Supposing versus learning: belief expression versus belief change

Giacomo Bonanno
University of California

Abstract

Consider two possible scenarios for belief "revision". Initially the agent believes that ϕ is not the case, that is, believes $\neg\phi$. In one scenario she receives reliable information that, as a matter of fact, ϕ *is* the case; call this scenario "learning that ϕ". In the other scenario she reasons about what she believes would be the case if ϕ were the case; call this scenario "supposing that ϕ". We argue that there are important differences between the two scenarios. Drawing on the analysis of [3, 4] we show that it is possible to view the AGM theory of belief revision ([1]) as a theory of hypothetical, or suppositional, reasoning, rather than a theory of actual belief change in response to new information.

1 Introduction

It is a pleasure to contribute to the Festschrift in honor of Andreas Herzig on the occasion of his 65th birthday. Andreas has produced an impressive number of contributions covering a wide spectrum of topics. Among the topics thoroughly studied by Andreas are belief revision, conditionals and modal logic. The contribution in this chapter touches upon these three topics.

I am grateful to two anonymous reviewers for helpful and constructive comments.

We consider two possible scenarios for belief "revision". Initially the agent believes that ϕ is not the case, that is, believes $\neg\phi$.[1] In one scenario she receives – from a reliable source – information that, as a matter of fact, ϕ *is* the case; call this scenario "learning that ϕ". In the other scenario she reasons about what she believes would be the case if ϕ were the case (counterfactual reasoning); call this scenario "supposing that ϕ". We argue that there are important differences between the two scenarios. Furthermore, drawing on the analysis of [3, 4], we show that it is possible to view the AGM theory of belief revision ([1]) as a theory of hypothetical, or suppositional, reasoning, rather than a theory of actual belief change in response to new information.

Several authors have pointed out that there is a significant difference between supposing that ϕ and learning that ϕ:

> "Merely suppositional change is essentially different from "genuine" change due to new information." [18, p. 410]

> "Supposing is like pretending, or making believe, in that suppositions do not call for justification in the way that beliefs do. We make them for the sake of argument. "[16, p.540]

> "There seems to be a need to distinguish actual belief revision from belief revision that is merely hypothetical. [...] Ordinary theories of belief change do not seem suited to handle the sort of hypothetical belief change that goes on, for example, in debates where the participants agree, "for the sake of argument", on a certain common ground on which possibilities can be explored and disagreements can be aired. *One need not actually believe what one accepts in this way.*" [19, p. 1, emphasis added]

> "In none of these contexts is supposing that ϕ equivalent to believing that ϕ... *Changing full beliefs calls for some sort of accounting or justification. Supposition does not...*" [15,

[1] We emphasize the scenario where, initially, the agent believes $\neg\phi$, because this is the case where the difference between supposing and learning seems to be more significant. However, the following analysis applies also to the case where, initially, the agent suspends judgment on ϕ (that is, considers both ϕ and $\neg\phi$ possible), as well as the case where the agent starts off believing ϕ.

p.5, emphasis added]

There is also empirical evidence that, even in the case where what is being supposed or learned is compatible with the initial beliefs, people treat supposition and information differently: Zhao *et al* found that there are

> "substantial differences between the conditional probability of an event A supposing an event B, compared to the probability of A after having learned B. Specifically, supposing B appears to have less impact on the credibility of A than learning that B is true." [21, p.373]

As an illustration of the difference between supposing and learning, consider the following example.

