Theories Looking for Domains
Fact or Fiction?
Reversing Structuralist Truth Approximation

Theo Kuipers

ABSTRACT. The structuralist theory of truth approximation essentially deals with truth approximation by theory revision for a fixed domain. However, variable domains can also be taken into account, where the main changes concern domain extensions and restrictions. In this paper I will present a coherent set of definitions of “more truth-likeness”, “empirical progress” and “truth approximation” due to a revision of the domain of intended applications. This set of definitions seems to be the natural counterpart of the basic definitions of similar notions as far as theory revision is concerned. The formal aspects of theory revision strongly suggest an analogy between truth approximation and design research, for example, drug research. Whereas a new drug may be better for a certain disease than an old one, a certain drug may be better for another disease than for the original target disease, a phenomenon which was nicely captured by the title of a study by Rein Vos [1991]: Drugs Looking for Diseases. Similarly, truth approximation may not only take the shape of theory revision but also of domain revision, naturally suggesting the phenomenon of “Theories looking for domains”. However, whereas Vos documented his title with a number of examples, so far, apart from plausible cases of “truth accumulation by domain extension”, I did not find clear-cut empirical instantiations of the analogy, only, as such, very interesting, non-empirical examples.

1 Introduction

This paper starts by recapitulating the structuralist theory of truth approximation, as developed in [Kuipers, 2000], and then elaborates a point that was already made explicit by Sjoerd Zwart [1998]. In [Kuipers, 2000, p. 207] I wrote, in concluding the second chapter on truth approximation (chapter eight):

Finally, variable domains can also be taken into account, where the main changes concern extensions and restrictions. We will not study this issue, but see [Zwart, 1998, chapters two to four],

Lorenzo Magnani, editor, Model–Based Reasoning in Science and Engineering, pp. 1–18
© 2006, T. Kuipers
for some illuminating elaborations in this connection, among other things, the way in which strengthening/weakening of a theory and extending/reducing its domain interact.

More specifically, in this paper I will present a coherent set of definitions of “more truthlikeness”, “empirical progress” and “truth approximation” due to a revision of the domain of intended applications. This set of definitions seems to be the natural counterpart of the basic definitions of similar notions in [Kuipers, 2000] as far as theory revision is concerned.

The formal aspects of theory revision strongly suggest an analogy between truth approximation and design research, for example drug research. Whereas a new drug may be better for a certain disease than an old one, a certain drug may be better for another disease than the original target disease, a phenomenon which was nicely captured by the title of a study by Rein Vos [1991]: Drugs Looking for Diseases. Similarly, truth approximation may not only take the shape of theory revision but also of domain revision, naturally suggesting the phenomenon of “Theories looking for domains”. However, whereas Vos documented his title with a number of examples, so far, apart from plausible cases of “truth accumulation by domain extension”, I did not find clear-cut empirical instantiations of the analogy, only, as such, very interesting, non-empirical examples.

Section two briefly states the structuralist theory of truth approximation by theory revision for a fixed domain and concludes with the formal analogy with design research. Suggested by this analogy, the corresponding theory of “truth approximation by domain revision for a fixed theory” is presented in section three, which ends with a synthesis of truth approximation by theory and/or domain revision. As a kind of meta- and self-application of the “Theories looking for domains” theme, section four is devoted to the search for historical cases exemplifying this theory. I conclude with arguing that it is plausible to expect that truth approximation by domain revision in the empirical sciences frequently occur, surprisingly enough, in particular as far as the instrumentalist method is practiced.

2 Truth approximation by theory revision

My favorite, qualitative theory of truth approximation is best represented within the structuralist theory of theory representation. Surprisingly enough, and if I am not mistaken, the latter theory of representation has not been presented in earlier MBR-conferences as a pre-eminent case of model-based reasoning in the sense that not the sentences but the models of a theory form the primary domain of discourse.

1With the kind help of Lorenzo Magnani, I only found a quasi-relevant reference by Ronald Giere [1999] in his MBR98-contribution, entitled “Using models to represent
Starting from a fixed vocabulary, and a suitable similarity type of structures, usually provided by a research program, let $M_p$ indicate the set of structures of that type, also called the potential models of the theory. Let the subset $M$ of $M_p$ indicate the set of models of the theory. Finally, assuming that our target is a fixed domain of physically or, more broadly, nomically possible phenomena, let $I$ indicate this domain “as seen through $M_p$”, hence a subset of $M_p$, and be called the set of (intended) nomic possibilities or the domain of intended applications. Note that we do not yet suppose to dispose of a general characterization of $I$ as a subset of $M_p$. What we only know is that each intended application can be represented as a potential model. A general characterization of $I$ is “the great unknown” where theories, as represented by their models, are looking for. More formally, a theory is a triple of the form $<M_p, I>$, together with the weak claim that $I$ is a subset of $M$ ($I \subseteq M$) or the strong claim that $I = M$. A theory is said to be true (false) in the weak sense when the weak claim is true (false). It is easy to check that (a general characterization of) $I$ represents the strongest true weak claim, and may be called “the truth” in this context, that is, the truth about the given domain in the given vocabulary.

