A Semantic-Modal View on Ramsey’s Test

Alexandru Baltag* and Sonja Smets†

Abstract

We present a semantic analysis of the Ramsey test, pointing out its deep underlying flaw: the tension between the “static” nature of AGM revision (which was originally tailored for revision of only purely ontic beliefs, and can be applied to higher-order beliefs only if given a “backwards-looking” interpretation) and the fact that, semantically speaking, any Ramsey conditional must be a modal operator (more precisely, a dynamic-epistemic one). Thus, a belief about a Ramsey conditional is in fact a higher-order belief, hence the AGM revision postulates are not applicable to it, except in their “backwards-looking” interpretation. But that interpretation is consistent only with a restricted (weak) version of Ramsey’s test (in-applicable to already revised theories). The solution out of the conundrum is twofold: either accept only the weak Ramsey test; or replace the AGM revision operator ⊞ by a truly “dynamic” revision operator ◐, which will not satisfy the AGM axioms, but will do something better: it will “keep up with reality”, correctly describing revision with higher-order beliefs.

Keywords: Logics for Belief Change, Doxastic Conditionals, Higher-order Belief Revision, Ramsey Test, Gardenfors Impossibility Result.

1 Introduction

The so-called Ramsey test, proposed in 1929 by F. P. Ramsey [20], was meant as an evaluation procedure for conditionals $A > B$, that would capture their role as dispositions for belief change induced by hypothetical assumptions.

*Computing Laboratory, Oxford University, UK. Alexandru.Baltag@comlab.ox.ac.uk
†Rijksuniversiteit Groningen and IEG, Oxford University. S.J.L.Smets@rug.nl
According to this reading, an agent accepts the conditional $A > B$ if and only if she is disposed to come to accept $B$ after learning $A$. The test was formalized by Gardenfors [14, 15] as saying that

$$A > B \in T \iff B \in T * A$$

should hold for all “theories” (beliefs sets) $T$ and all sentences $A, B$. Here, $*$ is a binary “belief revision” operator, meant to capture the revised set of beliefs $T * A$ of an agent whose initial belief set was $T$ after which she learned $A$. Gardenfors’ famous Impossibility (or Triviality) Result essentially shows that the existence of such ‘Ramsey conditionals’ is incompatible with the standard AGM postulates [1], assumed to govern any ‘rational’ belief revision operator $*$. But this leaves open the possibility that some reasonable modification of the AGM postulates might still be consistent with the Ramsey test.

While most discussions of Ramsey conditionals and most “solutions” to Gardenfors’ challenge assume a purely syntactic, and purely propositional, perspective (following in this sense Gardenfors himself and the original AGM approach), in this paper we adopt a semantic, and more specifically modal, point of view. While Ramsey, Gardenfors and others only give conditions for “acceptance” (i.e. belief by an implicit agent) of a Ramsey conditional, we are primarily interested in its truth conditions. Indeed, from a semantic and modal perspective, it makes no sense to ask for belief (or acceptance) of a statement whose extensional meaning was not yet defined. In this sense, Ramsey’s test cannot be taken in itself as a “definition” of a Ramsey conditional $A > B$, since it does not specify in what possible worlds might the conditional be true. So, in a semantic-modal setting, the important question is: are there any truth conditions for $A > B$ that would be compatible with Ramsey’s test (given the usual modal semantics for belief, and given some reasonable semantics for belief revision)?

Our answer is: in general, no. More precisely: if the Ramsey test is required to hold for all ‘theories’ (including the ones representing future belief sets, after possible revisions) and if some reasonable rationality conditions are assumed (such as full introspection of beliefs and of dispositions to believe, as well as un-revisability of beliefs that are in fact “known” to be true, in an absolute sense, e.g. by introspection), then the answer is “no”. So what we offer in this paper is a new, semantical Triviality Result.

Our argument can be summarized as follows. Assuming the usual Kripke semantics for belief operators $Bel$ and assuming any kind of “dynamic” se-
mantics for the belief-revision operator $*$, we argue that the only possible truth conditions for a Ramsey conditional $A > B$ make it equivalent to a type of dynamic-doaxastic conditional, expressing a hypothetical disposition to revise the agent’s beliefs. Hence contrary to the standard (static) interpretation of the Ramsey conditional, our dynamic-doaxastic Ramsey conditional captures a true revision operation for higher-order beliefs. The price we pay for a belief-revision operator that is compatible with our dynamic-doaxastic Ramsey conditional, is that it will not satisfy all the AGM revision postulates. This might sound worrying at first, but we show that the dynamic semantics for the belief revision operator will actually fall in place very naturally. Our argument will state clearly why the static semantics suffers from the mentioned Impossibility result, unless we restrict the application of the Ramsey test to its bare minimum, i.e. excluding its application to “already revised” theories would indeed make it is perfectly compatible with the (epistemic version) of the AGM postulates. This shows the two ways out of this Impossibility or Triviality Result: either we strip the Ramsey test till we are left with a weak version or restore its glory and yield a dynamic semantics for its matching belief-revision operator.

2 Dynamic Semantics for Belief Revision

Let $S$ be any kind of model endowed, among other things, with a set $S$ of possible worlds, or “states” of the world, a doxastic accessibility relation $\rightarrow \subseteq S \times S$ on worlds, an epistemic accessibility relation $\sim \subseteq S \times S$, and a valuation $\parallel \cdot \parallel : Prop \rightarrow \mathcal{P}(S)$ taking atomic sentences $p \in Prop$ into sets $\parallel p \parallel \subseteq S$ of possible worlds. We can also put

$$bel_s := \{ t \in S : s \rightarrow t \}$$

for the set of all “doxastic possibilities” at $s$.

