Erkenn (2011) 75:223–236
DOI 10.1007/s10670-011-9292-0

ORIGINAL ARTICLE

Basic and Refined Nomic Truth Approximation by Evidence-Guided Belief Revision in AGM-Terms

Theo A. F. Kuipers

Received: 2 February 2010 / Accepted: 10 November 2010 / Published online: 22 June 2011

Abstract Straightforward theory revision, taking into account as effectively as possible the established nomic possibilities and, on their basis induced empirical laws, is conducive for (unstratified) nomic truth approximation. The question this paper asks is: is it possible to reconstruct the relevant theory revision steps, on the basis of incoming evidence, in AGM-terms? A positive answer will be given in two rounds, first for the case in which the initial theory is compatible with the established empirical laws, then for the case in which it is incompatible with at least one such a law.

1 Introduction

AGM-style belief revision (AGM-BR; for an overview, see Hansson 1999), typically aims at coherence optimization between a given set of beliefs and new information in as conservative a way as possible, implicitly taking that new information as true, whatever distance its adherents take to matters of truth. However, as far as aiming at truth approximation at all, AGM-BR seems to be primarily aiming at the truth about the actual world, *actual truth approximation*, in short. Following Grove (1988), Niiniluoto (1999) and Cevolani and Calandra (2009) have been studying the prospects of belief revision for approximation of the actual truth. New and extended attempts focusing on this aim have been made at conferences in Trieste (2009) and Amsterdam (2009) by Cevolani, Crupi and Festa, Schurz, Niiniluoto, Smets, and Zwart and Renardel.

T. A. F. Kuipers (✉)
Department of Theoretical Philosophy, University of Groningen, Groningen, The Netherlands
e-mail: T.A.F.Kuipers@rug.nl
URL: www.rug.nl/staff/t.a.f.kuipers

However, theorizing and hence theory revision in the natural sciences typically aim at *nomic truth approximation*, that is, an approximation of the truth about what is nomically (e.g. physically) possible and what is not (Kuipers 2000). Nomic truth approximation by theory revision is guided by evidence, where evidence consists of case descriptions and induced empirical laws based on them. In addition to a basic ("content") kind of nomic truth approximation, there is a refined ("likeness") kind, as a concretization of the basic one (see Zwart (2001) for the distinction between content and likeness approaches[1]). Moreover, there exist (observationally-theoretically) stratified variants of both, probabilistic variants and, in principle, all kinds of combinations.

An instructive stratified toy example is the following: let there be a complex, but finite, electric network of switches and bulbs, and a battery. Let the network of (serial and parallel) connections be hidden and the task is to find out the precise structure of this network. The observational nomic truth about the network amounts to a characterization of the physically possible state of the network, as far as positions of the switches (on/off) and the bulbs (lighting or not) are concerned. This nomic truth can be expressed by a propositional formula, of which the disjunctive normal form has as disjuncts the constituents that represent these states. It may be possible to reconstruct from this formula the theoretical nomic truth, that is, the full network, including the hidden connections. However, it may also be that there are empirically equivalent networks, that is, networks generating the same observational nomic truth.

The question asked in this paper is: can something like AGM-BR be helpful for evidence-guided theory revision aiming at (some kind of) nomic truth approximation? In other words, is it possible to reconstruct plausible theory revision steps, on the basis of characteristic evidence, aiming at nomic truth approximation in AGM-terms?

In Sect. 2 it will first be argued that straightforward basic theory revision, taking into account as effectively as possible the established nomic possibilities and the, on their basis, induced empirical laws, guarantees (unstratified) basic nomic truth approximation. Then it will be shown that this revision can be reconstructed into two AGM-steps, in arbitrary order. One of these is straightforward expansion; the other is an extreme form of contraction, viz. so-called full meet contraction. This revision needs, however, refinement for the difficult but likely case that at least one of the induced laws is incompatible with the original theory.

In Sect. 3 it will first be shown that the spheres approach of theory revision developed by Adam Grove (1988) can be used to refine the above indicated two-step theory revision such that it can be used for the hard case, and reduces to the basic case when theory and all induced laws are compatible. Assuming the proper order, that is, first a refined kind of revision in the face of the induced laws, viz. a kind of partial meet revision, and then full meet contraction in the face of the remaining counterexamples, the resulting refinement is potentially conducive for basic truth

[1] Zwart, however, disagrees about calling the second a concretization of the first.

approximation. In terms of the likeness foundation of the spheres approach by Wlodek Rabinowicz (1995), based on a four-place similarity relation, it will be shown that even this refinement is potentially conducive for refined truth approximation.

In the concluding Sect. 4 the main (positive) conclusions will be followed by a number of debunking remarks about the presented AGM-style theory revision from a realist point of view.