Now it is plausible to define what it means, for fixed $<M_p, I>$, and hence theories of the same research program, that one theory $M_2$ is closer (or more similar) to the truth than another $M_1$. Intuitively, when $M_2$ is moving from $M_1$ in the direction of $I$. Formally, when $M_2-I \subseteq M_1-I$ ($\emptyset$-area empty in Figure 1) and $I-M_2 \subseteq I-M_1$ ($\emptyset$-area empty) and at least once it should be a proper subset relation (at least one $*$-area non-empty). In terms of symmetric differences ($\Delta B = \Delta B$ ($A-B) \cup (B-A))$, $M_2 \Delta I$ should be a proper subset of $M_1 \Delta I$. In [Kuipers, 2000] I have argued that, among other formulations, the two clauses of this definition amount to: (relative to $I$) $M_2$ has more true consequences than $M_1$ and $M_2$ has more correct models than $M_1$, respectively.

Knowing what it means that one theory is closer to the truth than another is one thing, in view of the fact that we don’t know the truth, judging that this is in fact the case, in the light of our evidence is another. Although there is no theorem guaranteeing that more-successfulness entails closer-to-the-truth, there is a theorem, the Success Theorem, guaranteeing almost the reverse entailment, viz., closer-to-the-truth entails at-least-as-successfulness.

Giere pays some attention to the pioneering work of Patrick Suppes in this respect. However, Suppes’ equally pioneering work with respect to “Using models to represent theories” is not touched upon. This work of Suppes was the starting point of Joseph Sneed and later followed by Stegmüller, Balzer, Moulines, and others, and called the structuralist representation of theories, i.e. theories as sets of structures, forming their models.
Let R represent the set of (documented) realized applications at a certain time, that is, the experimentally or otherwise realized nomic possibilities at that time. Let S indicate the strongest law induced on the basis of R. Of course, R will partly be the result of testing hypothetical laws. If no mistakes have been made in representing the realized possibilities, R is not only a subset of Mp, but even of I, for nomic impossibilities, by definition, can’t be realized. Moreover, in the structuralist representation S is also a subset of Mp, such that it has been concluded, at least provisionally, that conceptual possibilities outside S, that is, in Mp-S, are nomically impossible. Finally, if our inductive jump from R to S is correct, S has to be a superset of I. To sum up: $R \subseteq I \subseteq S \subseteq Mp$.

Let us call R/S the data set. We define what it means that M2 is, relative to the data set R/S, more successful than M1 as follows: $M1 \cap R \subseteq M2 \cap R$ and $S \cup M2 \subseteq S \cup M1$. The first condition amounts to “no loss of established examples” and the second to “no loss of explained established laws”.

Now the Success Theorem states, assuming that R/S is correct, that is, $R \subseteq I \subseteq S$, then M2 is at least as successful relative to R/S as M1 if M2 is closer to the truth. This theorem is such that persistently being-more-successful is functional or instrumental for truth approximation, though not guaranteeing it, in the sense that it is very difficult/special for M1 to be at least as successful as, let alone more successful than, M2 when M2 is not closer to the truth than M1. More in detail we may argue as follows. Let M2 be more successful than M1 relative to R/S. This suggests the Comparative Success Hypothesis, according to which it is hypothesized that this will remain the case, whatever experiments we design and perform.
Testing of this hypothesis may result in the conclusion, at least for the time being, that this is in fact the case: M2 persistently remains more successful than M1. This is the paradigm situation for speaking of empirical progress, and for applying the ‘instrumentalist rule of success’, viz. replace, for the time being, M1 in favor of M2. Under this condition it is very hard to imagine that M2 is, despite appearances, not closer to the truth than M1. To be sure, if this would nevertheless be the case, we have not been creative enough in designing new experiments. In other words, the situation even allows the tentative conclusion that M2 is in fact closer to the truth.