> My friend Bob has been complaining for years about his current house: it is too far from his workplace, it is too small, it is in a noisy neighborhood, etc. I accompanied Bob to view three houses: A, B and C, which differ, as shown in Figure 1, on the basis of two attributes: distance from workplace and size (H is Bob's current house). Later I hear that Bob made an offer on one of those three houses. Since, in the past, Bob mainly complained about the long commute, I believe that distance from the workplace was the main attribute in Bob's mind and thus I believe that Bob made an offer on house A. If asked to reason on the **supposition**, *for the sake of argument*, that Bob did not make an offer on house A, it would be defensible for me to maintain my belief that Bob's main concern was distance to the workplace and thus *believe that he made an offer on house B.* On the other hand, if I were to be reliably **informed** that Bob did not make an offer on house A, then I could react by abandoning my belief that distance was the dominant attribute in Bob's mind: size might also have been an important factor and, on the basis of these considerations, I could *believe that he made an offer on house C.*

What is the crucial difference between learning ϕ and supposing ϕ? In our view, it can be found in Levi's observation quoted above ([15,

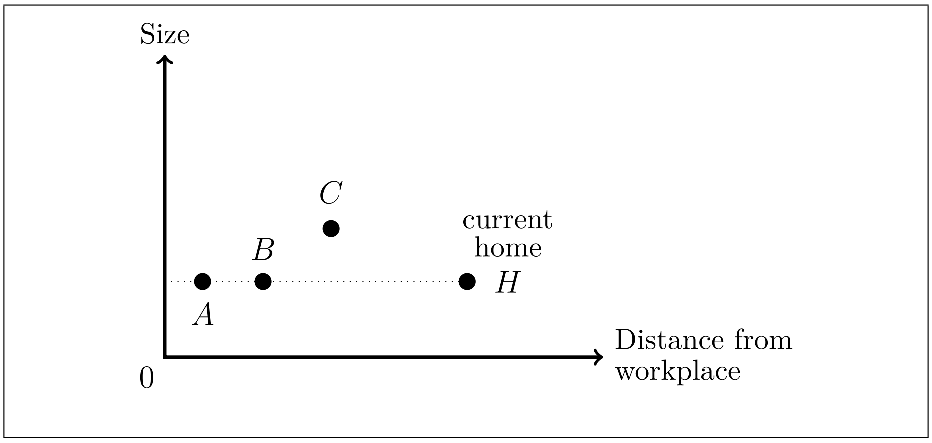

Figure 1: Four houses represented in terms of two attributes: distance from workplace and size. H is Bob's current house.

p. 7]) that *a supposition requires no explanation.* Indeed, it is common to state that a supposition is entertained "*for the sake of argument*"; similar expressions are: "suppose that, f*or whatever reason*, ϕ is the case ..." or "suppose that, *somehow*, ϕ were the case . . . ". In the example illustrated in Figure 1, if I suppose that my friend Bob did not make an offer on house A, I am not required to come up with an explanation for why that would be the case; indeed, there may be a number of plausible circumstances under which it could happen: perhaps somebody else made an offer before Bob, or perhaps the house was withdrawn from the market, etc. Since *no background scenario needs to be provided to account for the supposition*, I am fully justified in maintaining – to the extend possible – my initial beliefs, in particular, that distance to the workplace was still the dominant concern in Bob's mind; hence my suppositional belief that he made an offer on the next closest house, namely house B. On the other hand, *information that contradicts my initial beliefs calls for an understanding of, or explanation for, why my initial beliefs were wrong.*

As further illustration of the difference between supposing and learning, consider the following example discussed by Stalnaker:

> I initially believe the following three things: first, General Smith is a shrewd judge of character – he knows (better than

> I) who is brave and who is not. Second, the general sends only brave men into battle. Third, Private Jones is cowardly. It follows from these three propositions that Jones will not be sent into battle, so I also initially believe that. Let us assume that someone is cowardly if he would run away under fire. So I believe that Private Jones would run away if he were to be sent into battle (which, for that reason, he won't be). [20, p. 46]

Thus, I initially believe that "Private Jones will not be sent into battle" and also that "Private Jones would run away if he were to be sent into battle"; in light of the previous discussion, the latter belief can be interpreted as a suppositional belief: on the supposition that Private Jones is sent into battle I believe that he would run away. The supposition allows me to maintain my belief that he is cowardly, and running away is what a cowardly person would do. Stalnaker interprets this as a

> causal counterfactual – a belief about how Jones is disposed to behave – how he would behave in circumstances that I believe will not in fact rise. [20, ibid.]