We make four assumptions about these relations: $\rightarrow$ is serial ($\forall s \in S \exists t \in S : s \rightarrow t$); $\sim$ is reflexive; $\rightarrow$ is included in $\sim$ (i.e. $s \rightarrow t$ implies $s \sim t$); and the doxastic possibilities are the same in epistemically related states: i.e. $s \sim t$ implies $bel_s = bel_t$.

These conditions capture natural assumptions about belief and knowledge: seriality of $\rightarrow$ expresses Consistency of Beliefs; reflexivity of $\sim$ captures Veracity of Knowledge; $\rightarrow \subseteq \sim$ captures the fact that Knowledge implies...
Belief; and the last condition expresses “Full Introspection of Beliefs”: the agent knows what she believes and what she doesn’t believe.

Observe that the last two clauses together imply the transitivity and Euclideaness of $\rightarrow$ (i.e. the usual postulates of “Positive and Negative Introspection of Beliefs”). Hence, our belief operators satisfy the axioms of the modal system $KD45$, without us having to assume them (except for the axiom $D$, which corresponds to seriality).

Also observe that we did not assume (Positive or Negative) Introspection for Knowledge. Indeed, many philosophers deny that true “knowledge” has these properties, and we do not want to make our conclusions dependent on such doubtful assumptions. As for the fact that we assumed introspection of beliefs, one can argue that this is a much more realistic assumption. Unlike knowledge, belief does not depend on any external reality, but it is a notion that is “internal” to the agent; hence, it seems natural to assume that a perfectly rational agent will be “infallible” in her beliefs about her own beliefs: she will know what she believes and what she doesn’t believe.

Given some language $L$ consisting of a collection of well-formed formulae (which might be purely propositional, or might include belief operators, or other operators), we define, for every formula $\varphi \in L$, the interpretation $\|\varphi\|_S$ of $\varphi$ in the model $S$ to be the set of all states satisfying $\varphi$:

$$\|\varphi\|_S := \{ s \in S : s \models \varphi \}.$$  

When the model is fixed, we skip the subscript $S$.

As usual in Kripke semantics, we do not identify a state of the world with its purely ontic content (represented by the valuation: the atomic sentences $p$ satisfied by a state $s$ represent the “ontic facts” of the world $s$). The world consists of more than just ontic facts: for instance, an (implicit) agent’s knowledge and beliefs at world $s$ are also determined (via the doxastic accessibility relation) by the world $s$.

Indeed, we can always extend $L$ with knowledge and belief operators $Kn$ and $Bel$, using the usual Kripke semantics: $Kn$ is defined as the Kripke modality for the epistemic relation $\sim$, and $Bel$ is defined as the Kripke modality for the doxastic accessibility relation $\rightarrow$:

$$s \models KnA \text{ iff } t \models A \text{ for all } t \text{ such that } s \sim t,$$

$$s \models BelA \text{ iff } t \models A \text{ for all } t \text{ such that } s \rightarrow t.$$
This last one is equivalent to (quantifying over all the doxastic possibilities)

\[ \| \text{Bel} A \| = \{ s \in S : \text{bel}_s \subseteq \| A \| \}. \]

For a given language \( \mathcal{L} \), the agent’s belief set, or “theory”, (at any state \( s \in S \)) is the set \( T_s \) of \( \mathcal{L} \)-formulas that are believed at state \( s \):

\[ T_s := \{ \phi \in \mathcal{L} : s \models \text{Bel}\phi \} = \{ \phi \in \mathcal{L} : \text{bel}_s \subseteq \| \phi \| \}. \]

To give now a semantics to the revision operator, assume to be given, for each sentence \( A \in \mathcal{L} \), some partial, injective unary operator \( *A \) on the set \( S \) of states. Alternatively, this can be described relationally: for each sentence \( A \in \mathcal{L} \), we are given some binary relation \( *A \rightarrow \subseteq S \times S \), having four properties:

- **(Partial Functionality)** \( s *A t, t' \Rightarrow t = t' \);
- **(Injectivity)** \( s, s' *A t \Rightarrow s = s' \);
- **(Perfect Recall)** \( s * A \sim t * A \Rightarrow s \sim t \);
- **(Full Introspection of Revision)** \( s \sim t \Rightarrow \text{bel}_{s*A} = \text{bel}_{t*A} \).

If it exists, the unique state \( t \) such that \( s *A t \) is denoted by \( s * A \). The state \( s * A \) is meant to represent the state of the world after the agent performed a belief revision with \( A \) in an original state of the world \( s \).3

The Partial Functionality postulate expresses that a specific revision changes the state in a determinate way. The Injectivity postulate captures the “uniqueness of the past” (of a given state), a standard assumption in temporal logic. The “Perfect Recall” is the assumption that knowledge is never lost or forgotten after revision. Finally, “Full Introspection of Revision” is a strengthening of the postulate of “Full Introspection of Beliefs”, saying that the agent’s belief-revision dispositions are fully known to her: she knows what she would believe and what she would not believe after revising with a given sentence. One can justify this assumption in the same way as we justified Full Introspection of Beliefs: belief-revision plans or “dispositions” are internal to the agent, so the agent’s beliefs about these dispositions are “infallible”.

