PRINCIPIA **28**(1): 71–77 (2024) doi: 10.5007/1808-1711.2024.e96706
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

BELIEF REVISION IN A NUTSHELL

RAFAEL R. TESTA
Universidade de Campinas, BRAZIL
rafaeltesta@gmail.com

Abstract. Belief Revision studies how rational agents change their beliefs in response to new information. The main objective of the works in this area is modelling these dynamics by defining some formal operators of change. This paper is an overview on the AGM model for revision, introducing some of the major questions that are addressed in this area of study.

Keywords: belief revision • change operators • epistemological theories • rationality • formal epistemology

RECEIVED: 12/12/2022 REVISED: 23/01/2023 ACCEPTED: 01/05/2023

1. Introduction

Belief Revision (Belief Change) is an interdisciplinary research area that makes use of logic to produce models of how rational agents change their beliefs in response to new information. To better understand it, consider the following set of belief-representing sentences (belief state) in natural language (Gärdenfors 1992): all European swans are white; the bird caught in the trap is a swan; the bird caught in the trap comes from Sweden; and Sweden is part of Europe. From this set it can be deduced that the bird caught in the trap is white. But suppose that the bird caught is black. If this new piece of information is simply added to the latter set, it will become inconsistent. If it is intended that the new belief state remains consistent, how should one behave? Should the old beliefs be preserved by ignoring the new information? Or one should add the new information, removing some of the beliefs contained in the original set? In this case, should all the beliefs be retracted, or just some of them? Which ones?

In a wide sense, those questions have been subject of philosophical reflection since antiquity, as it is highlighted by Fermé and Hansson (2018), including discussions on the mechanisms by which scientific theories develop in the beginnings of the twentieth century. One of the main contributions to the study of belief dynamics is the so-called AGM model, proposed by Alchourrón, Gärdenfors and Makinson (1985) alongside a number of papers by the authors, mostly influenced by the works of Georg Henrik von Wright, Jaako Hintika, William Harper and Isaac Levi.

2. Rationality criteria

The main argument advocated by the AGM trio, mostly advanced by Gärdenfors extending and generalizing previous works on the area, is that there are certain basic principles in the metalevel of epistemological theories that must be satisfied — the so-called rationality criteria (Gärdenfors 1982). For instance, in the previous example the reader should agree that it does not seem rational to retract the fact that Sweden is part of Europe, instead of retracting for example the belief that all European swans are white. This means that there is some kind of preference when choosing in what to believe (and, accordingly, in choosing what to retract). Furthermore, another criterion tacitly present in that example is the preservation of consistency. Indeed, *consistency preservation* and *entrenchment order* are two of the most usual criteria, alongside *informational economy* (or *minimality*, which states that as much as possible of the previous information are retained), *logical closure* (the outcome of a belief change operation is a logically closed set, just like the original one), *irrelevance of syntax* (the outcome of a change does not depend on the formal representation of the beliefs), and *primacy of the new information*, among others.

Let us turn our attention to the so-called revision operation, namely the one that takes a set (the original set of beliefs) and a sentence (the new belief to be incorporated) and combines them into a new consistent belief set, possibly retracting parts of the original one — thus solving the puzzling situation in the example presented in the introduction. Is it possible to explicitly construct such belief change operator? How can we guarantee that such formal construction satisfies the rationality desiderata captured by the aforementioned criteria?

3. AGM model

In the AGM model, beliefs are expressed in a language L, closed under the usual truth functional connectives. Let Cn be a consequence operator on L satisfying the standard Tarskian properties, and $Cn(X)$ be all the logical consequences of a set X of sentences. A belief set representing an epistemic agent's belief state is a set of sentences logically closed such that $K = Cn(K)$.

3.1. Postulates for revision

If a new piece of information does not contradict K, then it can be directly incorporated by the agent — what is considered as a plain expansion, set-theoretically defined as $K + \alpha = Cn(K \cup \{\alpha\})$. However, in most situations (like the aforementioned example) a more complex operation is necessary. A revision of K by a belief-representing sentence α ($K * \alpha$) can be formally defined by the following basic postulates:

(∗1)	$K * \alpha = Cn(K * \alpha)$	(Closure)
(∗2)	$\alpha \in K * \alpha$	(Success)
(∗3)	$K * \alpha \subseteq K + \alpha$	(Inclusion)
(∗4)	If $\neg\alpha \notin K$, then $K + \alpha \subseteq K * \alpha$	(Vacuity)
(∗5)	If α is consistent, then $K * \alpha$ is consistent	(Consistency)
(∗6)	If $\alpha \leftrightarrow \beta$, then $K * \alpha = K * \beta$	(Extensionality)

As expected, those postulates formally captures the rationality desiderata previously advanced in natural language. (∗1) assures that the result of a revision is always a belief set, while (∗2) assures that the new piece of information is in the new belief state. (∗3) and (∗4) are reflexes of informational economy. (∗5) expresses consistency preservation. (∗6) is a formalization of the irrelevance of syntax criterion. It is worth mentioning that there are several alternative axiomatic characterizations of revision, some of them equivalent or even supplementary to those basic postulates — all these different characterization of revision can be found in the references.