So far I have presented the naïve or basic structuralist theory of truth approximation. It is basic in the sense that several idealizations have been made, requiring concretization. In Kuipers [Forthcoming] I have extensively illustrated the (philosophical) method of concept explication by idealization and concretization by the example of truth approximation. Let us review the main concretizations as they have been elaborated or indicated in [Kuipers, 2000].

Above we did not make a distinction between (relatively) observational and theoretical terms. In chapter nine I have spelled out how this distinction works out, with the main conclusion that empirical progress, of course, in observational terms, remains functional for truth approximation on the theoretical level, though with some greater risk of being wrong in this.

Another idealization was that we implicitly assumed that any established counterexample of a theory, that is, a realized possibility not being a model of the theory, is as bad for one theory as for any other. However, in chapter ten I have introduced the idea that one structure may be more similar to another than a third. In this way, the possibility arises that a counterexample is less dramatic for one theory than for another, because the former has a model that is more similar to the “countermodel” (i.e., the potential model representing the counterexample), than any model of the other theory. Adapted definitions of more-successfulness and closer-to-the-truth lead again to the conclusion that empirical progress is functional for truth approximation. As a matter of fact we only know of real life scientific examples of (potential) truth approximation, e.g. the Law of Van der Waals as successor of the ideal gas law, when this concretization is introduced. Without this, we only know of toy examples.

Finally, in [Kuipers, 2000, chapter thirteen] I have indicated that it is possible to introduce a domain vocabulary, as a subvocabulary of the (observational) vocabulary, such that it is possible to define the domain explicitly in that vocabulary, leaving the further behavior of the applications as the unknown to be specified. The advantage of this is that domains become comparable sets, that is, it enables us to say that an intended application
of one theory is or is not an intended application of another. This will be an important, though implicit, assumption in dealing with the main topic of this paper. But first I will introduce the analogy with design research in order to make that topic a plausible subject of theoretical and (meta-)empirical research.

**Analogy with design research**

The formal pattern of (the development of) design research, as presented in [Kuipers, 2001, chapter ten], is to some extent analogous to that of (descriptive or explanatory) “nomological” research. Starting from a set of relevant characteristics, let us call the subset of characteristics of the intended product (or process) its *intended profile*. As soon as we have a prototype we can delineate the subset of its actual characteristics, the *prototype profile*. A second prototype amounts to genuine progress when the symmetric difference between its profile and the intended profile is a proper subset of that between the profile of the first and the intended profile. In this way, for example, improving a (medical) drug for a given disease becomes formally analogous with improving a theory for a specific domain. See [Forbus and Gentner, 1997] for another illustration of progress in design research, viz., designing “qualitative simulators” to be evaluated in terms of useful and problematic properties derived from qualitative reasoning.

Rein Vos [1991], who elaborated the formal description of progress in design research for the case of drugs, entitled his book *Drugs Looking for Diseases*, because by historical research he came to the conclusion that pharmaceutical research often goes in the other direction. That is, there may be an interesting chemical substance that does not do what it is supposed to do for a given disease, even after some manipulation. However, instead of discarding the substance for this reason, one starts to look for other diseases, whether or not related to the first one, for which the substance may have curing effects. In this way, for example, bêta blockers have been discovered as a drug for certain heart diseases.

Hence the plausible question arises whether in the context of nomological research something similar, viz., “theories looking for domains” or, equivalently, “models looking for applications”, does also exist. One type of example is well known, viz., the case that a theory which is or is not successful for a certain domain turns out to be (also) successful for a totally unrelated domain. In this case we have a purely formal analogy. A famous example is Keynes’ transfer of hydraulic modeling from fluid mechanics to macroeconomics. Our interest in this paper concerns however cases in which the new and old domains are to some extent related, that is, genuine domain revision.
3 Truth approximation by domain revision

It is plausible to define what it means that, for fixed \( <M, M> \), (the truth about) domain \( I_2 \) is closer (or more similar) to \( M \) than (the truth about) domain \( I_1 \). Intuitively, when \( I_2 - M \subseteq I_1 - M \) (\( \emptyset_2 \)-area empty in Figure 2) and \( M - I_2 \subseteq M - I_1 \) (\( \emptyset_1 \)-area empty) and at least once it should be a proper subset relation (at least one *-area non-empty). Note that the *-areas differ from those in the case of theory revision, since \( I_2 \) is now moving from \( I_1 \) in the direction of \( M \), whereas \( M_2 \) was moving from \( M_1 \) in the direction of \( I \). In terms of symmetric differences, \( M \Delta I_2 \) should be a proper subset of \( M \Delta I_1 \). Now it is not difficult to argue that this definition amounts to: \( M \) has more true consequences and more correct models relative to \( I_2 \) than relative to \( I_1 \).