But how would I react to the *information* that Private Jones was in fact sent into battle? In this case, Stalnaker suggests that

> since I think the general is a better judge of character than I, I would revise [my initial beliefs] by giving up my belief that Jones is cowardly. Of the three beliefs mentioned above, the first two are more robust than the third. [20, ibid.]

Stalnaker interprets the above example as an illustration of the difference between causal counterfactuals and epistemic counterfactuals. He views the 'if' in the expression "if Private Jones were sent into battle, he would run away" as a *causal* 'if', that is, an 'if' used to express Private Jones' disposition to act in a situation that I believe will not arise. On the other hand, the 'if' in the expression "if Private Jones *is* sent into battle, then he won't run away" is an *epistemic* 'if': it concerns my belief revision policy, in particular, how I would revise my beliefs if I were to *learn* that Private Jones was in fact sent into battle.

We propose a different interpretation: one that pertains to the difference between supposition and learning. Both 'ifs' in the above example are epistemic 'ifs' in that they relate to my beliefs. The statement "I believe that if Private Jones were sent into battle, he would run away" represents suppositional reasoning, in which I *express*, or elucidate, my initial beliefs, reiterating my belief that he is cowardly. The statement "if Private Jones *is* sent into battle, then I believe that he won't run away" represents my reaction to learning that Private Jones was in fact sent into battle, and my reaction is to *change* my beliefs to accommodate, and account for, what I just learned.

In the next section we argue that the AGM theory of belief revision can be viewed as a theory of hypothetical, or suppositional, reasoning, rather than a theory of actual belief change in response to new information.

2 The AGM theory of belief revision

Alchourrón, Gärdenfors and Makinson in [1] propose eight axioms for belief revision, which are listed in the Appendix. Their approach is syntactic. They take as starting point a consistent and deductively closed set of Boolean formulas K representing the agent's initial beliefs. A belief revision function based on K is a function that associates with every Boolean formula ϕ a set of Boolean formulas $K * \phi$, interpreted as the change in K prompted by the input ϕ.

The dominant interpretation of the AGM theory is in terms of *belief change in response to reliable information*, so that $K * \phi$ is understood as the modified belief set after the information represented by the formula ϕ has been made compatible with the initial belief set K. This interpretation is apparent in the way in which the Success Axiom (AGM axiom $(K * 2)$: $\phi \in K * \phi$) is described or criticized in the literature:[2]

> "The Success postulate says that the new information ϕ should always be included in the new belief set. [It] places enormous faith on the reliability of ϕ. The new information

[2] To address these criticisms, a more recent literature ([2, 5, 9, 12, 13]) has dropped the Success Axiom by allowing the agent to discard some pieces of information as not credible or to accept the information only in a limited way.

> is perceived to be so reliable that it prevails over all previous conflicting beliefs, no matter what these beliefs might be." [17, p. 319]
>
> "In AGM revision, new information has primacy. This is mirrored in the Success postulate for revision. At each stage the system has total trust in the input information, and previous beliefs are discarded whenever that is needed to consistently incorporate the new information. This is an unrealistic feature since in real life, cognitive agents sometimes do not accept the new information that they receive."[8, p. 65]
>
> "A system obeying [the Success axiom] is totally trusting at each stage about the input information; it is willing to give up whatever elements of the background theory must be abandoned to render it consistent with the new information. Once this information has been incorporated, however, it is at once as susceptible to revision as anything else in the current theory. Such a rule of revision seems to place an inordinate value on novelty, and its behaviour towards what it learns seems capricious." [6, p. 251]

On the hand – as shown below – when interpreted in the context of suppositional reasoning, the Success Axiom becomes entirely trivial and devoid of any substantive content.