The partiality of our operation \( * \) is explainable by the epistemic constraints posed to revision by any notion of un-forgettable knowledge: in contrast to the classical AGM setting, agents should not be able to revise
with “information” they know to be false! (Any such revision would amount either to forgetting prior knowledge, or to acquiring inconsistent beliefs.) As we’ll see later (and as argued in [4]) this requires replacing the AGM “Vacuity Axiom” by a weakened version (essentially saying that a revised theory $T \ast A$ is consistent iff $T$ did not entail knowledge of the falsity of $A$). In terms of the epistemic relation, this says that $s \ast A$ exists if and only if $s \models \neg K \neg A$. We do not make this requirement yet here, but we only leave open this possibility by allowing the operation $\ast$ to be partial on states.

Given such a setting, we can add to our language dynamic operators $[\ast A]B$ (in the style of Propositional Dynamic Logic), expressing the way in which revising one’s beliefs with $A$ changes (the state of) a given world. More precisely, $[\ast A]$ is the Kripke modality for the $\ast A$-transitions $\ast A$:

$$s \models [\ast A]B \text{ iff } s \ast A \models B \text{ whenever } s \ast A \text{ exists}.$$ 

So $[\ast A]B$ says that, if the agent can consistently revise with $A$, then $B$ will be true after the revision.

Dually, one can use the injectivity of $\ast A$ to define a backward-looking temporal operator $\text{Before}$, as follows:

$$w \models \text{Before} B \text{ iff } s \models B \text{ for all } s \text{ such that } w = s \ast A.$$ 

Given our operators, the semantical assumptions we made can be converted into a number of validities:

**Normality:** $Kn, Bel, [\ast A], \text{Before}$ satisfy Kripke’s axiom and Necesitation Rule,

**Consistency of Beliefs:** $\neg Bel \bot$,

**Veracity of Knowledge:** $Kn \varphi \Rightarrow \varphi$,

**Knowledge Implies Belief:** $Kn \varphi \Rightarrow Bel \varphi$,

**Full Introspection of Beliefs:**

$$Bel \varphi \Rightarrow Kn Bel \varphi,$$

$$\neg Bel \varphi \Rightarrow Kn \neg Bel \varphi,$$

**Partial Functionality of Revision:** $\neg [\ast A] \varphi \Rightarrow [\ast A] \neg \varphi$,

**Injectivity (Unique Past):** $\neg (\text{Before} \varphi) \Rightarrow (\text{Before} \neg \varphi)$,
Perfect Recall: $Kn\varphi \Rightarrow [\ast A]Kn(Before \varphi)$,

Full Introspection of Belief-Revision Plans:

$[\ast A]Bel\varphi \Rightarrow Kn[\ast A]Bel\varphi$

$[\ast A]\neg Bel\varphi \Rightarrow Kn[\ast A]\neg Bel\varphi$

Adjunction: $\varphi \Rightarrow [\ast A](Before \varphi)$.

The last expresses the temporal duality between the dynamic revision operator $[\ast A]$ (going forward in time) and the $Before$ operator (going backwards).

As a side note, let us mention that this general setting subsumes the usual Grove-sphere semantics for belief revision as a special case: Grove sphere models can be equivalently presented (“Stalnaker-style”) as Kripke models $S$ with a family of (conditional) doxastic accessibility relations $\rightarrow^A$ (one for each set $A$ of possible worlds), satisfying some conditions (and such that belief $Bel$ is the Kripke modality for the relation $\rightarrow^S$ associated to the set $S$ of all possible worlds): this is the setting of “conditional doxastic models” [4], a modal equivalent of Grove-sphere models. In their turn, these models are equivalent with the so-called preference models (or “plausibility models”) [5, 6, 7, 8], in which a (well-founded, total) preorder $\leq$ is given on the worlds: essentially, $s \leq t$ means that world $s$ is at least as plausible (for our agent) as world $t$. The equivalence is established by having the conditional doxastic relations $\rightarrow^A$ defined by quantifying over all the “minimal” (=most plausible) worlds satisfying the condition $A$: $s \rightarrow^A t$ iff $t \in Min_{\leq}A$. The semantic revision operator $\ast$ can also be defined on such models, using e.g. either Boutilier’s “minimal revision” with $A$ (defined by changing the plausibility order, in such a way that all the “most plausible” $A$-worlds become the most plausible overall, while in rest nothing changes), or Spohn’s “lexicographic revision” (defined by changing the plausibility order, in such a way that all the $A$-worlds become strictly more plausible than all the non-$A$-worlds, while keeping the same order within any of the two zones), or any of the other of the many proposals encountered in the literature on Belief Revision. (See [21] for an overview of many such proposals.)
3 The Modal Nature of Ramsey Conditionals

Until now, we only formalized the revised state of the world $s * A$ after an agent’s revision with $A$. In fact the semantic conditions we required are consistent with interpreting $*A$ as any action that (1) respects perfect recall (no knowledge is forgotten after $*A$ is performed), and (2) whose belief-changing effects are already known to the agent before the action. But what does make this action into an action of belief revision with sentence $A$?

More specifically, how should we define the revised theory, or set of beliefs $T * A$ of the agent in a world $s$? Recall that the original belief set $T$ at state $s$ was given by the set

$$T := T_s = \{ \varphi \in L : s \models Bel(\varphi) \}$$

of all the sentences believed by the agent at $s$. What then should be the theory $T_s * A$?