![Figure 2. (The truth about) domain \( I_2 \) is closer to theory \( M \) than (the truth about) domain \( I_1 \).](image.png)

Empirical progress by domain revision becomes a bit more complicated because we have to compare two data sets, relative to \( I_1 \) (\( R_1/S_1 \)) and to \( I_2 \) (\( R_2/S_2 \)). \( M \) is at least as successful relative to \( R_2/S_2 \) as to \( R_1/S_1 \) if and only if \( M \cap R_1 \subseteq M \cap R_2 \) & \( S_2 \cup M \subseteq S_1 \cup M \). The first condition now amounts to the claim that \( M \) does not loose established examples of \( I_1 \) by going from \( R_1 \) to \( R_2 \); and the second condition amounts to the claim that \( M \) does not loose explained established laws with respect to \( I_1 \) by going from \( S_1 \) to \( S_2 \).

Assuming again correct data, that is, \( R_1 \subseteq I_1 \subseteq S_1 \) and \( R_2 \subseteq I_2 \subseteq S_2 \), it is now easy to derive the Success Theorem for Domain Revision: “\( M \) is closer to \( I_2 \) than to \( I_1 \)” entails “\( M \) is at least as successful relative to \( R_2/S_2 \) as to \( R_1/S_1 \)”.

Starting from one set of realized possibilities \( R \) and one strongest estab-
lished empirical law $S$, that is, when they are not yet specified relative to the
two domains, we may also define these specifications for $i = 1, 2$ as follows:
$R_i = d_f R \cap I_i$ and $S_i = d_f S \cup I_i$, which even guarantees the condition $R_i \subseteq I_i \subseteq S_i$.

The present *Success Theorem* is again such that *persistently* being-more-
successful relative to one domain than to another is functional or instrumen-
tal for truth approximation. Let $M$ be more successful relative to $R_2/S_2$
than to $R_1/S_1$. This suggests the *Comparative Success Hypothesis*, accord-
ing to which it is hypothesized that this will remain the case, whatever
experiments we design and perform. Testing of this hypothesis may result
in the conclusion, at least for the time being, that this in fact is the case:
$M$ persistently remains more successful relative to $R_2/S_2$ than to $R_1/S_1$.
Whether we should call the relevant domain revision a case of making em-
pirical progress may be disputed. However, in the indicated situation it is
very hard to imagine that $I_2$ is, despite appearances, not closer to $M$ than
$I_1$, and this certainly is progress in knowledge of a kind, again guided by
the rule of success.

**Synthesis**

Again we have restricted ourselves to the naïve or basic theory of truth
approximation by domain revision, containing several idealizations. Similar
concretizations as in the case of theory revision can be made. However, we
will turn to a plausible formal synthesis of basic truth approximation by
theory and domain revision, being a generalization of both.

As mentioned before, the basic definitions of truth approximation by
theory revision and domain revision, respectively, amount to comparisons
of symmetric differences. Both can be generalized qualitatively, leading to
the same result:

*Definition:* $<M_2, I_2>$ *fits better than* $<M_1, I_1>$ if and only if $M_2 \Delta I_2 \subset M_1 \Delta I_1$.

This leads to four special cases, viz., when $M^* \subset M$ and $I^* \subset I$ (as depicted
in Figure 3) then there is “better fit” if and only if, for some $i, j = 1, 2, 3, 4$, the
$i$- and $j$-area are empty ($i \& j E$) and the $k$- and/or $l$-area are non-empty
($k \lor l N$), as follows:

**I** $<M^*, I^* >$ is a double specialization of $<M, I>$ iff $3 \& 4 E / 1 v 2 N$

**II** $<M, I>$ is a double generalization of $<M^*, I^*>$ iff $1 \& 2 E / 3 v 4 N$

**III** $<M^*, I>$ is a double strengthening of $<M, I^*>$ iff $1 \& 3 E / 2 v 4 N$

**IV** $<M, I^*>$ is a double weakening of $<M^*, I>$ iff $2 \& 4 E / 1 v 3 N$
We obtain some interesting “special-special cases” of “better fit” if we assume in addition that I* = I or M* = M. Let us start with the first possibility, depicted in Figure 4. When M* ⊂ M and I* = I there is “better fit” if and only if the i-area is empty (iE) and the j-area is non-empty (jN), as follows:

I/III ⇒ <M*,I> is a theory (specialization or) strengthening of <M,I> iff 3E/2N

II/IV ⇒ <M,I> is a theory (generalization or) weakening of <M*,I> iff 2E/3N

where the first condition amounts to: “not fewer correct models and fewer incorrect models” and the second to “no new incorrect models and some new correct models”.