3.2. Constructive model for revision

The basic postulates presented in the previous section fully characterize an explicit construction for revision. The one that has been chosen to be explored in this paper is the so-called *partial meet revision*. This construction is based in the concept of *remainder set*, that is a set of maximal subsets of a given set that fail to imply a given sentence.

Definition 3.2.1 (Remainder Set) Let A be a set of sentences and α a sentence. The set $A\perp\alpha$ ("remainder of A by α" or simply "A remainder α") is the set of sets such that $B \in A\perp\alpha$ if and only if:

- $B \subseteq A$
- $B \nvdash \alpha$
- For all B' such that $B \subset B' \subseteq A, B' \vdash \alpha$

The elements of $A\perp\alpha$ are called the remainders of A by α. The partial meet revision of K by α is then obtained by expanding by α the intersection of some of the remainders of K by $\neg\alpha$, that is, expanding by α some subset of K that does not contradict with α. The choice of those elements is performed by a *selection function*. Intuitively, the selection function is expected to choose those beliefs that are more entrenched in the agent's belief state.

Definition 3.2.2 (Selection Function) Let A be a set of sentences. A selection function for A is a function γ such that, for all sentences α:

- If $A \perp \alpha \neq \emptyset$ then $\gamma(A \perp \alpha)$ is a non-empty subset of $A \perp \alpha$.
- If $A \perp \alpha = \emptyset$ then $\gamma(A \perp \alpha) = \{A\}$

Definition 3.2.3 (Partial Meet Revision) Let K be a belief set and γ a selection function for K. For any sentence α, the partial meet revision operator over K determined by γ is defined as follows:

$$K *_\gamma \alpha = Cn(\bigcap \gamma(K \perp \neg\alpha) \cup \{\alpha\})$$

An operator $*$ on K is a partial meet revision if and only if there is a selection function γ for K such that for all sentences $\alpha : K * \alpha = K *_\gamma \alpha$.

It is worth mentioning that when revising K by α, the previous retraction of those elements that contradicts the sentence α is indeed the so-called contraction operation — accordingly, a partial meet contraction of K by α is constructed as $K -_\gamma \alpha = Cn(\bigcap \gamma(K \perp \alpha))$ (given a selection function γ). Thus, contraction and revision are inter-definable: $K * \alpha =_{def} (K - \neg\alpha) + \alpha$ (Levi identity); and $K - \alpha =_{def} K \cap (K * \neg\alpha)$ (Harper identity).

3.3. From postulates to construction, and vice-versa

The class of partial meet revision functions coincides exactly with the class of revisions defined by the aforementioned basic postulates — a result called Representation Theorem.

Observation 3.3.1 (Representation Theorem) Let K be a belief set. An operator on K is a partial meet revision function if and only if it satisfies the postulates $(*1)$ to $(*6)$.

The proof of this theorem has two parts. In the construction-to-postulates part, it is shown that the partial meet revision satisfies the given postulates — what follows almost directly by construction and the properties of remainder sets. In the postulates-to-construction part, it is shown that if an operation satisfies the given postulates, then that operation can be constructed as a partial meet revision. The very Harper identity can be of use: let $*$ be an operator for K that satisfies the listed postulates and let γ be defined as follows:

$$\begin{aligned} \gamma(K \perp \alpha) &= \{X \in K \perp \alpha : K \cap K * \neg\alpha \subseteq X\} \text{ if } K \perp \alpha \neq \emptyset \\ &= \{K\} \text{ otherwise.} \end{aligned}$$

Along these lines, the proof follows by proving that (1)γ is a selection function and (2) that $K *_\gamma \alpha = K * \alpha$ for all α. The details of this proof can be found in the references.

4. Equivalent characterizations and alternative models

Since the publication of the AGM model in 1985, its major constructions and concepts have been the starting point to significant ramifications as well as to criticisms, including the elaboration of alternative models that have been proposed to extend or replace the AGM framework. Some of the reasons for this impact are the main features of the latter model — its simplicity, generality and capability of being expressed in several different ways, including *possible world models, epistemic entrenchment, specified meet contraction, kernel contraction and safe contraction* (Fermé and Hansson 2018).

Other reason for the lasting influence of the AGM model is that it provided a simple input–output framework for modelling change, with potential applications to a wide range of important areas and connections with other fields of research, for example *Nonmonotonic* and *Defeasible Reasoning; Argumentation, Social Choice* and *Game Theories; Truth Maintenance Systems; Database Update*; just to name a few.

