis*.
7. Rott, H. "Two Methods of Constructing Contractions and Revisions of Knowledge Systems", *Journal of Philosophical Logic,* **20**, pp. 149-173, 1991.
8. Rott, H. "Preferential Belief Change Using Generalized Epistemic Entrenchment", *Journal of Logic, Language and Information*, **1**, 45-78, 1992.
9. Schlechta, K. "Theory Revision and Probability", *Notre Dame Journal of Formal Logic*, **32**, 307-319, 1991.
10. Spohn, W. "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States", in W. L. Harper and B. Skyrms (eds), *Causation in Decision, Belief Change and Statistics,* **2**, pp. 105-134, 1988.
11. Williams, M. A. "Transmutations of Knowledge Systems", in *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, J. Doyle and E. Sandewall eds., (forthcoming).