Let us postpone the answer and simply assume for the moment that we are given, in addition to the partial operator $*A$ on states in $S$, an operator $*A$ on “theories” (sets of sentences), satisfying the AGM postulates, and such that $T_s * A$ intuitively captures the agent’s revision with $A$ of her original theory $T_s$ (at state $s$).

Let us suppose now that our language $L$ contains a Ramsey conditional operator $A > B$, satisfying the Ramsey test, in particular for theories of the form $T = T_s$:

$$A > B \in T_s \iff B \in T_s * A.$$ 

By using our definition for the belief set $T_s$ at a state $s$, we can unfold this to obtain:

$$s \models Bel(A > B) \iff B \in T_s * A.$$ 

But this gives us only a semantical clause for believing a Ramsey conditional, while a true, full semantics should give us a truth clause for this conditional: when does $s \models A > B$ hold?

As it has been argued by a number of authors (and in particular most convincingly by H. Leitgeb [17]), there is no way to define such a Ramsey conditional $A > B$ in a “purely ontic”, non-doxastic way. (In particular, one cannot identify $A > B$ with an “objective” counterfactual conditional.) In addition to the numerous arguments in the literature, we add here two more arguments in this sense.
Our first argument goes in a sense beyond this paper, since it belongs to a multi-agent setting: it is very easy to construct a formal model (as well as to give realistic examples) in which two agents have exactly the same “first-level” beliefs (about the ontic facts of the world), but have different belief-revision dispositions (i.e. different ways to revise their beliefs when given the same information). This shows that belief revision dispositions are not uniquely determined by the facts of the world and the agent’s beliefs about these facts.

A more basic argument is that, in principle, the meaning of a Ramsey conditional should be completely given by the Ramsey test. But this test, as seen above, does not refer at all to ontic facts of the world: it’s a purely “internal” test, having to do with the agent’s own dispositions to revise her beliefs. The meaning of $A > B$ should thus be recoverable only by looking at the agent’s belief-revising plans or dispositions. Taken as such a truth clause, the test Ramsey test becomes stronger:

$$s \models A > B \text{ iff } B \in T_s * A.$$  

**Ramsey Conditionals are Modal, Higher-Order Statements** The conclusion is that a statement of the form $A > B$ is non-ontic in nature, a modal, indeed, a dynamic-doxastic statement, talking about (potential changes of) beliefs! Such a statement is internal to the agent, hence fully introspective: the agent’s beliefs about it are infallible. Like belief-revisions dispositions themselves, the Ramsey conditional is “known” (or believed) to hold if and only if it actually holds:

$$A > B \iff Kn(A > B) \iff Bel(A > B).$$

(This follows from our other assumptions if we accept the above truth clause for $A > B$.) Since in conclusion a Ramsey conditional $A > B$ is a dynamic-doxastic operator, any belief set $T$ such that $A > B \in T$ will represent a higher-order belief: a belief about the agent’s own (potential changes of) beliefs.

**The Catch: AGM Fails for Higher-Order Beliefs** However, as often remarked in the Dynamic Epistemic Logic literature [12, 13, 4, 7], the AGM postulates are not adequate for revision of higher-order beliefs! In particular, there are problems with the the AGM “(Non-)Vacuity” Postulate, as well as with the “Success” Postulate.
The Problem with (Non-)Vacuity

The AGM “(Non-)Vacuity” postulate says a revised theory is inconsistent only if the new information is inconsistent:

\[ T \ast A = \bot \iff T = \bot \]

In other words: revision with a consistent sentence can always be performed, and leads to a consistent theory.

But this postulate is obviously inadequate for a language that includes knowledge operators and higher-order beliefs! Indeed, if the negation of \( A \) was already “known” , in an “absolute” sense, then the agent simply should not accept to revise with \( A \): one cannot accept to come to believe \( A \), when already knowing \( \neg A \). The revision should fail in this case! One might object that such “absolute knowledge” is unrealistic and incompatible with the spirit of the AGM postulates: according to that spirit, the only things that are “known” in this absolute sense are tautologies (and hence, revision only fails when revising with inconsistent sentences!)

However, this objection cannot be applied to languages that can express higher-level beliefs! If \( A \) is itself of the form \( \neg \text{Bel } p \), and \( p \) was a sentence that was already believed, then an introspective agent knows \( \neg A \) in an “absolute” sense, by introspection, even if it isn’t a tautology (since it’s equivalent to \( \text{Bel } p \))!

The unavoidable conclusion is that, in its original version, the (Non-)Vacuity postulate is inapplicable to languages that (can express doxastic-epistemic operators and thus) have higher-order beliefs, and in particular inapplicable to our language \( \mathcal{L} \) (assumed to contain a Ramsey conditional \( A > B \), which as we saw is a higher-level doxastic statement!).

Restricting the (Non-)Vacuity Postulate: “Epistemic AGM”

Still (as already proposed in [4]), there is an obvious way to solve this problem: restrict the Non-Vacuity postulate, by simply replacing in it “inconsistent statements” with “statements known to be false”. The “Epistemic Non-Vacuity” postulate will say now that: the revised beliefs \( T \ast A \) are consistent only if the new information \( A \) was not already “known” to be false (i.e. only if \( Kn\neg A \notin T \)). Formally:

\[ T \ast A = \bot \iff Kn\neg A \in T. \]

In [4] we proposed this axiom, and we called the resulting modified system “epistemic AGM”.