2 The Basic Account

2.1 Basic Definitions and Basic Theory Revision

According to the structuralist theory of truth approximation (Kuipers 2000), nomic truth approximation more specifically aims at the strongest true theory T about the set of nomic possibilities within the set of conceptual possibilities Mp generated by a chosen vocabulary for a chosen domain. Nomic truth approximation by evidence-guided *theory* revision requires definitions of 'being closer to the truth' and 'being more successful', or rather primarily their 'at least as'-versions. A theory X amounts to a specified subset of Mp with the weak claim that it is a superset of T ($T \subseteq X$) and the strong claim that it is equal to it ($T = X$), resulting from adding the claim that X is a subset of T ($X \subseteq T$). The weak claim may also be called the *necessity claim* and the extra one the *sufficiency claim*, corresponding to whether the claim states that belonging to X is necessary or sufficient for being nomically possible.

Informally we can summarize the point of departure as follows: we have a domain Mp of possibilities and every theory 'amounts to' a subset of this domain. This applies also to the strongest true theory, T. The elements of T are the 'real' possibilities, so to speak. All the possibilities outside T are not real.

The weak claim concerning a theory X is that this theory does not leave out any real possibilities. The strong claim is that it in addition does not allow for any unreal possibilities.

The (qualitative) basic definition of '***Y is at least as close to T as X***' amounts to: $Y\Delta T \subseteq X\Delta T$ (where Δ stands for symmetrical difference, i. e. $Y\Delta T = Y - T \cup T - Y$), and hence to:

(i^b) $T - Y$ is a subset of $T - X$
(ii^b) $Y - T$ is a subset of $X - T$

and '***closer to***' iff, in addition, in at least one case it is a *proper* subset.

This is the model version; there is also a consequence version and a mixed version (see Kuipers 2000, Chap. 8).

Not knowing T, we have to try to improve our guesses (theories) of what T is on the basis of, or guided by, (new) evidence. Evidence typically comes in by experimentally realizing conceptual possibilities, say R(t) up to time t. They are, of course, nomic possibilities, hence, if we have not made mistakes, R(t) is a subset of

T ($R(t) \subseteq T$), whatever T is. Neglecting mistakes and forgetfulness, R(t) is an increasing set of established nomic possibilities.

R(t) will grow in particular due to testing general hypotheses, each one claiming that all nomic possibilities satisfy it. They may have been derived from the weak claim of theory X or may have been put to the test in order to test some other theory or for still other reasons. At each point of time we may assume that one or more of them are considered to have been sufficiently established as empirical laws by inductive generalization. Let subset S(t) of Mp represent at time t the resulting strongest, induced empirical law, which amounts to the claim that S(t) is the smallest induced superset of T, whatever T is ($T \subseteq S(t)$). Neglecting mistakes and forgetfulness, S(t) is a decreasing set.

In sum: $R(t) \subseteq T \subseteq S(t)$, assuming no mistakes. From now on t will be omitted. The following definition is now plausible:

The (qualitative) basic definition of '***Y is at least as successful as X relative to R/S***' amounts to:

(i^b-sf) $R - Y$ is a subset of $R - X$

(ii^b-sf) $Y - S$ is a subset of $X - S$

and '***more successful***' by requiring in addition that in at least one case it is a *proper* subset.

The first clause can be rephrased as: all established counterexamples to Y are counterexamples to X; and the second as: all established laws (represented by supersets of S!) explained by X are explained by Y. Note that the above definition implies that a theory Y is maximally successful relative to R/S iff $R \subseteq Y \subseteq S$. For then, and only then, both $R - Y$ and $Y - S$ are empty sets, which means that Y is at least as successful as any theory X.

In general, it is crucial for the proper explication of qualitative notions of more truthlikeness and (corresponding) more successfulness or greater success to be able to prove the following theorem, with or, as in the present basic unstratified case, without further conditions:

Success Theorem: If Y is closer to T than X then Y will always be at least as successful as X and become more successful in the long run.

Proof of "Y will always be at least as successful as X". First clauses, assuming $R \subseteq T$, $R - Y \subseteq T - Y$, and by (i^b), $R - Y \subseteq T - X$. But we also have that $R - Y \subseteq R$, hence $R - Y$ is a subset of the intersection of R and $T - X$, which equals $R - X$. Second clauses, assuming $T \subseteq S$, $Y - S \subseteq Y - T$, and by (ii^b), $Y - S \subseteq X - T$. But we have also $Y - S \subseteq Mp - S$, hence $Y - S$ is a subset of the intersection of $Mp - S$ and $X - T$, which equals $X - S$. Q.e.d.

Proof sketch of "Y will …. become more successful [than X] in the long run". When (i^b) or (ii^b) can be strengthened to proper subsets, in the long run, in which R approaches T by steadily growing and S approaches T by steadily shrinking, there will be realized (hence, nomic) possibilities belonging (to T) and to Y, but not to X, or there will be laws induced that assign the status of nomic impossibilities (in $Mp - T$) to conceptual possibilities that are excluded by Y, but not by X. Of course, a straightforward proof requires precise assumptions about the way in which the experiments are going through T and how and when laws are induced. Q.e.d.