The second possibility, viz., M* = M and I* ⊂ I, leads to the following “special special cases”, depicted in Figure 5. In this case there is “better
fit” if and only if the i-area is empty (iE) and the j-area is non-empty (jN), as follows:

\[
\begin{align*}
I/IV & \Rightarrow <M,I*> \text{ is a domain restriction of } <M,I> \text{ iff } 4E/1N \\
II/III & \Rightarrow <M,I> \text{ is a domain extension of } <M,I*> \text{ iff } 1E/4N
\end{align*}
\]

where the first condition amounts to “not fewer examples and fewer counterexamples” and the second to “no new counterexamples and some new examples”. In the first case we might also speak of “domain specialization” or “domain weakening” but I prefer to speak of “domain restriction”. Similarly for the second case. Although one might speak of “domain generalization” or “domain strengthening”, I prefer to speak of “domain extension”.

![Figure 5. Truth approximation by domain restriction or extension.](image)

4 In search of applications of “Truth approximation by domain revision”

So far we have shown that it is perfectly possible to explicate the idea of “Truth approximation by domain revision”, and hence of the possibility of “Theories looking for domains”. Now the question arises, as a kind of meta- and self-application of the “Theories looking for domains” theme, whether there are historical cases exemplifying this possibility.

Apart from a plausible type of cases the “predicted” phenomenon may seem rare in the empirical sciences. Of course, there are well-known examples of theories of which the formal core, i.e., \( <M_p,M> \), finds a totally new area of applications. It may be that the new type of applications behave in relevant respects analogously to the originally intended ones, in which case we have a kind of domain extension, say, domain extension by analogy.

However, and here I depart from the Drugs Looking for Diseases perspective, in which it doesn’t really matter whether the diseases are or are
Theories Looking for Domains

not related, my main interest is in truth approximation by genuine domain revision in the sense of domain manipulation within an area of, in some sense, related empirical phenomena, that is, the vocabulary constituting Mp should have the same kind of interpretation. Given this assumption, we may distinguish between basic (or set-theoretic) kinds of domain revision and refined kinds.

Of the basic kinds we can distinguish domain restriction, domain extension, and their combinations: domain restriction followed by domain extension or vice versa. Note that speaking of “truth approximation by domain extension from I* to I”, in accordance with our definition, viz., I* is a proper subset of I, such that I-I* is a (proper) subset M (the 1-area empty and the 4-area non-empty in Figure 5), may sound somewhat strange in the case that the weak claim that I* is a subset of M (I* ⊆ M) is true and hence the weak (but stronger) claim that I is a subset of M is true as well. For this amounts to the case that we correctly extend the domain for obtaining a true weak claim without having to correct the original domain. Let us call this “truth accumulation by domain extension” (with the counterpart of “truth accumulation by theory strengthening”). Historical cases of this type are also well known: successful extensions of the domain with respect to which the theory was already (assumed to be) successfully applied. The successful search for new domains of application of Newton’s theory, by Newton and his followers, belongs to this type²³.

Of course, later findings may relativize the correctness of the original and/or later claims. In this case one option is theory revision, as in the case of the discovered deviations from Newton’s predictions, leading to Einstein’s special and general theory of relativity. The special theory is, relative to Newton’s theory, a typical case of “truth approximation by (idealization followed by a, to be sure, fundamental) concretization of the theory”. In order to define this type of truth approximation by theory revision I have developed in part three of [Kuipers, 2000] a general theory of “refined truth approximation by theory revision”, which was indicated in section two. As a side remark, in contrast to Rivadulla’s general claim [Rivadulla, 2006],

²It is not always unambiguous whether we have a case of pure domain extension or a case of domain extension by analogy. For example, from the point of view of Newton’s theory of gravitation the Rutherford-Bohr planetary model of the atom is a kind of domain extension by analogy with the planetary system, since the Coulomb forces behave analogously to the gravitational forces. However, looked from Newton’s general theory of motion we have a proper case of pure domain extension.