10
The Problem with Success

Intuitively, there also seems to be a tension between the AGM “Success” postulate and the presence of higher-order doxastic statements in our language $\mathcal{L}$. The Success postulate

$$A \in T * A$$

can be interpreted as saying that: “after revising with $A$, the agent will believe $A$”. But this is intuitively wrong, when applied to higher-order statements, such as Moore sentences $p \land \neg\text{Bel} p$ (“you don’t believe $p$, but nevertheless $p$ is the case”). No matter what you do after revising your beliefs with such a Moore sentence, you shouldn’t try believing this sentence. Indeed, believing it leads to inconsistency: the sentence $\text{Bel}(p \land \neg\text{Bel} p)$ is inconsistent, given our above assumptions about belief. By the (classical AGM) Non-Vacuity postulate, such an inconsistency should not be believed after revising with the (consistent) sentence $p \land \neg\text{Bel} p$. But even applying our revised “Epistemic Non-Vacuity” principle, the same conclusion follows: since (if $p$ was indeed not believed before the revision, i.e. $p \not\in T$, then) the Moore sentence was not known to be false before revision (i.e. $K(p \land \neg\text{Bel} p) \not\in T$), and hence the revision should not lead to inconsistency!

However, this argument is not really showing the failure of the Success postulate, but only of a certain interpretation of it! Indeed, when applied to the Moore sentence, the Success postulate simply says that

$$(p \land \neg\text{Bel} p) \in T * (p \land \neg\text{Bel} p)$$

One cannot actually derive from this a contradiction inside the theory $T * (p \land \neg\text{Bel} p)$. The contradiction can be derived only externally (semantically) and only if we interpret theory revision $T * A$ as above; more precisely, assuming that the original state is $s$ (and thus the original belief set is given by $T_s = \{\varphi \in \mathcal{L} : s \models \text{Bel}\varphi\}$), then the above (doubtful) interpretation takes $T_s * A$ to consist of the agent’s beliefs after revision about the new state of the world $s * A$ (as it is after revision). Formally, this doubtful interpretation is embodied by the following “definition”:

$$T_s * A \triangleq \{\varphi \in \mathcal{L} : s \models [*A \text{Bel}\varphi]\}.$$

However, the informal argument above shows that, if we want to preserve the AGM axiom of “Success” for a language $\mathcal{L}$ that contains a Ramsey conditional $A > B$, we simply cannot accept this tentative “definition” (?) for theory revision, on pain of contradiction!
Indeed, we can make this into a formal argument: if we accept (?), then by unfolding the “Success” postulate

\[ A \in T_s * A \]

for the sentence \( A := p \land \neg \text{Bel } p \), we obtain

\[ \left[ * (p \land \neg \text{Bel } p) \right] \text{Bel} (p \land \neg \text{Bel } p) . \]

Since (as we already saw) \( \text{Bel} (p \land \neg \text{Bel } p) \) is inconsistent, hence we obtain \( \left[ * (p \land \neg \text{Bel } p) \right] \bot , \) and as a consequence \( \left[ * (p \land \neg \text{Bel } p) \right] \text{Bel} \bot , \) and thus (again by the “definition” ?) \( \bot \in T * (p \land \neg \text{Bel } p) , \) and hence \( T * (p \land \neg \text{Bel } p) = \bot . \) This contradicts, not only the classical AGM Non-Vacuity postulate, but even our “Epistemic Non-Vacuity” version: if \( s \) is any state at which \( p \) is not believed, i.e. \( T_s \) is any consistent theory not containing \( p , \) then \( K(p \land \neg \text{Bel } p) \notin T , \) and so by Epistemic Non-Vacuity \( T * (p \land \neg \text{Bel } p) \neq \bot . \)

The conclusion is that, if we want to maintain the AGM postulates (even in the restricted form of “Epistemic AGM”) in the presence of higher-level doxastic statements in our language, then the interpretation given to \( T * A \) by the definition (?) is simply wrong: it just does not give us an AGM-type of theory revision. It’s not really a failure of the “Success” postulate in itself, it’s a failure of a specific interpretation of *.

The Correct Interpretation of AGM Revision Nevertheless, as we argued elsewhere [4, 7, 5, 8], the AGM belief revision operator can be easily interpreted in such models, in a way that satisfies all the postulates of “Epistemic AGM”. Moreover, this interpretation is “natural”, in the sense that it captures the “static” spirit of the AGM axioms: AGM is a system for revising about a “static” world (assumed not to change during revision). Since, in a modal (higher-level) setting, the world is always changed by our changes of beliefs (since “the world” is not reduced to ontic facts, but includes our beliefs), the only way to consistently interpret AGM revision in such a setting is to simply insist on not changing the world: just interpret \( T_s * A \) as consisting of the agent’s revised beliefs (after revision with \( A \)) about the old state of the world \( s \) (as it was before revision). So roughly, \( B \in T * A \) means that the agent has the following disposition: in case she’d have to revise with \( A , \) she’ll come to believe that \( B \) was the case (before the revision).

The Correct Definition of Theory Revision Formally, instead of the “definition” (?) above, we adopt the following revised version:

\[ T_s * A = \{ \varphi \in \mathcal{L} : s \models [A] \text{Bel} (\text{Before } \varphi) \} . \]
Unfolding this definition, we get:

\[ T_s \ast A = \{ \varphi \in \mathcal{L} : s \ast A \models Bel(Before \varphi) \} \],

whenever \( s \ast A \) exists,

and \( T_s \ast A = \bot \), otherwise.