This theorem gives good reasons to abduce, under certain conditions and for the time being, that theory Y is closer to the truth than theory X when Y is persistently more successful than X, i.e., when we typically speak of *empirical progress*. Or, conversely: 'truth approximation' provides the *default-explanation* of 'empirical progress'. For the basic case the good reasons are threefold (Kuipers 2000, 162), in brief: (1) it is still possible that Y is closer to the truth than X, which would explain the persistent greater success, (2) it is impossible that X is closer to the truth than Y, (3) if neither holds, the persistent greater success so far requires a (test history) specific explanation. We can also paraphrase the overall conclusion by saying that persistent greater success is conducive for truth approximation and hence that greater success is *potentially* conducive for truth approximation.

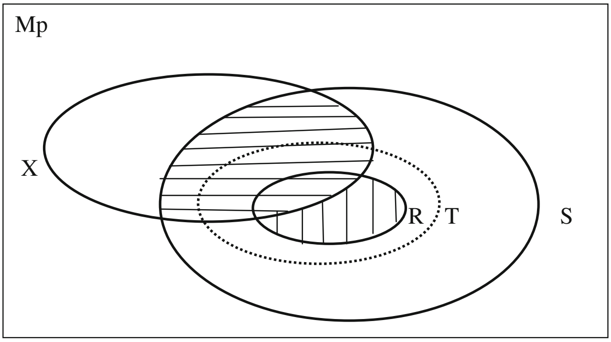
From the above definitions and the theorem we may also draw the preliminary conclusion that, assuming the data are correct, theory revision which not only realizes empirical progress but also nomic truth approximation is at least formally possible. Moreover, as is easy to check, both are even realistic, in the case of finitely many conceptual possibilities, as in the electric network, and, in general, when a finite propositional language can be used.

It is now easy to show that there is a unique way to revise a theory X in the face of evidence R/S such that the revision is, as a rule, not only more successful but even closer to the truth than X. We call this *basic theory revision of X by R/S*. The revised theory is $(X \cap S) \cup R$ or, equivalently, $(X \cap R) \cup S$ and will be indicated by $X^b_{R/S}$. Note that $X^b_{R/S}$ equals X when $R \subseteq X \subseteq S$, i.e., when X is maximally successful. Moreover, it is easy to see that $R \subseteq X^b_{R/S} \subseteq S$, i.e., that $X^b_{R/S}$ is maximally successful.

Basic Revision Theorem: Assuming correct data ($R \subseteq T \subseteq S$), '*basic theory revision of X by R/S*', resulting in $X^b_{R/S}$, guarantees that $X^b_{R/S}$ is (basically) at least as close to T, and hence at least as successful as X relative to R/S. Moreover, it is even closer to T, and more successful, than X when X is not maximally successful.

Note that the condition that X is not maximally successful amounts to the claim that R is not a subset of X or X is not a subset of S, i.e., R includes counterexamples of X or X cannot explain all laws derivable from S, in sum, X is not 'between' R and S (while the revision is!).

The validity of the theorem can easily be checked on the basis of the following picture, in which the shaded areas together indicate the revised theory $X^b_{R/S}$, the horizontal one the expansion step and the vertical one the contraction step.



Note that there are two extreme cases in which the role of X essentially vanishes and which are for that reason of special interest:

$$\text{If } X \cap S = \emptyset, \text{ then } X^{b}_{R/S} = R, \text{ hence further roles of X and S vanish}$$

$$\text{If } X \cup R = M_p, \text{ then } X^{b}_{R/S} = S, \text{ hence further roles of X and R vanish}$$

In particular the first case is of a great interest, for though extreme in some sense, it is certainly not exceptional. It simply amounts to the case in which a theory X is incompatible with at least one established law, and hence with the strongest established law. But first we will deal with the question put in this paper as far as non-extreme cases are concerned.

2.2 Basic Theory Revision in Light of AGM-Belief Revision

Now we turn to the main question of the paper: is it possible to reproduce the theory revision from X to $X^{b}_{R/S}$ by AGM-style belief revision? As is well-known, *AGM-belief revision* centers around three (partially related) operations (Alchourrón et al. 1985), see also e.g. Cevolani and Calandra (2009). A belief set, that is, a deductively closed set of sentences of a given language, is confronted with some 'input sentence' that, by minimal further changes of the original belief set, either should become a consequence or no longer be a consequence of the revised belief set. For the first case, it makes an important difference whether or not the input sentence is compatible with the belief set. In the first subcase we get so-called *expansion*, viz., the belief set is strengthened to the set of consequences of the union of the belief set and the input sentence. Regarding the input sentence it leads from suspension of judgment about that sentence to its acceptance. In the second subcase the belief set has to be adapted in a more complicated way, satisfying certain axioms. It is called *revision* (in the narrow sense). Regarding the input sentence revision leads from its rejection to its acceptance, except when the input sentence is inconsistent. Finally, in the second main case, the input sentence is supposed to belong to the belief set, but should no longer belong to the revised set. Hence, now the belief set has to be weakened in a minimal way, again in line with some axioms. It is called *contraction*. Regarding the input sentence it leads from its acceptance to suspension of judgment, except when the input sentence is logically true, in which case it remains accepted after contraction.[2]

The focus in the belief revision program has been the axiomatic characterization of the three indicated operations. Whereas this kind of explication of expansion is relatively simple, it is rather complicated for revision and contraction.