³Of course, starting from an uninteresting (purported) truth, domain extension will be equally uninteresting. In other words, it is perfectly possible to let the search for domain revision be guided by “interesting (purported) truths”, which reflects, apart from the terminology, a typically instrumentalist perspective. It may well lead to domain restriction, rather than domain extension.
for refined truth approximation it is not at all relevant that the theories concerned are incompatible, e.g., those of Newton and Einstein, of Copernicus and Kepler, and of Boyle/Gay-Lussac (ideal gas law) and Van der Waals. Unfortunately, it needs some preparation to explain this possibility of refined truth approximation in detail. The reader is referred to [Kuipers, 2000, chapter ten, pp. 268–271], where the gas example is presented in detail.

Besides concretization of the theory, when later findings relativize the correctness of the original and/or later claims related to “truth accumulation by domain extension”, another option is to concretize the (specification of the) domain. More generally, it is plausible that also a general form of “refined truth approximation by domain revision” can be defined with “truth approximation by (idealization followed by) concretization” as a special subtype.

Hence, the remaining question is whether there are historical examples of truth approximation by basic or refined kinds of domain revision, apart from cases of (putative) “truth accumulation by domain extension”. As far as the basic kind is concerned it may be not surprising if there are none. Also in the case of “truth approximation by theory revision” basic kinds are either toy examples or examples of “truth approximation by theory strengthening”, that is, the counterpart of “truth accumulation by domain extension”. However, so far, I also did not find convincing refined cases, but I think this is a matter of insufficient knowledge of the history of science. Moreover, from the structuralist perspective it is tempting to focus on empirical theories of a mathematical nature, while it may be easier to find cases in areas where this is less typically the case, e.g., biology and psychology, because the target phenomenon of domain revision may occur in these areas at least as frequently.

It is easier to find related types of non-empirical cases. Let us begin with cases of “truth accumulation by domain extension”. Mathematics provides evident cases. For example, every extension of the set of natural numbers for which Goldbach’s conjecture has been proved belongs to it. Similarly, even before the full proof of the four-color problem, every new type of map for which the four or, for that matter, the five color problem was positively solved, provides another illustration. Mathematical economics also provides beautiful illustrations. As Bert Hamminga [1983] has documented, in the development of the theory of international trade new results frequently amounted to an extension of the proven domain of validity of a certain “interesting theorem”, such as the “factor prize equalization theorem”; for a brief further indication see [Kuipers, 2001, section 1.2.7]. By the way, this type of research, though not empirical as such, can be motivated in terms of
truth approximation with respect to the actual world, because any extension of the proven domain of validity increases the chance that the theorem also holds in the actual world.

I do not yet know of basic cases of “truth approximation by domain revision” in mathematical economics other than the cumulative type. However, (mixed) refined cases, based on distances between structures in Mp, have been studied. In particular, “truth approximation by (idealization followed by) concretization of the domain and the theory”, where the theory is an interesting theorem in the sense indicated above. In Cools et al. [1994], see also [Kuipers, 2000, chapter eleven, section two], important developments in the theory of the capital structure of firms have been reconstructed along these lines, more specifically, the transition from the theory of Modigliani and Miller to that of Kraus and Litzenberger.

At the moment I only know of one, beautiful, case study of (presumably⁴) basic truth approximation by domain revision in mathematics, viz., Lakatos’ [1976] reconstruction of the history of Euler’s theorem about polyhedra. As is well-known, Lakatos shows convincingly, using a lively class room setting, with historical notes, that the history of Euler’s claim that the polyhedra satisfy $V - E + F = 2$ ($V$ the number of vertices, $E$ that of edges, and $F$ that of faces) is a history of domain revision. More specifically, it is domain revision by all kinds of revision of the definition of a polyhedron. Hence, the case also illustrates that basic truth approximation by domain revision may well take the form of “truth approximation by definition revision”.

David Atkinson provided me with a clear case in physics of “truth approximation by domain extension” in the form of definition revision, more precisely, in the form of successive definition extension: (putative) truth accumulation of the law of conservation of energy by extending the definition of “energy”. With his kind permission, I quote his description of the case in full.

