The Ramsey test as an inference rule

Gabriella Crocco and Andreas Herzig

April 30, 1997

1 Introduction

One of the main features of pragmatism concerns the definition of the notion of belief, which is generally considered to be central in knowledge theory. Recent developments in AI can be seen from this philosophical background. There are mainly two basic ideas that AI retains from philosophical pragmatism:

- Assertive sentences are the linguistic expressions of beliefs.

- Beliefs have to be defined by reference to the actions to which their assertion them would lead.

A set of beliefs is thus a mental state, a sort of cognitive map of "reality". As paths can be marked or taken away from a real map, beliefs can be added or subtracted from a cognitive map [Sah88]. Cognitive maps are then judged in relation to their causal properties.

An important consequence of the pragmatical redefinition of the notion of belief is a new theory of meaning, defined in relation to possible causes and effects to which the assertion of a belief leads. A major example of such a theory of meaning is Ramsey's account of the meaning of conditionals. He said in a footnote in [Ram31, pp. 247, 248], which became a classic in AI approaches to revision :

> "If two people are arguing about 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q".

This so-called Ramsey test (RT henceforth) has been used as a way of giving meaning to conditionals by basing them on an operation of belief change. We call such changes updates. Given a set of beliefs B and a new piece of belief A (that we call the input), we speak of the update of B by A. Moreover, we note C the consequence of an update.

Hansson [Han95] claims that the RT was at the base of the three main tentatives to give meaning to conditionals, viz. Goodman's metalinguistic (derivability) theory, Stalnaker and Lewis' world selection analysis, and Gärdenfors and colleagues' belief revision account. Without any doubts the RT has a lot of intutive appeal. But if we take it for granted, then what does it mean to update a set of beliefs, or - using our metaphore - what are the effects of adding a new path on a map?

We can put the question in another way: Does the RT explain a complex notion (viz. conditionals) by a simpler one (viz. belief change)? The answer seems to be negative: Apparently both notions are equally vague (in the sense of [Lew73, 4.2, p. 95]), and equally complex. We have thus two actions expressed by linguistic devices: adding a belief to a set of beliefs and checking for consequences, and asserting a conditional on the background of a set of beliefs. These actions have equal rights.

What can be said about the correlation of these actions, and what are the consequences from a pragmatical point of view? This is studied in the present paper.

2 Belief sets vs. belief bases

We may take the first argument of updates on any of three levels: as individual sentences, as finite sets of sentences (belief bases), or as arbitrary sets of sentences. (This holds for the input as well: see [FH94].) It is only in the last case that we can require sets of beliefs to be closed under logical consequences, and this was the choice made in the early times of the AGM approach.

In this paper we shall suppose that sets of belief are finite, and we shall identify such belief bases with the conjunctions of their elements.

There are at least three arguments for that. The first is a practical one: If we want to represent beliefs on a computer, we are obliged to do with finite objects.

The second argument is a philosophical one: A real map is not a complete and thus infinite description of an area, but rather the result of a reasoned balance between the requirement of being informative and that of being read-able. In the same way it is natural to think of a belief set as a finite number of sentences that roughly correspond to explicit beliefs in a mental state. [1]

A belief base is supposed to describe “the actual mental state”. As we claimed via the above map metaphore, belief sets, intended as mental states, are usually incomplete. Thus they do not completely describe all the consequences of a set of beliefs, but rather several of them that might be obtained by completing the actual belief base. (Every possible complete description coming out from the actual belief base will be called a possible world in the sequel.)

The last argument is formal: If we worked with infinite belief sets, we would be obliged to situate the update operation on the metalevel. This is discussed in more detail in the next section.

3 Language vs. metalanguage

Both for conditionals and updates we have two options: We might consider that they correspond to operations in the metalanguage, but we might as well consider them to be non-classical connectives in the object language. If the operation is on the metalevel, an analysis cannot be given in terms of axioms and inference rules (which are constructive), but must be formulated in terms of postulates, which are non-constructive and somewhat more informal. Putting the update operation in the object language allows us to identify it with a connective, to which a pragmatical meaning can be ascribed, i.e. a meaning in terms of the consequences of the updating action on the belief state.