Given this reinterpretation, the resulting semantics for Ramsey conditionals is the one of a “backward-looking” dynamic-doxastic conditional

\[ A > B := [\ast A]Bel(Before B) \]

So \( A > B \) means that, in case the agent would have to revise with \( A \), she’d come to believe that \( B \) was the case (before the revision). \(^4\)

\[
\begin{array}{ccc}
T_s & \left\{ \psi, A > \varphi, \ldots \right\} & \overset{\ast A}{\rightarrow} \left\{ \varphi, \ldots \right\} & T_s \ast A \\
\downarrow & \downarrow & \downarrow & \downarrow \\
\relphantom{\{ \psi, A > \varphi, \ldots \}} & \relphantom{\left\{ \varphi, \ldots \right\}} & \relphantom{T_s \ast A} & \relphantom{s \ast A}
\end{array}
\]

In the above diagram, time flows from left to right; the theories above represent belief sets containing beliefs hold by the agent at the corresponding states straight below them; the continuous arrows indicate temporal transitions (from the old state to the new state, and from the old theory to the new theory); the dotted arrows go from any theory to the state/world that it refers to, and that gives its truth condition: e.g. the belief in \( \varphi \) (held at state \( s \ast A \), and thus belonging to the theory above \( T_s \ast A \)) is correct iff \( \varphi \) holds at the original world \( s \).

If we iterate now AGM belief revision, we see even better its “static” character: all theories \( T_n = T \ast A_1 \ast \cdots \ast A_n \) obtained by iterated revision (theories held as beliefs at successive states of the world \( s_n = s \ast A_1 \ast \cdots \ast A_n \)) refer back to the original world \( s_0 = s \) (before the start of the revision process).

\[
\begin{array}{c|c|c}
Belief Sets & Initial(0) & Next(1) & Next(2) \\
\hline
T_0 \overset{\ast A_1}{\rightarrow} T_1 = T_0 \ast A_1 & \overset{\ast A_2}{\rightarrow} T_2 = T_1 \ast A_2 \\
\hline
Reality & s_0 \overset{\ast A_1}{\rightarrow} s_1 = s_0 \ast A_1 & \overset{\ast A_2}{\rightarrow} s_2 = s_1 \ast A_2 \\
\end{array}
\]
This “static”, “back-referential” character of AGM belief revision underlies the following observation (O):

\[(\text{Bel} \varphi) \in T_s * A_1 * \cdots * A_n \text{ iff } \varphi \in T_s.\]

Indeed, since \(T_s * A_1 * \cdots * A_n\) contains (re-revised) beliefs about \(s_0\), we have \((\text{Bel} \varphi) \in T_s * A_1 \cdots A_n\) if and only if at state \(s_n = s_0 * A_1 * \cdots * A_n\) the agent believes that \(\text{Bel} \varphi\) was true at state \(s_0 = s\), i.e. if and only if \(\varphi \in T_s.\)

Another important remark is that all the axioms of “Epistemic AGM” are consistent with this definition, even if we interpret them as postulates about “iterated theories”, i.e. theories \(T\) of the form

\[T = T_s * A_1 * A_2 * \cdots * A_n\]

for any state \(s \in S\), natural number \(n\) and sentences \(A_1, \ldots, A_n\) in any language \(L\). (Here, \(T_s := \{ \varphi \in L : s \models \text{Bel} \varphi \}\), as usual.) These axioms don’t all hold automatically, but we can add them as extra postulates, and they will simply correspond to adding a number of (consistent) semantic constraints.

4 Triviality/Impossibility Result and Two Ways Out

In the previous sections we argue that: (1) in a semantic setting, the only possible interpretation of a Ramsey conditional is as a dynamic-doxastic operator (encoding the agent’s belief-revision dispositions); (2) classical AGM postulates cannot be applied to an object-language containing doxastic operators (and hence, in particular, to one containing Ramsey conditionals), unless the (Non-)Vacuity postulate is restricted (obtaining “Epistemic AGM”); (3) the only natural semantics for theory revision operator \(T \ast A\) for such (theories formulated in) languages containing doxastic operators is a “static” one (encoded in our “correct definition” above): interpreting \(T_s \ast A\) as the agent’s revised beliefs (after revision with \(A\)) about the original state of the world \(s\) (as it was before revision).

In conclusion, we do have now a setting (given by our semantic postulates and definitions) for knowledge, belief, state revision \(S \ast A\) and theory revision \(T \ast A\) that is consistent with the axioms of “Epistemic AGM”, and that seems to be the natural semantical setting for a theory of higher-level belief revision.

However, we will now show that these semantic assumptions are “inconsistent” with the existence of a Ramsey conditional in our language! More precisely, we
only obtain a “triviality” result (showing belief revision collapses in triviality). To prove it in its strongest form, we will only use the following important consequences of our semantic clauses above:

**Closure under Logical Consequence (C):**

\[ A \in T \text{ and } \vdash A \Rightarrow B \text{ imply } B \in T \]

**Normality (K), Consistency (D), Positive (4) and Negative Intropection (5):**

\[ \vdash Bel(A \Rightarrow B) \Rightarrow (Bel A \Rightarrow Bel B) \]

\[ \vdash \neg Bel \perp \]

\[ \vdash Bel A \Rightarrow Bel Bel A \]

\[ \vdash \neg Bel A \Rightarrow Bel \neg Bel A \]