To be sure, we did not present the previous subsection in terms of sentences of a language but in terms of (sets of) conceptual possibilities or structures generated by a language. But we could translate, for example, theory X in terms of Th(X), i.e., the (deductively closed) set of sentences that are true of all structures in X. In this way the set of structures X becomes the set of models of Th(X). However, it is characteristic of the structuralist approach to identify a sentence or theory X with its set of models and to consider the set of (subsets of Mp being) supersets of X as representing the set of consequences of X. In the present context of nomic theories, this essentially model-theoretical notion of consequence is, directly be transmitted to the weak claim of a theory. To be precise, if Y is a superset of X, $X \subseteq Y$, the weak claim of theory Y, "$T \subseteq Y$", is a consequence of the weak claim of theory X, "$T \subseteq X$". Note that the strong claims of theories X and Y are incompatible as soon as Y is a *proper* superset of X.

In this way we not only get 'model versions' of (sets of) sentences and consequences, but we can also form model versions of the three operations (Hansson 1999, 220–225). For expansion this is almost trivial, for revision and contraction some extreme forms are also rather easy, precisely the ones we need in this section.

Expansion of theory X by input 'sentence' A amounts to $X \cap A$. The so-called *full meet* (fm-)revision of X by A amounts to $X \cap A$ when X is compatible with A ($X \cap A$ is non-empty) and to A when X is incompatible with A. Finally, the so-called *full meet* (fm-)contraction of X by A amounts to $X \cup cA$ when X entails A and to X when it does not. Note that fm-revision of X by A not only entails A, as informally required of revision, but also coincides with expansion of X by A when X and A are compatible and fully jumps to A when they are incompatible. In the latter case it is an extreme form of (AGM-)revision. Note also that fm-contraction of X by A no longer entails A, when X does entail A, as informally required of contraction, but that it then fully allows *all* possibilities in cA. In this sense it is an extreme form of (AGM-)contraction. Note, moreover, that it remains X when X does not entail A.

Since the AGM-operations typically deal with the consequences of the relevant belief set, it is plausible to focus our leading question first on the way in which the weak or necessity claim of a theory X has to be adapted. To obtain a nomic theory in our sense we have to finally add the sufficiency claim to the adapted version of X. Recall that we exclude in this subsection two extreme cases, of which the most important one is that S and X are incompatible.

[2] The formal definition also leaves room for the case in which the input sentence does not belong to the original belief set. Then the outcome of contraction is simply the original belief set, i.e. judgment about the input sentence was and remains suspended.

From the indicated perspective it is immediately clear that the first step in the basic revision, from X to $X \cap S$, is a clear case of (the model version of) *expansion* of X by S. Similarly, if we consider the transition from $X \cup R$ to $(X \cup R) \cap S$. Recall, for later purposes, that expansion and fm-revision of X by S coincide when they are compatible, which we are assuming.

Regarding the transition from X to $X \cup R$, or from $X \cap S$ to $(X \cap S) \cup R$, the situation is a bit more complicated. Focusing on the transition from X to $X \cup R$, that is, from X to $X \cup (R - X)$, of which $R - X$ amounts to the set of realized counterexamples of X, we see that whereas X is a subset of, and hence entails, $c(R - X)$, $X \cup R$ does no longer entail this consequence. It even allows all possibilities in $R - X$. This amounts to fm-contraction of X by $c(R - X)$. Similarly, the transition from $X \cap S$ to $(X \cap S) \cup R$ amounts to fm-contraction of $X \cap S$ by $c(R - (X \cap S))$.

In sum, we may now conclude that basic theory revision of X by R/S, leading to $(X \cap S) \cup R$ or, equivalently, $(X \cup R) \cap S$, can be seen as the successive application of expansion and fm-contraction (or vice versa), followed by adding the sufficiency claim to the resulting theory.

To be sure, we need to add the sufficiency claim, of which it is not easy to see how it can be represented in AGM-terms, starting from the present perspective on 'nomic theories'. It seems that the totally different approach by Cevolani, Crupi, and Festa (this volume) opens a new perspective avoiding this closure operation of sorts. However, that approach seems restricted to finite propositional languages and does not seem to have a clear alternative for the 'refined account' that we will soon start to motivate and to develop.