Notwithstanding, just like any other logical formalization of philosophical arguments, the AGM model of belief revision describes an idealized scenario. The epistemic agent's rationality presupposes a certain logical omniscience (as it is required by *closure*) and always prioritizes the new information in prejudice of old beliefs (by *success*), among other compromises. Thus, alternative models are considered when distinct contexts are assumed (including the aforementioned connections).

Some of these models are, for instance, *Belief Bases*, in which belief states are not closed, thus differentiating beliefs accepted by the agent and their logical consequences, among several other features (Hansson 1999a), *Paraconsistent Belief Revision*, in which consistency desideratum can be refined into at least two distinct ones, namely non-contradictoriness and non-triviality (Testa et. al. 2018), *Non-Prioritized Belief Change*, in which it is possible to reject or to accept only parts of the new information (Hansson 1999b), *Multiple Change* (Fuhrmann and Hansson 1994), among many others — a comprehensive overview is given by Fermé and Hansson 2018.

In fact, not all rationality principles are desirable at all times, and certain criteria contradict with others when distinct cases are considered, as it is explained by Testa (2015) in the context of *Paraconsitent Belief Change* (as studied by Testa et. al. 2017). In the latter model, it is highlighted the fact that there is a kind of trade-off between consistency and minimality — contradictions are not sufficient conditions for retracting a belief when minimality principle is held in higher regard, given the possibility of contradictory yet non-trivial theories in a paraconsistent setting (Testa 2023).

Indeed, there are several trade-offs when advancing representational formalisms aimed at various uses or applications, mostly related with two important concepts: the *expressiveness* of the representation language and the *tractability* of the associated reasoning task (Brachman and Levesque 2004). Accordingly, the more fine-grained a formalization is (with a more expressive language), the more complex it will also be (and vice versa). This fact can be perceived in most formalizations of philosophical arguments present in this special issue.

5. Final Remarks

The introduction of formal models of belief change has opened up a new field of philosophical reflection, along with discussions on the very nature of epistemic states and the underlaying rationality of its dynamics. Furthermore, new light has been shed on some classical questions, for instance the use of the formal tools and methods of belief change for a number of investigations in the philosophy of science (Olsson and Enqvist 2011).

Other important impact is fomenting the confluence of Philosophy with several areas of study, including Artificial Intelligence — mostly in the field of *Knowledge Representation and Reasoning*. Indeed, several studies in Computer Science contributed, and still contribute, to the flourishing of the Belief Revision area, such as the field of *Database Update*.

The fact is that the area turned out to be a useful discipline in several distinct fields, including *Economics, Formal Epistemology, Ontology, Legal Theory*, among many others. Moreover, its formal tools and methods allow intersecting the latter disciplines, thus fomenting a cross-fertilization of ideas and results.

References

Alchourrón, C. E.; Gärdenfors, P.; Makinson, D. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* **50**(2): 510–30.

Brachman, R. J.; Levesque, H. J. 2004. *Knowledge Representation and Reasoning*. Morgan Kaufmann.

Fermé, E.; Hansson, S. O. 2018. *Belief Change: Introduction and Overview*. Springer Verlag.

Fuhrmann, A.; Hansson, S. O. 1994. A survey of multiple contractions. *Journal of Logic, Language and Information* **3**(1): 39–75.

Gärdenfors, P. 1982. Rules for rational changes of belief. In: T. Pauli (ed.) *Philosophical Essays Dedicated to Lennart Åqvist on his Fiftieth Birthday*, n.34, pp.88–101.

Gärdenfors, P. 1992. Belief revision: An introduction. In: P. Gärdenfors (ed.) *Belief Revision*, no. 29, pp.1–28. Cambridge University Press.

Hansson, S. O. 1999a. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Dordrecht and Boston: Kluwer Academic Publishers.

Hansson, S. O. 1999b. A survey of non-prioritized belief revision. *Erkenntnis* **50**(2-3): 413–27.

Olsson, E.; Enqvist, S. 2011. Belief Revision meets Philosophy of Science. *Logic, Epistemology, and the Unity of Science*, volume 21. 10.1007/978-90-481-9609-8.

Testa, R. 2015. The cost of consistency: information economy in Paraconsistent Belief Revision. *South American Journal of Logic* **1**(2): 461–80.
Testa, R. 2023. Paraconsistency. In: J. Mattingly (ed.) *The SAGE encyclopedia of theory in science, technology, engineering, and mathematics*, Vol. 1, pp.629–32.
Testa, R.; Coniglio, M. E.; Ribeiro, M. M. 2017. AGM-Like Paraconsistent Belief Change. *Logic Journal of the IGPL* **25**(4): 632–72.
Testa, R.; Fermé, E.; Garapa, M.; Reis, M. 2018. How to construct Remainder Sets for Paraconsistent Revisions: Preliminary Report. *Proceedings of the 17th International Workshop on Non-monotonic Reasoning*, pp.125–31.

Acknowledgments

Rafael R. Testa is funded by a grant from *Carlos Chagas Filho Research Support Foundation* of the State of Rio de Janeiro – FAPERJ.