**Internality of Ramsey Conditional (I):**

\[ \vdash A > B \Leftrightarrow Bel(A > B) \]

**Preservation of Higher-Order Beliefs (P):**

\[ (Bel \varphi) \in T \text{ implies } (Bel \varphi) \in T * A \]

**Determinacy of Higher-Order Beliefs (DOB):**

either \((Bel \varphi) \in T\) or \((\neg Bel \varphi) \in T\)

To show that these are satisfied in our setting, notice that we already justified (C), (KD45) and (I). As for (P), we need to check for theories of the form \(T = T_s * A_1 * \cdots * A_n\): by Observation(O) in the previous Section, \((Bel \varphi) \in T = T_s * A_1 * \cdots * A_n\) iff \(\varphi \in T_s\), and again by (O), we similarly have \((Bel \varphi) \in T * A = T_s * A_1 * \cdots * A_n * A\) iff \(\varphi \in T_s\). Hence, \((Bel \varphi) \in T * A\) iff \((Bel \varphi) \in T\).

As for (DOB), to show it for \(T = T_s * A_1 * \cdots A_n\): suppose that \((Bel \varphi) \notin T_s * A_1 * \cdots * A_n\). By the same Observation (O), this implies \(\varphi \notin T_s\), i.e. \(s \not\models Bel \varphi\), and hence \(s \models \neg Bel \varphi\). By (4), \(s \models Bel(\neg Bel \varphi)\), and hence \((\neg Bel \varphi) \in T_s\). By (O) again, this is equivalent to \((\neg Bel \varphi) \in T_s * A_1 * \cdots * A_n\), and hence by (D45) and (C) we get \((\neg Bel \varphi) \in T_s * A_1 * \cdots * A_n\).

On the other hand, these consequences can be taken as natural postulates in their own respect: they are justifiable directly. We already argued extensively in the previous sections for (C), (KD45) and (I). As for (P) and (DOB), they follow directly from our interpretation of (iterated) revised theories as beliefs about the original state of the world (before any iterated revision) together with the assumptions of Full Introspection of Beliefs and Perfect Recall.
Proposition (Triviality Result 1) Assume ((C), (KD45), (I), (P) and (DOB), and also assume that our Ramsey conditional satisfies the Ramsey test. Then, for all theories \( T \) and sentences \( A \) and \( C \), we have:

\[
T \ast C \neq \bot \text{ implies } (T \ast C) \ast A = T \ast A
\]

Proof. From right to left: Let \( B \in T \ast A \). By the Ramsey test, \( (A > B) \in T \). By (I) and (C), we get \( \text{Bel}(A > B) \in T \). By (P), \( \text{Bel}(A > B) \in T \ast C \). By the Ramsey test again, we obtain \( B \in (T \ast C) \ast A \).

From left to right: We prove the counter-positive. Let \( B \notin T \ast A \). By the Ramsey test, \( (A > B) \notin T \). By (I) and (C), we get \( \text{Bel}(A > B) \notin T \). By the Expansion postulate, we get \( \text{Bel}(A > B) \notin T \ast C \). Using (KD45) again and (C), we get \( \text{Bel}(A > B) \notin T \ast C \). From this together with the consistency of \( T \ast C \) (the assumption that \( T \ast C \neq \bot \)), it follows that \( \text{Bel}(A > B) \notin T \ast C \). By (I) again, we obtain \( (A > B) \notin T \ast C \), and hence (by the Ramsey test) \( B \notin (T \ast C) \ast A \). QED

Corollary (Triviality Result 2) In addition to the assumptions in the previous Proposition, assume the following easy, weak consequence of the AGM Expansion Axiom\(^6\): \( T \ast \text{true} = T \), i.e. revision with a tautology preserves the original theory \( (T \ast \text{true} = T) \). Then revision either fails or is trivial:

\[
T \ast C \neq \bot \text{ implies } T \ast C = T
\]

Proof. Take \( A =: \top \). QED.

Corollary (Triviality Result 3) In addition to the assumptions in the Proposition above, assume either AGM Expansion Axiom, or else the assumption in the previous Corollary \( (T \ast \text{true} = T) \) together with the AGM “Success” Axiom \( (C \in T \ast C) \). Then revision with new information always fails:

\[
C \notin T \text{ implies } T \ast C = \bot
\]

Proof. Assume \( C \notin T \). If we had \( T \ast C \neq \bot \), then by Corollary 2, \( T \ast C = T \). Hence \( C \notin T \ast C \) (since \( C \notin T = T \ast C \)), which contradicts both the “Success” postulate, and the Expansion postulate (since, for \( T \ast C \neq \bot \), Expansion tells us that \( T \ast C \) is the closure of \( T \cup \{C\} \), and hence that \( C \in T \ast C \). QED

First Way Out: Stick with Restricted Ramsey! Finally, we propose two different ways to avoid the Triviality result, based on our semantic perspective. The first is to simply restrict the validity of Ramsey’s test only to “theories” that correspond to actual belief sets (in a possible world \( s \)) about the current world \( (s \) itself): i.e. theories of the form \( T = \{ \varphi : s \models \text{Bel}\varphi \} \). This excludes the application of the test to “already revised” theories \( T \ast C \) (since as we saw,
these represent a very different kind of theories, expressing revised beliefs about a past state of the world, before revision). We call this the weak Ramsey test, and it is perfectly compatible with the “epistemic AGM” postulates. Indeed, the conditional belief operator $Bel^A B$ (defined over Grove-sphere models in the usual way, or over conditional doxastic models as the Kripke modalities for the relations $\rightarrow^A$, in the way mentioned above) will satisfy the weak Ramsey test.