Up to now it has mostly been considered that updates are in the metalanguage, and conditionals in the object language, and that the Ramsey test stipulates a formal link between both, in much the same way as the predicate of consistency in Gödel's work. Under this view, there is a formal result that is sometimes called Gärdenfors' impossibility theorem [Gär78, Gär88] which roughly says that there is no (object language) conditional operator that

[1] Hence logically equivalent sets of beliefs might be expressed in different ways. Note that this does not force us a priori to take position in the debate between so-called foundationalists and coherentists (because even if we work with belief bases we can still require that the update of logically equivalent belief bases with the same input leads to equivalent results). Nevertheless, we shall see that under the RT the coherentist position seems to be more adequate.

corresponds with (metalinguistic) AGM-revision operations via the Ramsey test. This negative result has led to a big discussion in the literature on belief revision and several proposals to overcome it. [2]

As we work with finite belief bases (that we identify with sentences), we can place both updates and conditionals in the object language. It has been pointed out in [Gra91] that if we do not take revision operations but update operations in the style of Katsuno and Mendelzon [KM92], the proof of Gärdenfors' impossibility theorem does not go through. This paper can be seen as going beyond Grahne's work by showing that indeed all the standard systems of conditional logic support update operations.

4 The formal view

The RT explicits that both conditionals and updates involve three entities, viz. an initial belief base B, an input A, and a resulting belief base C (or rather a consequence of it). Therefore we may say that these two operations are different ways of seeing algebraically a basic ternary relation $R(B, A, C)$ between beliefs. In fact, we have at least two possibilities to express R by means of binary operators:

- We may group B and A, and thus explicitly construct the new belief base resulting from the change. In this case, what we choose is a binary *update operator* $\hookleftarrow\!\diamond$. $B \hookleftarrow\!\diamond A$ can be read "B has been updated by A". Then to check whether $R(B, A, C)$ holds amounts to prove that $B \hookleftarrow\!\diamond A \vdash C$.

- We may group A and C. In this case, we choose a *conditional operator* $\Box\!\!\rightarrow$ which acts on an input A and a question C. $A \Box\!\!\rightarrow C$ can be read "if an update by A is performed then C follows". This may be called a hypothetical update. To check whether $R(B, A, C)$ holds amounts to prove that $B \vdash A \Box\!\!\rightarrow C$. [3]

Astonishingly enough, this duality between the two connectives has not be explored further. This might either be due to a general prejudice against

[2] In particular, taking the conditionals out of the object language (the language of beliefs) allows to avoid the impossibility theorem (see [Lev88, LR95]).

[3] Note that we may as well group B and C. This should be related to plan generation as studied in AI, but has not been investigated up to now in the literature on updates.

the translation of metalinguistic operations in the object language via modal connectives, or to a peculiar prejudice coming with Gärdenfors' impossibility result. We claim that stressing this duality we can get a better account of the notion of belief change. Let us try to make this point clearer.

Having introduced these two non-classical operators, we are now able to write the RT more formally:

$B \vdash A \Box\!\!\rightarrow C$ iff $B \leftarrow\!\!\diamond A \vdash C$.

Suppose $\vdash$ is monotonic. It follows immediately from the RT that the update operator is monotonic in the left argument, and that the conditional operator is antitonic in the right argument. [4]

In this paper, the inference relation will be monotonic, and consequently we buy the entailed monotonicity properties. [5]

5 Normal update logics

Our formal starting point is thus to state the RT as an inference rule. More precisely it corresponds to two inference rule schemas:

(R.Conv1) $\dfrac{B \leftarrow\!\!\diamond A \ \rightarrow \ C}{B \ \rightarrow \ A \Box\!\!\rightarrow C}$

(R.Conv2) $\dfrac{B \ \rightarrow \ A \Box\!\!\rightarrow C}{B \leftarrow\!\!\diamond A \ \rightarrow \ C}$

that we call the Ramsey rules of conversion.

What is the minimal logic of updates and conditionals that supports these two inference rules?