From our perspective the above analysis completes our task of an AGM-presentation of basic theory revision for the non-extreme cases in which X and S are compatible ($X \cap S \neq \emptyset$) and in which X and R do not exhaust Mp ($X \cup R \neq M_p$), respectively. In both extreme cases the role of X essentially vanishes. Whereas the second extreme case ($X \cup R = M_p$) seems rather rare, the first extreme case ($X \cap S = \emptyset$) certainly is not: it merely assumes that X is incompatible with at least one induced empirical law. Hence, the main remaining task is to refine the expansion of X by S in some way for the case in which X and S are incompatible.

One might suggest that another aspect of the transition from X to $X \cup R$ (or from $X \cap S$ to $(X \cap S) \cup R$) may require refinement. Instead of fm-contraction of X by $c(R - X)$ one might think of so-called partial meet contraction of X by $c(R - X)$, in which case not all possibilities in $R - X$ are allowed. This would require some kind of degree of trustworthiness of the various experimentally realized conceptual possibilities. A similar kind of refinement of the transition from X to $X \cap S$ arises when we would assume a degree of trustworthiness of the induced laws, in which case S is no longer taken for granted. However, these kinds of refinement, which amount to weakening of the correct data assumption, go beyond the scope of the present paper.

Finally, here, and for later purposes, it is interesting to see what would have resulted when we would have defined basic theory revision of X by R/S in terms of fm-revision of X by S followed by the relevant fm-contraction, or vice versa (in both cases, followed by adding the sufficiency claim). By the indicated alternative definition we would obtain $(X \cap S) \cup R$ when $X \cap S$ is non-empty, and $S \cup R$, that is, S, when $X \cap S$ is empty, hence deviating from our primary definition in, and only in, the second case. By the 'vice versa' definition, fm-contraction by c(R − X), followed by fm-revision by S, we would obtain $(X \cup R) \cap S$, which reduces to R when $X \cap S$ is empty. Hence, in this case the result would not differ from our primary definition. Let us call the deviating alternative definition the *fm-definition* of basic theory revision of X by R/S.

3 The Refined Account

The main problem of basic revision of X by R/S is that it reduces to R when X and S do not overlap, i.e., contradict each other. Expansion of X by S then gives the empty set, the subsequent weakening with R just amounts to R, hence a result that in no way reminds us of X. It is easy to check that the other order leads to the same result. For this route it is crucial to note that R is a subset of S. The plausible direction for refinement is to try to concretize basic revision of X by R/S in terms of a likeness approach that reduces to the basic (content) approach under the appropriate idealization conditions (IC-test), i.c. when X and S are compatible.

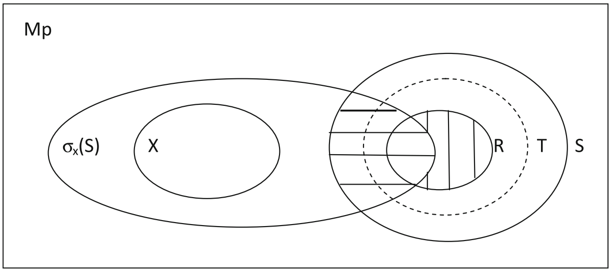
From the AGM-BR-perspective and our structuralist view of theories the spheres approach of Adam Grove (1988) is highly plausible. The spheres may seem to fall rather out of the air, but later we will see that they can be given a plausible 'similarity foundation' which, moreover, enables us to connect the spheres approach more specifically to the (structuralist) likeness approach of qualitative truth approximation.

The basic idea of Grove is to postulate nested spheres around X, satisfying a number of conditions, notably, and plausibly, that X is the smallest and Mp is the largest sphere. Consider the smallest sphere $\sigma_X(S)$ around X overlapping with S. It is now plausible to define (refined) theory revision of X by S (X^r_S) as the intersection of S and $\sigma_X(S)$, i.e. as $\sigma_X(S) \cap S$. It is easy to check (IC-test!) that, when X and S are compatible, $\sigma_X(S) = X$ and hence $X^r_S = X \cap S (= X^b_S)$. Grove has proved that X^r_S satisfies the original AGM-axioms of belief revision presented in (Alchourrón et al. 1985). This corresponds to what later has been called 'transitively relational partial meet revision' (Hansson 1999, p. 223), which we will simply abbreviate by 'pm-revision'.

In sum, $X^r_S = \sigma_X(S) \cap S$ is the most straightforward AGM-way to deal with the revision of X by S, but how to take R into account now? Recall that R is a non-empty subset of S. Recall also that the transition from X to $X \cup R$ amounted to fm-contraction of X by c(R − X) and the transition from $X \cap S$ to $(X \cap S) \cup R$ to fm-contraction of $X \cap S$ by c(R − (X ∩ S)). As Hansson (1999, 224–225) describes, contraction can also get a spheres interpretation, giving rise to partial meet (pm-)contraction. However, this would mean that we have to make a selection of members of R − X or of R − (X ∩ S), respectively. In the present context there is not much reason for this kind of refinement.