The standard semantics for AGM belief revision ([11, 14, 7]) is in terms of a total pre-order $\succsim$ on the set W of "possible worlds" (that is, maximally consistent sets of formulas) with the interpretation of $w_1 \succsim w_2$ as "world w_1 is at least as plausible as world w_2". In such a model, the agent initially believes ψ (that is, $\psi \in K$) if and only if ψ is true at all the most plausible worlds in W, and the agent believes ψ in response to input ϕ (that is $\psi \in K * \phi$) if and only if ψ is true at all the most plausible ϕ-worlds (a world w is a ϕ-world if ϕ is true there, that is, if $\phi \in w$).

In [3] an alternative semantics for AGM belief revision is put forward, based on Kripke-Lewis frames.

Definition 1. *A* Kripke-Lewis frame *is a triple* $\langle S, \mathcal{B}, f \rangle$ *where*

1. *S is a set of* states*; subsets of S are called* events.

2. *$\mathcal{B} \subseteq S \times S$ is a binary* belief relation *on S which is serial: $\forall s \in S, \exists s' \in S$, such that $s\mathcal{B}s'$ ($s\mathcal{B}s'$ is an alternative notation for $(s,s') \in \mathcal{B}$). We denote by $\mathcal{B}(s)$ the set of states that are reachable from s by $\mathcal{B}$: $\mathcal{B}(s) = \{s' \in S : s\mathcal{B}s'\}$. $\mathcal{B}(s)$ is interpreted as the set of states that initially the agent considers doxastically possible at state s.*

3. *$f : S \times (2^S \setminus \varnothing) \to 2^S$ is a* selection function *that associates with every state-event pair (s,E) (with $E \neq \varnothing$) a set of states $f(s,E) \subseteq S$. $f(s,E)$ is interpreted as the set of states that are closest (or most similar) to s, conditional on event E.*

Under this semantics – interpreted by adding a valuation – the agent initially believes ψ at state s if and only if $\mathcal{B}(s) \subseteq \|\psi\|$ (where $\|\psi\|$ denotes the set of states where ψ is true), and the agent believes ψ in response to input ϕ at state s if and only if, for all $s' \in \mathcal{B}(s)$, $f(s', \|\phi\|) \subseteq \|\psi\|$. Given the customary interpretation of selection functions in terms of conditionals, the latter condition can be interpreted as stating that, at state s, $\psi \in K * \phi$ if and only if the agent believes that "if ϕ is (were) the case then ψ is (would be) the case".[3] This interpretation is made explicit in [4] within a language with three modal operators: a unimodal belief operator B, a bimodal conditional operator $>$ and the unimodal global operator $\square$. The interpretation of $B\phi$ is "the agent believes ϕ", the interpretation of $\phi > \psi$ is "if ϕ is (or were) the case then ψ is (or would be) the case" and the interpretation of $\square\phi$ is "ϕ is necessarily true". It is shown in [4] that, for every AGM axiom there is a property of Kripke-Lewis frames that characterizes that axiom; in turn, that property characterizes an axiom in the modal language, so that to each AGM axiom there corresponds a modal axiom; in other words, each AGM axiom can be "translated" into a corresponding modal axiom.

For example, the axiom that "translates" the Success Axiom ($(K{*}2)$: $\phi \in K * \phi$) into the modal language is $B(\phi > \phi)$. This axiom says that the agent believes that "if ϕ is (or were) the case, then ϕ is (or would

[3]Note that we allow for both the indicative and the subjunctive conditional. The indicative form (if ϕ *is* the case then ψ *is* the case) seems to be more appropriate when the initial belief set does not contain $\neg\phi$ (that is, the agent initially considers ϕ possible), while the subjunctive form (if ϕ *were* the case then ψ *would be* the case) seems to be more appropriate when the agent initially believes $\neg\phi$.

be) the case". This translation makes the Success Axiom entirely trivial: any meaningful reading of the conditional $\phi > \phi$ makes the formula $B(\phi > \phi)$ necessarily true. Contrast this with the observations quoted above about on how "restrictive" or "unreasonable" the Success Axiom is!