Suppose $\rightarrow$ is material implication. Surprisingly enough, we shall see that we obtain in this way all the axioms and inference rules of Chellas's basic seminormal conditional logic $\mathcal{C}$k. This makes us call this logic the basic semi-normal update logic $\mathcal{UC}$k. In detail, the following can be derived [6] from

[4] Suppose $B \vdash B'$ and $B \leftarrow\!\!\diamond A \vdash C$. By the RT, $B \vdash A \Box\!\!\rightarrow C$. Hence $B' \vdash A \Box\!\!\rightarrow C$, and again by the RT $B' \leftarrow\!\!\diamond A \vdash C$. The proof for the conditional is similar.

Note also that in any case, both operators can be monotonic in the input part only on pain of triviality.

[5] Note nevertheless that as we have both the update and the conditional operator at the object level, this leaves us open the possibility of a nonmonotonic inference relation at the metalevel.

[6] These theorems and rules can be established using the technique of [RS97]. There, axioms and inference rules have been transformed to sentences with only one non-classical

(R.Conv1) and (R.Conv2) together with the axioms and inference rules of classical propositional logic:

(RC.EC) $\dfrac{C_1 \ \leftrightarrow \ C_2}{A \Box\!\!\rightarrow C_1 \ . \leftrightarrow . \ A \Box\!\!\rightarrow C_2}$

(RC.M) $\dfrac{C_1 \ \rightarrow \ C_2}{A \Box\!\!\rightarrow C_1 \ . \rightarrow . \ A \Box\!\!\rightarrow C_2}$

(C.C) $(A \Box\!\!\rightarrow C_1) \wedge (A \Box\!\!\rightarrow C_2) \ . \rightarrow . \ A \Box\!\!\rightarrow (C_1 \wedge C_2)$

(C.N) $A \Box\!\!\rightarrow \top$

As well, the following update counterparts of the above can be derived:

(RU.EC) $\dfrac{B_1 \ \leftrightarrow \ B_2}{B_1 \leftarrow\!\!\diamond A \ . \leftrightarrow . \ B_2 \leftarrow\!\!\diamond A}$

(RU.M) $\dfrac{B_1 \ \rightarrow \ B_2}{B_1 \leftarrow\!\!\diamond A \ . \rightarrow . \ B_2 \leftarrow\!\!\diamond A}$

(U.C) $(B_1 \vee B_2) \leftarrow\!\!\diamond A \ . \rightarrow . \ (B_1 \leftarrow\!\!\diamond A) \vee (B_2 \leftarrow\!\!\diamond A)$

(U.N) $\neg(\bot \leftarrow\!\!\diamond A)$

Thus the RT gives us a principle of extensionality for belief bases, but not for inputs. Just as the basic normal conditional logic $\mathcal{CK}$ is obtained from $\mathcal{C}$k by adding a principle of extensionality for inputs, the basic normal logic of updates and conditionals $\mathcal{UCK}$ will be obtained in the same way by adding

(RC.EA) $\dfrac{A_1 \ \leftrightarrow \ A_2}{A_1 \Box\!\!\rightarrow C \ . \leftrightarrow . \ A_2 \Box\!\!\rightarrow C}$

(Note that the same logic is obtained by adding a principle of extensionality (RU.EA) for the input argument of the update operator.)

Let us briefly recall Chellas' selection function models for $\mathcal{CK}$, which will do as well for $\mathcal{UCK}$.

There is a set of possible worlds W and a selection function such that for a given world w, the set $w \cdot U$ collects all those worlds of U that are closest to w. To every sentence A there is associated the belief state $[A]$ consisting of all those possible worlds where A is true. In particular, the belief state associated to a conditional $A \Box\!\!\rightarrow C$ is the set of those possible worlds w such that C is true in all those A-worlds that are closest to w:

operator, to which the conversion rules are then applied in order to 'strip off' that operator from the relevant part of the sentence and apply classical principles. (It can also be found in [Herar]).