Another Way Out: Move to Dynamic Belief Revision! The second way out is to simply replace the “backward-looking” revision operator $*$ by its truly “dynamic” counterpart: an operator $T \otimes A$, capturing the agent’s revised beliefs about the “new” state of the world, as it is after the revision. Formally, the semantics of $\otimes$ is simply the one we previously used as a tentative definition for $*$: if $s$ is the original state of the world and $T_s = \{ \varphi : s \models Bel \varphi \}$ is the original belief set, then $T_s \otimes A := \{ \varphi : s \models [\ast A]Bel \varphi \}$. We can see that $T_s \otimes A = T_{s \ast A}$ whenever $s \ast A$ exists. In a sense, $\otimes$ is the “true” revision operation for higher-order beliefs, since it takes into account the changes induced (on the truth values of doxastic sentences) by the revision itself! The revised theories $T \ast A$ are now of the same type as the original theories $T$ (namely, representing the agent’s beliefs at some given moment about the state of the world at that very moment). Unsurprisingly, dynamic belief revision $\otimes$ is consistent with the unrestricted Ramsey test! (Indeed, just take $A \triangleright B := [\ast A]Bel B$ as your “conditional”.) Unlike the AGM revision $\ast$, the dynamic revision $\otimes$ will not satisfy the “epistemic AGM” postulates. But it will do something better: it’ll keep up with the change of reality!

<table>
<thead>
<tr>
<th>Initial(0)</th>
<th>Next(1)</th>
<th>Next(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belief</td>
<td>$T^0 \otimes A_1 \rightarrow T^1 = T^0 \otimes A_1 \otimes A_2 \rightarrow T^2 = T^1 \otimes A_2$</td>
<td></td>
</tr>
<tr>
<td>Reality</td>
<td>$s_0 \rightarrow s_1 = s_0 \ast A_1 \rightarrow s_2 = s_1 \ast A_2$</td>
<td></td>
</tr>
</tbody>
</table>

Conclusions and Comparison with Other Work Our semantical analysis points out to the deep underlying reasons for the failure of the Ramsey test: (1) any conditional satisfying the test will be of a dynamic-doxastic nature, representing hypothetical beliefs, or rather dispositions for belief change; (2) the test treats hypothetical beliefs about beliefs in the same way as hypothetical beliefs about facts. The test would succeed only if, when making a hypothesis, agents would
revise their beliefs about their own belief revision in the same way they revise their factual beliefs. But this requirement is inconsistent with the restrictions posed by introspective knowledge to belief revision: introspective agents know their own hypothetical beliefs, and so cannot accept hypotheses that go against their knowledge. Due to introspection, beliefs about one’s own belief revision policy cannot be revised, at least not in the sense of the “static” (AGM) belief revision. Only a “dynamic” kind of belief revision, representing the revised beliefs about the situation after the revision, can satisfy the unrestricted version of the Ramsey test.

To compare our solution with other approaches in the literature [17, 22, 18, 19], let us focus on one of the major differences. Indeed, the doxastic, non-ontic nature of Ramsey conditionals has already been noted by a number of authors, and most recently and most clearly was articulated by H. Leitgeb [17]. But, first, Leitgeb extracts a different, and rather debatable, conclusion from this, namely that Ramsey conditionals have “no truth values”, and that they “do not denote propositions”. In contrast, our conclusion is much simpler and (in our view) more natural: these conditionals are modal (in fact, dynamic-doxastic) operators, which form modal statements (denoting modal propositions), that refer to the agent’s own hypothetical beliefs; and thus any belief set $T$ that includes such conditional statements represents a higher-order belief. Secondly, we go further than any of the authors who wrote on the doxastic nature of Ramsey conditionals, by combining this observation with another (independently known) fact, namely that the AGM postulates must fail for higher-order beliefs. Our original conclusion is that this inner conflict lying at the core of the (unrestricted) Ramsey test provides the deeper explanation for its failure.

Notes

1 Indeed, we did not assume that the epistemic relation is transitive or Euclidean, but only reflexive, which corresponds to assuming only Veracity of Knowledge.
2 On the other hand, Williamson and others give arguments against these assumptions.
3 In the Dynamic Epistemic Literature, the state $**A$ is described as living in a different, “updated” model $S * A$ than the original model $S$. But one could of course always assume that all required states live in one big model (by taking as model the disjoint union of the original model $S$ with all the models $(S * A_1) * \cdots * A_n$ obtained by repeated updates). So one can always assume that the original model is closed under the $*A$-transitions $*A$. We choose to do this here, so we can keep track of the past history of any revised state.
4 In the standard Grove-sphere models, this is simply equivalent to the “conditional belief” operator $Bel^A B$, also denoted sometimes by $Bel(B|A)$, and defined as the Kripke modality for the doxastic accessibility relation $\rightarrow^A$; see 11, 9, 4.
5 In fact, the actual proof of this Observation is a more intricate induction argument and uses most our postulates, including Perfect Recall, Full Introspection of Beliefs and
Belief-Revision Plans, Injectivity and the properties connecting knowledge and belief.

The Expansion Axiom is one of the few quasi-universally accepted AGM postulates: if $T$ is consistent with $A$, then $T * A$ is the logical closure of $T \cup \{A\}$.

References