As another example, the axiom that "translates" the Vacuity Axiom ($(K * 4)$: if $\neg\phi \notin K * \phi$ then $K \subseteq K * \phi$) into the modal language is

$$(\neg B \neg\phi \wedge B(\phi \to \psi)) \to B(\phi > \psi)$$

which says that if, initially, the agent considers ϕ possible and believes that, whenever ϕ is the case then ψ is also the case, then the agent believes that "if ϕ is (or were) the case then ψ is (or would be) the case". For more details and the complete list of translations of the AGM axioms into modal axioms the reader is referred to [4].

3 Conclusion

As noted in the Introduction, the dominant interpretation of an input to AGM revision is in terms of *reliable new information*, so that $K * \phi$ is interpreted as the revised belief set after the information represented by the formula ϕ has been made compatible with the initial belief set K.

Some authors, although certainly in the minority, have argued that the AGM axioms for belief revision are suitable for modeling suppositional beliefs but not for belief change in response to learning new information. For example, Levi writes "the contribution of Alchourrón, Gärdenfors and Makinson is best seen as a contribution to an account of reasoning for the sake of the argument and not as an account of the logic of belief change" ([15, p. 117]).

We have argued that our proposed Kripke-Lewis semantics provides an alternative interpretation of AGM revision in terms of *supposition* rather than information, and that supposition and information are conceptually very different. If we interpret the sentence 'on the supposition that ϕ, the agent believes that ψ' as 'the agent believes that if ϕ is (or were) the case then ψ is (or would be) the case', then the characterization of AGM belief revision in terms of the Kripke-Lewis semantics provided in [3, 4] shows that AGM belief revision can indeed be given a precise and consistent interpretation in terms of supposition rather than information.

There are several questions to be addressed in future work:

1. In what contexts is suppositional belief revision (as opposed to "genuine" belief change) useful and relevant?

2. From a practical point of view, e.g. in the context of decision making or within the logic of action, how does suppositional belief revision differ from belief change in response to new information?

3. Is there a precise way in which one can answer the question whether the AGM belief revision axioms are more appropriate for suppositional reasoning than for belief change in response to new and reliable information?

4. If, in fact, the AGM theory is best understood in terms of suppositional reasoning, what would constitute an appropriate axiomatization of "genuine" belief change?

A The AGM axioms

Consider a propositional logic based on a countable set $\mathtt{At}$ of atomic formulas. Denote by Φ_0 the set of Boolean formulas constructed from $\mathtt{At}$ as follows: $\mathtt{At} \subset \Phi_0$ and if $\phi, \psi \in \Phi_0$ then $\neg\phi$ and $\phi \vee \psi$ belong to Φ_0. Define $\phi \to \psi$, $\phi \wedge \psi$, and $\phi \leftrightarrow \psi$ in terms of $\neg$ and $\vee$ in the usual way (e.g. $\phi \to \psi$ is a shorthand for $\neg\phi \vee \psi$).

Given a subset K of Φ_0, its deductive closure is denoted by $Cn(K)$. A set $K \subseteq \Phi_0$ is *consistent* if $Cn(K) \neq \Phi_0$; it is *deductively closed* if $K = Cn(K)$. Given a set $K \subseteq \Phi_0$ and a formula $\phi \in \Phi_0$, the *expansion* of K by ϕ, denoted by $K+\phi$, is defined as follows: $K+\phi = Cn\,(K \cup \{\phi\})$.

Let $K \subseteq \Phi_0$ be a *consistent and deductively closed* set representing the agent's initial beliefs. A *belief revision function* based on K is a function $* : \Phi_0 \to 2^{\Phi_0}$ (where 2^{Φ_0} denotes the set of subsets of Φ_0) that associates with every formula $\phi \in \Phi_0$ a set $K * \phi \subseteq \Phi_0$, interpreted as the change in K prompted by the input ϕ.[4] The axioms proposed

[4]We follow the common practice of writing $K * \phi$ instead of $*(\phi)$ which has the

by Alchourrón, Gärdenfors and Makinson in [1], known as the AGM axioms, are as follows:

$(K*1)$ (Closure) $K*\phi = Cn(K*\phi)$.