Therefore, the only question that remains is the order in which the refined revision by S and the basic revision by R should take place. According to a first alternative one can first apply fm-contraction of X by c(R − X), followed by pm-revision of the result $(X \cup R)$ by S, leading to $X^{r}_{S(R)} =_{def} (X \cup R)^{r}_{S} = (X \cup R) \cap S = X^{b}_{R/S}$, since $X \cup R$ itself is the smallest sphere around $X \cup R$ overlapping with S. Hence, $X^{r}_{S(R)} = X^{b}_{R/S}$, i.e., basic revision of X by R/S, even if $X \cap S = \emptyset$. Hence, the first alternative is no solution of the main problem.

The second alternative fares better: first pm-revision of X by S, followed by fm-contraction of the result $(\sigma_X(S) \cap S)$ by $c(R - (\sigma_X(S) \cap S))$. In this way we get: $X^{r}_{R(S)} = {}_{def} X^{r}_{S} \cup R = (\sigma_X(S) \cap S) \cup R$. Note that (IC-test) $X^{r}_{R(S)}$ reduces to $(X \cap S) \cup R = X^{b}_{R/S}$ when X and S overlap. However, $X^{r}_{R(S)} \neq X^{b}_{R/S}$ when X and S do not overlap. Hence, the order matters a lot, and we will opt for this second alternative. Of course, refined revision is only rounded off by adding the sufficiency claim. Of the suggested overlaps in the figure below, the only required overlap is that of $\sigma_X(S)$ with S, not that with T, let alone that with R. However this may be, the horizontally shaded area indicates the revision step and the vertically shaded area the contraction step.



Note that pm-revision of X by S reduces to fm-revision when there are just two spheres, viz. X and Mp. Hence, when there are just two spheres, the result of pm-revision of X by S followed by the relevant fm-contraction, and closed by adding the sufficiency claim, reduces to (the result of) the fm-definition of basic theory revision of X by R/S.

Let us now evaluate refined revision first in terms of basic truthlikeness and basic successfulness. Let us begin by the latter. It is not difficult to check that $X^{r}_{R(S)}$ is basically at least as successful as X. It is even maximally successful, for it holds that $R \subseteq X^{r}_{R(S)} \subseteq S$, hence, $X^{r}_{R(S)}$ has no established counterexamples and it explains the strongest established law, hence it explains all established laws. However, already in view of being basically at least as successful, the proposed revision is, due to the (basic) success theorem, potentially conducive for basic truth approximation, even if X is incompatible with S. But in this extreme case, the proposed revision is not basically at least as close to the truth, except in a very extreme, lucky case. The reason is that, as a rule, the revision introduces new mistakes, viz. it includes models of S outside T that did not belong to X, i.e., $\sigma_X(S) \cap (S - T)$ will be non-empty.

This is typically grist to the mill of refined truth approximation, for in that approach new mistakes are allowed as long as they are less bad than old ones. Hence, the question is how the revision fares in terms of refined truthlikeness and corresponding refined successfulness.

Refined truth approximation, as presented in (Kuipers 2000), is a qualitative likeness approach to truth approximation. It is based on a three-place 'structure-likeness' relation on the set of structures:

$$\mathrm{s}(\mathrm{x},\mathrm{y},\mathrm{z})\ \text{y is at least as similar(close) to z as x}$$

When s(x,y,z) holds, y is also said to be, qua kind of structure, *between* x and z. It is supposed to satisfy some plausible minimal (s-) conditions.[3] Moreover, we need not assume that all pairs of structures are comparable in the sense of being related by some intermediate structure. Hence we define: x and z are *related*, r(x,z), iff $\exists$ y s(x,y,z). Finally, we say that s is trivial if: for all x, y, and z $\mathrm{s}(\mathrm{x},\mathrm{y},\mathrm{z})$ iff $\mathrm{x}=\mathrm{y}=\mathrm{z}$.

Before we introduce further definitions, let us introduce the likeness foundation of spheres and indicate the connection with the likeness approach to truth approximation. Not all of Grove's sphere axioms are very plausible. Wlodek Rabinowicz (1995) provided plausible foundations in terms of a four-place similarity relation:

$$\mathrm{sim}(\mathrm{x},\mathrm{y};\mathrm{u},\mathrm{v})\ \text{x is at least as close (similar) to y as u is to v}$$

satisfying four plausible conditions and one Limit Assumption (see below). Given a set of structures X, Rabinowicz now defines a binary relation between structures

$$\mathrm{x} \le_{\mathrm{X}} \mathrm{y}\ \text{iff}\quad \forall \mathrm{y}' \in \mathrm{X}\, \exists \mathrm{x}' \in \mathrm{X}\ \mathrm{sim}(\mathrm{x}',\mathrm{x};\mathrm{y}',\mathrm{y})$$

This relation might be paraphrased by: X has at least as similar representatives of x as of y.