- $[A \Box\!\!\rightarrow C] = \{w \in W : w \cdot [A] \subseteq [C]\}$

In order to validate the Ramsey rules of conversion we need the following truth condition for the update operator:

- $[B \leftarrow\!\!\diamond A] = \{w \in W : \exists u \in B, w \in u \cdot [A]\}$

It is straightforward to establish that $\mathcal{UCK}$ is complete. [7]

What about extensions of $\mathcal{UCK}$ by standard conditional logic axioms, such as (ID) $A \Box\!\!\rightarrow A$, (MP) $A \Box\!\!\rightarrow B \,.\rightarrow.\, A \rightarrow B$, or (ASC)$(A \Box\!\!\rightarrow B) \wedge (A \Box\!\!\rightarrow C) \,.\rightarrow.\, (A \wedge B) \Box\!\!\rightarrow C$?

It is known that extensions of $\mathcal{CK}$ by any combinations of these axioms are complete. Although a general transfer result is impossible (see [Wol97]), in [Herar] completeness has been established for a large class of extensions of $\mathcal{UCK}$ by standard conditional axioms (e.g. by any combination of (CS), (CA), (ASC), (RT), (CSO), (CV)).

Equivalent axioms and rules in terms of updates can be obtained via the Ramsey rules, as done in [RS97]. In this way, completeness results for extensions of $\mathcal{UCK}$ for the update counterparts (U.ID) $B \leftarrow\!\!\diamond A \rightarrow A$, (U.MP) $B \wedge A \,.\rightarrow\, B \leftarrow\!\!\diamond A$, and (RU.ASC) $\dfrac{B \leftarrow\!\!\diamond A \rightarrow A'}{B \leftarrow\!\!\diamond (A \wedge A') \rightarrow B \leftarrow\!\!\diamond A}$ of the above conditional axioms can be obtained.

Note that all these axioms either directly correspond or are derivable from the update postulates of [KM92].

6 Conclusion

We have tried to answer the question: What does it mean to add a belief to a stock of beliefs? On the background of a pragmatical account of beliefs, we have made three methodological assumptions:

1) We start with belief bases, which are intended to correspond to mental belief states.

2) We express the operation of updating in the object language.

3) We take the Ramsey test RT as an if-and-only-if rule. Interpreted in this way, it gives us at the same time an account of the meaning of conditionals through updates, and an account of the meaning of updates through conditionals.

[7] This done in [Herar] by exploiting the connection with tense logic pointed out in [RS97].

Using RT as an inference rule and adding it to classical logic we have obtained a basic logic of updating containing Chellas' basic semi-normal conditional logic $\mathcal{C}$k. We have thus put to work the semantical analysis of conditionals to interpret changes of belief in an elegant and natural way, using all the syntactical and semantical tools at disposal.

This technical result seems to open some more general questions about a common distinction in the AI literature, viz. that betweeen updates and revisions: E.g. in [KM92] it is stated that when a belief base is notified of a change occurring in something that we could call "the external world" or "the reality", then the possible worlds in the actual belief state must be modified accordingly. "Since we are confined to our set of possibilities, we must make the change come true in all of our candidate worlds. Semantically, we change each of the possible worlds *'as little as possible'* in order to make the new state of affairs hold. Our new syntactic description of the worlds of interest should now correctly reflect the outcome of this *set* of changes. The function that maps the old description to the new is called an *update*." [Gra91]. On the contrary, if we suppose that nothing changes in the external word, and that an incoming piece of belief is supposed just to extend or correct our current belief base then the picture is different, and it is claimed that one should use revision (understood as AGM-revision of [AGM85]).

Formally, update operations as defined in [KM92] are not only different from AGM-revision operations, but even incompatible with them. That incompatibility leads to the question whether such a distinction is really acceptable. How can we decide whether an incoming information corresponds to a change in the real world? Things seem to be clear only in particular cases such as fault diagnosis or detective stories, where the real world can be supposed to be statically described by the belief base (e.g. when a murder or a fault in a circuit has occurred, and the detective or engineer revises his beliefs in the light of new information about that fixed picture). Can we presuppose that in any case we have enough information to make the right choice between AGM-revision and updates a la [KM92], i.e. between a change concerning "the real world" and a change concerning our beliefs?

This seems to be too strong an assumption. The clear account of update operation that we give taking RT as a rule of inference seems at least to ask for a more attentive analysis of the notion of revision and of the AGM assumptions from which the Gärdenfors impossibility result is drawn.