In the elastic collision of balls, there is a conserved scalar quantity, the kinetic energy: if we know the mass of each ball, and measure its speed at a given time, we can calculate from these numbers the total kinetic energy of the system, and this does not change in time, at least if the effects of inelasticity and friction are negligible. If the balls are not free, however, but are attached to one another by an array of springs, the total kinetic energy of the balls does not remain constant in time, i.e. the law of conservation of energy breaks down. Or does it? By extension

⁴To be honest, I did not (yet) check whether there is an episode that satisfies the basic definition perfectly, without being merely a case of domain extension, but this seems very likely.
of the definition of energy to include the potential energy that is stored in an extended or a compressed spring – which can be calculated if one knows the (Hooke) spring constant –, one finds that the law of conservation of energy holds, on condition that one defines the energy as the sum of the total kinetic energy of all the balls, and the sum of the potential energies stored in all the springs. For the motion of cannon balls in the gravitational field of the earth, one finds that the same law is valid, where now the potential energy is stored in the gravitational field, and is proportional to the height of the ball at any instant.

A proviso in the above is that friction should be negligible. If it is not – or if one makes measurements with sufficiently great accuracy, or waits long enough – one finds that energy, as defined above, does not seem after all to be conserved, but appears to diminish slowly with time. The law is reinstated by a further extension of the notion of energy to include heat. A precise equivalence was found in the XIX century (Joule) between heat and work, which is a form of energy. If one adds the heat produced by friction, and that produced by any inelasticity in collisions, to the kinetic and the potential energy to produce the total energy, one finds that this is indeed conserved.

The conservation of kinetic energy + potential energy + heat was thought to be exact until, at the end of the XIX century, it was found that, in the decay of radioactive substances, there is a small mismatch in energy before and after decay. The law was reinstated by Einstein, who showed in his theory of special relativity that mass and energy are equivalent. In radioactive decay the sum of the masses of the constituents before and after decay was not the same, and if one extends the notion of energy once more to include mass, \( m \), and adds \( mc^2 \) to the kinetic energy + potential energy + heat, one finds that the law of conservation of energy is reinstated.

Since it is a case of successive domain extension by successive definition extension, it is, however, not the type of whimsical domain revision as in the case of polyhedra.

Gordana Dodig-Crnkovic provided me with an example that may be less straightforwardly cumulative with respect to the (definition of the) domain, viz., the study of magnetism. With her kind permission, I quote her indication of this case as well.

First, a number of observations (starting in ancient Greece) are
made about what magnets are and how they behave. Later on the concept of magnetic field is introduced. Then the fact that the moving electric charge induces magnetic field was established. Earlier concepts about magnetism were transplanted to the new environment. Maxwell predicted the existence of electromagnetic waves and identified light as an electromagnetic phenomenon. One has noticed that even subatomic particles possess magnetic properties, so the concept and the theory have shifted the physical domain again. Evidently, the domains of the first proto-theories of magnetism were very different from the domains of the theory of electromagnetism of today, that is, electromagnetism and quantum electrodynamics.

Only a full structuralist reconstruction could show whether this is a case of not purely cumulative domain revision. However, without that we have to face that there is no fixed theory, that is, there is no fixed theory of magnetism originating in ancient Greece.

5 Concluding remarks

The question remains whether there are convincing cases in the empirical sciences of basic or refined truth approximation by domain revision, other than pure domain extension, whether or not in the form of definition revision, other than successive definition extension. One reason to expect many explicit or hidden cases of this kind is the frequently confessed instrumentalism; for domain revision in practice evidently is a plausible form of instrumentalist practice. Happily enough for epistemological realists, it is possible to show [Kuipers, 2000] in general that the instrumentalist methodology with respect to theory revision is in theory, and hence in practice, functional for truth approximation and, even more surprisingly, more efficient than the falsificationist methodology. It is rather easy to see that a variant of the argumentation for this claim can be used for domain revision. The first claim is argued by realizing that, starting from a falsified theory, it may well be that the shortest route to the truth will go via intermediate theories that are also false. The falsificationist method instead forces one to start over again with a new theory as soon as the next theory again gets falsified. Similarly, given a theory, starting from a “non-fitting” domain, that is, a domain with respect to which the (weak claim of the) theory is false, it may well be that the shortest route to a maximally fitting domain, that is, a domain with respect to which the strong claim of the theory is true, hence the truth for that domain, will go via intermediate nonfitting domains. Again, the plausibly adapted version of the falsificationist method to domain revision, if that is acceptable from the falsificationist perspective
at all, forces one instead to start over again with a new domain as soon as the next domain again turns out not to fit the given theory.

Of course, in practice, neither the domain nor the theory will be fixed stubbornly, while the other is revised. As I wrote at the end of [Kuipers, 2000, p. 332], “science in the making”, that is, truth approximation in practice, is a matter of dialectical interaction between vocabulary, domain and theory. The above mentioned example of magnetism illustrates this process nicely.