The relation enables the definition of a sphere (Rabinowicz 1995, p. 92):

$$\text{Y is a } \textit{sphere} \text{ around X}\quad \text{iff}\quad \begin{array}{l}\text{(i) if } \mathrm{X} \neq \emptyset \text{ then } \mathrm{Y} \neq \emptyset \\ \text{(ii) } \forall \mathrm{x} \forall \mathrm{y} \in \mathrm{Y} \text{ if } \mathrm{x} \le_{\mathrm{X}} \mathrm{y} \text{ then } \mathrm{x} \in \mathrm{Y}\end{array}$$

It is not difficult to check that this definition satisfies Grove's four axioms, among them that X and Mp are the smallest and the largest sphere, respectively.

Recall that $\sigma_{\mathrm{X}}(\mathrm{S})$ was the 'smallest' sphere around X that overlaps with S and that $\mathrm{X}^{\mathrm{r}}_{\mathrm{S}} = \sigma_{\mathrm{X}}(\mathrm{S}) \cap \mathrm{S}$ was defined as the refined revision of X by S. Rabinowicz proved that $\mathrm{X}^{\mathrm{r}}_{\mathrm{S}} = \{\mathrm{x}' \in \mathrm{S} \mid \exists \mathrm{x} \in \mathrm{X}\ \forall \mathrm{y} \in \mathrm{X}\ \forall \mathrm{y}' \in \mathrm{S}\ \mathrm{sim}(\mathrm{x}',\mathrm{x};\mathrm{y}',\mathrm{y})\}$, where the latter set corresponds to Rabinowicz' version of $\mathrm{X}^{\mathrm{r}}_{\mathrm{S}}$. The idea behind this version is that it forms "the set of S-worlds that are as similar to some worlds in X as possible, as compared with other worlds in X" (Rabinowicz (1995, p. 82, S substituted for Y). The Limit Assumption that now is needed instead of a, here not presented, very

[3] They are: centered, centering and conditionally left and right reflexive. Here s is centered iff s(x,x,x) and centering iff s(x,y,x) implies $\mathrm{x}=\mathrm{y}$. s is conditionally left/right reflexive if s(x,y,z) implies all kinds of left and right reflexivity, i.e., s(x,x,y), s(x,x,z), s(y,y,z) and s(x,y,y), s(x,z,z), s(y,z,z), respectively. Note that this conditional form leaves room for incomparable structures (see text), which otherwise would not be the case.

arbitrary assumption of Grove is not at all that arbitrary: if X and S are non-empty then X^r_S is non-empty.

Now we can turn to the connection between s and sim. Assuming that z in s(x,y,z) is a kind of target the most plausible one certainly is:

$$s(x, y, z) \text{ iff } sim\,(y, z; x, z) \quad \text{i.e., } y \text{ is at least as similar to } z \text{ as } x \text{ (is to } z)$$

With this connection in mind we now arrive at the crucial definition of refined truth approximation.

Definition: ***Y is refined at least as truthlike as X*** iff

(i^r) $\forall x \in X\ z \in T\ r(x,z) \rightarrow \exists y \in Y\ s(x,y,z)$
(ii^r) $\forall y \in Y - (X \cup T)\ \exists x \in X - T\ \exists z \in T - X\ s(x,y,z)$

It is easy to check that (i^r) is a strengthening of (i^b) of the basic definition and that (ii^r) is a weakening of (ii^b). (i^r) roughly says that every comparable pair of structures, one of X and one of T, has an 'intermediate' in Y. (ii^r) states that if Y − (X ∪ T) is at all non-empty, which is excluded in the basic case, these structures are 'useful'. The definition reduces to the basic one when s is trivial.

Whereas the basic revision $X^r_{R(S)}$ was easily seen to be basically at least as truthlike as X, the refined revision $X^r_{R(S)}$ is now not necessarily at least as truthlike as X in the refined sense. Hence, there is now even more reason to turn to successfulness.

Definition: ***Y is refined at least as successful as X, relative to R/S***, iff

(i^r-sf) $\forall x \in X\ z \in R\ r(x,z) \rightarrow \exists y \in Y\ s(x,y,z)$
(ii^r-sf) $\forall y \in Y - (X \cup S)\ \exists x \in X - S\ \exists z \in S - X\ s(x,y,z)$

The Refined Success Theorem tells now that, assuming correct data, 'refined at least as truthlike' entails 'refined at least as successful'. Again the proof is not difficult. However, for the general proof of (ii^r)'s entailment of (ii^r-sf) we need to assume that, if Y − (X ∪ S) is non-empty, S is convex (i.e., if x, z ∈ S and s(x,y,z), that is, when y is qua kind of structure between x and z, then y ∈ S). Similar to the basic case, the consequence of the theorem is that being persistently more successful in the refined sense is conducive for refined truth approximation (provided S is convex, if relevant).

The final crucial question now is whether the (AGM-interpretable) refined revision $X^r_{R(S)}$ of X by R/S is at least as successful as X in the refined sense. In that case it would be potentially conducive for truth approximation for it may become persistently more successful in the refined sense and hence conducive for refined truth approximation. This happens to be the case according to the following:

Main Theorem: $X^r_{R(S)}$ is refined at least as successful as X, relative to R/S.