References

[AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. of Symbolic Logic*, 50:510–530, 1985.

[Bur84] John P. Burgess. Basic tense logic. In Gabbay and Günthner [GG84], pages 89–134.

[CFdCH95] Gabriella Crocco, Luis Fariñas del Cerro, and Andreas Herzig, editors. *Conditionals: from philosophy to computer science.* Number 5 in Studies in Logic and Computation. Oxford University Press, 1995.

[Che75] Brian Chellas. Basic conditional logics. *J. of Philosophical Logic*, 4:133–153, 1975.

[Che80] Brian Chellas. *Modal logic: An introduction.* Cambridge University Press, 1980.

[EG92] Thomas Eiter and Georg Gottlob. On the complexity of propositional knowledge base revision, updates, and counterfactuals. *Artificial Intelligence J.*, 57:227–270, 1992.

[FdCH97] Luis Fariñas del Cerro and Andreas Herzig. All other things being equal: on a notion of inertia in conditional logic. In Eva Ejerhed and Sten Lindström, editors, *Logic, Action and Cognition - Selection of papers of the Umea colloquium on Dynamic approaches in Logic, Language and Information (UmLLI-93)*, Trends in Logic. Kluwer Academic Publishers, 1997. (in press).

[FH94] André Fuhrmann and Sven-Ove Hansson. A survey of multiple contraction. *J. of Logic, Language and Information*, 3:39–76, 1994.

[Gär78] Peter Gärdenfors. Conditionals and changes of belief. In I. Niiniluoto and R. Tuomela, editors, *The Logic and Epistemology of Scientific Change*, volume 30, pages 381–404. Acta Philosophica Fennica, 1978.

[Gär88] Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* MIT Press, 1988.

[GG84] Dov M. Gabbay and Franz Günthner, editors. *Handbook of Philosophical Logic*, volume II. D. Reidel, Dordrecht, 1984.

[Gra91] Gösta Grahne. Updates and counterfactuals. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning (KR'91)*, pages 269–276. Morgan Kaufmann Publishers, 1991. extended version to appear in the J. of Logic and Computation.

[Han95] Sven-Ove Hansson. The emperor's new clothes: some recurring problems in the formal analysis of counterfactuals. In Crocco et al. [CFdCH95], pages 13–31.

[Herar] Andreas Herzig. Logics for belief base updating. In Didier Dubois and Henri Prade, editors, *Handbook of defeasible reasoning and uncertainty management, volume 'Belief change'.* Kluwer Academic Publishers, to appear.

[KM92] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gärdenfors, editor, *Belief revision*, pages 183–203. Cambridge University Press, 1992. (preliminary version in Allen, J.A., Fikes, R., and Sandewall, E., eds., Principles of Knowledge Representation and Reasoning: Proc. 2nd Int. Conf., pages 387–394. Morgan Kaufmann Publishers, 1991).

[Lev88] Isaac Levi. Iteration of conditionals and the Ramsey test. *Synthese*, 76:49–81, 1988.

[Lew73] David Lewis. *Counterfactuals*. Basil Blackwell, Oxford, 1973.

[LR95] Sten Lindström and Wlodek Rabinowicz. The Ramsey test revisited. In Crocco et al. [CFdCH95], pages 147–192.

[Ram31] Frank P. Ramsey. The foundations of mathematics and other logical essays. Routledge and Kegan Paul, New York, 1931.

[RS97] Mark Ryan and Pierre-Yves Schobbens. Intertranslating counterfactuals and updates. *J. of Logic, Language and Information*, 1997. special issue on TARK'96, to appear; preliminary version in: W. Wahlster (ed.), Proc. ECAI'96. John Wiley & Sons, Ltd.

[Sah88] N.-E. Sahlin. *The philosophy of F. P. Ramsey*. Cambridge University Press, 1988.

[vB95] Johan van Benthem. Temporal logic. In Dov M. Gabbay, Anthony Galton, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4, pages 241–350. Oxford University Press, 1995.

[Wol97] Frank Wolter. A counterexample in tense logic. *Notre Dame J. of Formal Logic*, 37(2), 1997. (special issue on combining logics).