$(K*2)$ (Success) $\phi \in K*\phi$.

$(K*3)$ (Inclusion) $K*\phi \subseteq K+\phi$.

$(K*4)$ (Vacuity) if $\neg\phi \notin K$, then $K \subseteq K*\phi$.

$(K*5)$ (Consistency) $K*\phi = \Phi_0$ if and only if $\neg\phi$ is a tautology.

$(K*6)$ (Extensionality) if $\phi \leftrightarrow \psi$ is a tautology then $K*\phi = K*\psi$.

$(K*7)$ (Superexpansion) $K*(\phi\wedge\psi) \subseteq (K*\phi)+\psi$.

$(K*8)$ (Subexpansion) if $\neg\psi \notin K*\phi$, then $(K*\phi)+\psi \subseteq K*(\phi\wedge\psi)$.

For a discussion of the above axioms, see, for example, [8, 10].

References

[1] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50:510–530, 1985.

[2] G. Bonanno. Filtered belief revision: Syntax and semantics. *Journal of Logic, Language and Information*, 31(4):645–675, 2022.

[3] G. Bonanno. A Kripke-Lewis semantics for belief update and belief revision. *Artificial Intelligence*, 339, 2025.

[4] G. Bonanno. A modal logic translation of the AGM axioms for belief revision. Technical report, arXiv preprint, 2025.

[5] R. Booth, E. Fermé, S. Konieczny, and R. P. Pérez. Credibility-limited revision operators in propositional logic. In *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*, pages 116–125. AAAI Publications, 2012.

advantage of making it clear that the belief revision function refers to a given, *fixed*, K.

[6] C. Cross and R. Thomason. Conditionals and knowledge-base update. In P. Gärdenfors, editor, *Belief Revision*, volume 29 of *Cambridge Tracts in Theoretical Computer Science*, pages 247–275. Cambridge University Press, 1992.

[7] F. Falakh, S. Rudolph, and K. Sauerwald. Agm belief revision, semantically. *ACM Trans. Comput. Logic*, 2025.

[8] E. Fermé and S. O. Hansson. *Belief change: introduction and overview.* Springer, 2018.

[9] M. Garapa. Two level credibility-limited revisions. *The Review of Symbolic Logic*, pages 1–21, 2020.

[10] P. Gärdenfors. *Knowledge in flux: modeling the dynamics of epistemic states.* MIT Press, 1988.

[11] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.

[12] S. O. Hansson. A survey of non-prioritized belief revision. *Erkenntnis*, 50:413–427, 1999.

[13] S. O. Hansson, E. Fermé, J. Cantwell, and M. Falappa. Credibility limited revision. *The Journal of Symbolic Logic*, 66:1581–1596, 2001.

[14] H. Katsuno and A. O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263 – 294, 1991.

[15] I. Levi. *For the sake of the argument.* Cambridge University Press, 1996.

[16] M. Morreau. For the sake of the argument. *Journal of Philosophy*, 95(10):540–546, 1998.

[17] P. Peppas. Chapter 8: Belief Revision. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*, pages 317–359. Elsevier, 2008.

[18] H. Rott. Coherence and conservatism in the dynamics of belief. *Erkenntnis*, 50:387–412, 1999.

[19] K. Segerberg. Irrevocable belief revision in Dynamic Doxastic Logic. *Notre Dame Journal of Formal Logic*, 39(3):287–306, 1998.

[20] R. Stalnaker. Belief revision in games: forward and backward induction. *Mathematical Social Sciences*, 36:31–56, 1998.

[21] J. Zhao, V. Crupi, K. Tentori, B. Fitelson, and D. Osherson. Updating: Learning versus supposing. *Cognition*, 124:373–378, 2012.