Let us consider from this point of view the various attitudes that are possible with regard to a counterexample of the theory, assuming that we at least fix the basic vocabulary. Hence, I leave it to the reader to think about the way in which revision of the vocabulary can be brought into the picture [Kuipers, 2000, pp. 129–130]. So far we have suppressed the plausible distinction between the “core theory” and auxiliary hypotheses. Hence, in the spirit of the Duhem-Quine thesis, we may first look for auxiliary hypotheses that might be blamed for the counterexample, which may lead to revision of one or more of the auxiliary hypotheses, leaving the “core theory” intact. However, as soon as we accept the counterexample as a counterexample of the core theory for the domain, we may, secondly, look for an appropriate theory revision, or, thirdly, for an appropriate domain revision.

A similar story can be told about dealing with explanatory problems of a theory. An explanatory problem amounts to an established empirical law that cannot be derived from the theory; hence it is too weak for the domain. Again we may revise an auxiliary hypothesis, the core theory, the domain, the vocabulary, or a combination of these.

Of course, in all cases, we want to avoid ad hoc revision; we prefer revisions that lead to successful new predictions. However, we have not paid attention to “ad hoc-ness”, because from the truth approximation analysis it is clear that ad hoc changes are unproblematic if and only if they survive the severe testing of the relevant “comparative success hypothesis” [Kuipers, 2000, pp. 168–169], and this we have assumed throughout.

In sum, truth approximation by revision of theories, domains and vocabularies is perfectly possible, in particular when choices are guided by the instrumentalist rule of success. This is in sharp contrast to one of the main claims of Rivadulla [2006, end of section two]: empirical success is no indicator of truth, verisimilar or truth approximation. As a matter of fact, his illustrations of “historical failures of the Newtonian gravitational model in astrophysics and cosmology” nicely provides an historical case (section 4.1) I have been looking for. According to the Kelvin-Helmholtz gravitational collapse hypothesis the luminosity of the stars originates from the fact that
a star is the result of a collision of smaller bodies, due to mutual gravitation. This collision produces, according to the theory of Joule, an amount of heat equivalent to the kinetic energy lost in the collision. As Eddington noted, this explanation leads to bizarre conclusions about the age of the sun, reason for which “the source of stellar energy”, was no longer considered to belong to the domain of the theory of gravitation: hence, here we have a case of domain restriction.

It is important to note that the conclusion was not that the theory should be revised. As is well known, and reported by Rivadulla in section 4.2, this was essentially done in a rather drastic way by Einstein’s general theory of relativity, in particular, in light of the classical failure of Newton’s theory, viz., to explain the advance of Mercury’s perihelion. This was soon followed by the successful prediction of light deflection by the Sun and, much later, by the successful explanation of the phenomenon of red shift. Of course, these phenomena disappeared out of the domain of Newton’s theory, if they were ever considered as such at all, which is only the case for Mercury’s behavior. Moreover, one might say that the domain of Newton’s theory became essentially restricted to phenomena with small velocities and weak gravitational fields, that is, velocities and fields for which the theory is approximately true.

For completeness, I add that Rivadulla also shows that Newton’s theory can be successfully used to calculate the (relative) masses of binary stars (section 3.1), illustrating that the orbiting of binary stars, as might be expected, belongs to its domain. Moreover, he reports (section 3.2) the way in which Newton’s theory (or model, as he prefers) can explain the stability of stars, that is, that they do not collapse. A clear case of domain extension.

The illustration of domain revision by restriction, i.e., the exclusion of the problem of the source of stellar energy from the domain of Newton’s theory, suggests that several examples of so-called Kuhn-loss might be reconstructed as a case of domain restriction. Hence, we may conclude that truth approximation by domain revision is, at least in the form of domain restriction, fact rather than fiction in the empirical sciences.

Acknowledgements

This paper was first presented on the Workshop “Structures and Dynamics of Science”, in Paris in 2004, organized by Ulises Moulines. I am grateful for all comments that I received in Groningen, in Paris, and, last but not least, in Pavia. Special thanks go to David Atkinson, Gordana Dodig-Crnkovic, Nico Krijn, Andrés Rivadulla, Jan-Willem Romeyn and Rein Vos for their useful comments and suggestions on an earlier draft.
BIBLIOGRAPHY


Theo Kuipers
Department of Philosophy
University of Groningen
Groningen, The Netherlands
Email: T.A.F.Kuipers@philos.rug.nl
URL: http://www.rug.nl/filosofie/Kuipers