Let us look at the specific claims:

$$\left(i^r\text{-sf-wrt } X^r_{R(S)}\right) \quad \forall x \in X\, z \in R\, r(x, z) \rightarrow \exists y \in X^r_{R(S)} s(x, y, z)$$

This is trivial, for R is a subset of $X^r_{R(S)}$ and r(x,z) → s(x,z,z) is a (plausible) minimal s-condition.

$$\left(\mathrm{ii^r\text{-}sf\text{-}wrtX^r_{R(S)}}\right) \quad \forall y \in X^r_{R(S)} - (X \cup S)\exists x \in X - S\exists z \in S - X\, s(x, y, z)$$

This is also trivial, for $X^r_{R(S)}$ is a subset of S, hence $X^r_{R(S)}$ − (X∪S) is empty. The latter fact has even the consequence that the convexity of S is not required for the applicability of the Refined Success Theorem.

4 Conclusions

The main conclusions of this paper are:

First, basic revision of theory X in light of evidence R/S, assuming X and S compatible, based on expansion by S, leading to X ∩ S, followed by fm-contraction by c(R − (X ∩ S)), leading to (X ∩ S) ∪ R, and closed by adding the sufficiency claim, is basically at least as successful as X and even basically at least as close to X in the nomic sense.

Second, refined theory revision in light of evidence R/S, assuming X and S incompatible, based on pm-revision by S, along Grove-Rabinowicz lines, leading to $\sigma_X(S) \cap S$, followed by fm-contraction by c(R − ($\sigma_X(S) \cap S$)), leading to ($\sigma_X(S) \cap S$) ∪ R, and closed by adding the sufficiency claim, is at least as successful as X in the refined sense, and hence potentially conducive for refined nomic truth approximation.

At this point a number of debunking remarks are in order:

(a) Having to focus in both cases first on the necessity claim and to add at last the sufficiency claim is not very elegant.
(b) Both revisions are rather ad hoc. However, as in general for ad hoc changes in a theory, the crucial question is whether they can be put to new (HD-) tests, and this is evidently the case. After all, it could even be the case that all further tests indicate that no new ad hoc maneuvers have to be made.
(c) Both revisions are rather diehard empiricist or instrumentalist. The 'instrument' X is precisely so adapted that it just saves the phenomena, not only with respect to R but also with respect to S. Note that this character will not change by weakening the correct data assumption, as suggested at the end of Sect. 2.
(d) If there is something like well-formed theories, there do not seem to be good reasons to expect that the two revisions will satisfy the criteria, even if R and S satisfy some derived criteria.
(e) Last, but not least, what remains of the idea behind X? A proper theory, even if it is without theoretical terms, in some sophisticated sense, is usually based on one or two ideas. It is difficult to imagine that such ideas do not become 'mutilated' by the revision.

Be this as it may, the two results may stimulate the interaction between truth approximation and belief revision approaches for they fundamentally show that AGM belief revision provides means for nomic truth approximation.

Acknowledgments The author wishes to thank Roberto Festa for the occasion to present this paper in 2009 at a symposium in Trieste, Gustavo Cevolani, Gerhard Schurz and Wlodek Rabinowicz for very useful comments, and the *Netherlands Institute for Advanced Study* (NIAS), Wassenaar, for providing again paradisiacal conditions for research and writing, including linguistic correction by Anne Simpson.

References

Alchourrón, C., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory revision: partial meet contraction and revision functions. *Journal of Symbol Logic*, *50*(2), 510–530.

Cevolani, G., & Calandra, F. (2009). Approaching the truth via belief change in propositional languages. In M. Suárez, M. Dorato, & M. Rèdei (Eds.), *EPSA epistemology and methodology of science. Launch of the European philosophy of science association, Ch. 5*. Berlin: Springer.

Cevolani, G., Crupi, V., & Festa, R. (this volume). Verisimilitude and belief change for conjunctive theories.

Grove, A. (1988). Two modelings for theory change. *Journal of Philosophical Logic, 17*, 157–170.

Hansson, S. O. (1999). *A textbook of belief dynamics*. Dordrecht: Kluwer AP.

Kuipers, T. (2000). *From instrumentalism to constructive realism. On some relations between confirmation, empirical progress, and truth approximation, Synthese Library 287*. Dordrecht: Kluwer AP.

Niiniluoto, I. (1999). Belief revision and truthlikeness. In: *Spinning ideas, electronic essays dedicated to Peter Gärdenfors on his fiftieth birthday.*

Rabinowicz, W. (1995). Global belief revision based on similarities between possible worlds. In Hansson and Rabinowicz (Eds.), *Logic for a change. Essays dedicated to Sten Lindström on the occasion of his fiftieth birthday* (pp. 80–105). Uppsala Prints and Preprints in Philosophy 1995–1999, Department of Philosophy, Uppsala University.

Zwart, S. D. (2001). *Refined verisimilitude. Synthese Library 307*. Dordrecht: Kluwer AP.