Truth Demystified

[Truth in the Model-Based Model of Cognition]

Karlis Podnieks

University of Latvia

What is true, will not change in the future (for some time, at least). Else, how could we benefit from knowing the truth? However, is truth any nobler than "something that will not change in the future"?

Let us start with the most fundamental philosophical position that is using a minimum of metaphysical assumptions, the so-called (subtleties aside) *methodological solipsism*: I have direct access only to my world of sensations; what is behind my sensations (and is there anything at all), I cannot know that with absolute certainty.

I will try to show how to convert this negative position into a positive philosophy of *demystified realism*.

First, let us adopt an appropriate model of cognition – the following *model-based model of cognition* (MBMC).

A model is anything that is (or could be) used, for some purpose, in place of something else. In this definition, models are meant to be concrete systems that serve as replacements of concrete target systems (for some concrete purposes).

Models are the ultimate goal of cognition. Humans and robots need models (in the above sense) to manage what is happening in the world around them.

Hence, cognition should be assessed, first of all, as the production of models and means of model-building. Means of model-building can be further subdivided into theories, methods, research programs, doctrines, paradigms, ontologies, logics, languages, etc.

According to this definition, the Standard Model of particle physics is not a model; being a theory, it belongs to means of model-building.

All the ideas contained in this definition were ready for use towards the end of 1980s. They can be traced to Hestenes (1992), Rothenberg (1989), Cartwright (1983), Minsky

(1965), Stachowiak (1965), Apostel (1960), Craik (1943), Einstein (1918), Poincaré (1904), Vaihinger (1876), etc. For most of references, see my (2009: Towards Model-Based Model of Cognition. *The Reasoner*, 3(6),5–6) and (2014: The Dappled World Perspective Refined. *The Reasoner*, 8(1), 3–4). For the reference to Kenneth Craik, I'm indebted to Philip Johnson-Laird. The reference to David Hestenes (1992: Modeling games in the Newtonian world. *American Journal of Physics*, 60(8), 732-748).

These ideas can be found also in the work of the actual community of modeling philosophers, see, for example, the volume (2009: M. Suárez (Ed.), *Fictions in Science. Philosophical Essays on Modeling and Idealizations*).

How could one reconstruct the notion of truth in the above model of cognition? How could we recognize truth, if we only have models, means of model-building, and the history of their evolution? Where is the truth in the cloud of models – with so many of them already gone with the wind?

First of all, truth must be considered not as a property of *propositions*, but as an assessment applicable to *knowledge constructs of any complexity* – to entire models and to means of model-building. Knowledge cannot be reduced to a heap of unrelated propositions.

Secondly, in MBMC, we cannot start our analysis by asking about truth as the correspondence between models and "reality". Reality is itself a construct used in model-building.

Thirdly, after so many cases of a strongly believed knowledge that was rejected afterward, how can we hope to determine *in advance* what exactly is "true" (will be retained) and what is "false" (will be rejected) in our *actual* knowledge?

The fourth argument is due to Craik (1943: The Nature of Explanation, *Cambridge University Press*, Cambridge, Ch. 5):

"... the model need not resemble the real object pictorially; Kelvin's tide predictor, which consists of pulleys on levers, does not resemble a tide in appearance, but it works in the same way in certain essential respects..."

How big could be the "pulley on lever" part of our best physical theories? Of the eleven-dimensional super-gravity, for example?

And finally, is anything achieved by qualifying some *non-observable* piece of knowledge as "approximately true"? Predictions generated by a model may be approximately true, in a well-defined sense, indeed. But what can be derived from the

"approximate truth" of the model structure? Or, from the "approximate truth" of the methods used to build the model?

However, if some knowledge construct really "works", we may decide to continue applying it. We may decide even to try extending the application to some other situations. In this way some knowledge constructs (models, means of model-building) become more or less persistent *invariants* of the knowledge evolution. It is this re-usability of knowledge constructs that inspires us to believe that these constructs are "true". But the real thing is the very phenomenon of invariance, and not the way how people categorize it. Truth is not a justified belief, on the contrary, *belief is an unquestioned truth*. (By the way, from this point of view, truths contained in the means of model building are more valuable than truths contained in individual models.)

Thus, in MBMC, we can define truths *as more or less persistent invariants of successful evolution* of models and means of model-building. What is true, will not change in the future (for some time, at least).

Hans Vaihinger identified the following fundamental successful knowledge constructs (he qualified them as fictions, of course): Kant's fiction of *Ding an sich* (i.e., the fiction of independent reality), the fiction of thing-attributes, the fiction of whole-parts, and the fiction of cause-effect. See Ch. XVII, XXV in Vaihinger (1911: Die Philosophie des Als Ob, *Reuther & Reichard*, Berlin). Do we intend to continue using of these constructs in our model building? If we intend, then we have here the first few truths to believe in: let us think of our world of sensations *as caused* by an independent reality, and let us use the concepts of thing-attributes, whole-parts and cause-effect for structuring it.

This approach to truth could be named *demystified realism* (or, demystified theory of truth) – the kind of realism based on a minimum of metaphysical assumptions. ("Robotic realism" also would be appropriate, but the term is occupied already.)

Three great pragmatist thinkers have expressed similar opinions:

"The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real." **C. S. Peirce** (1878: How to Make Our Ideas Clear. *Popular Science Monthly*, 12, 286-302)

"The 'absolutely' true, meaning what no farther experience will ever alter, is that ideal vanishing-point towards which we imagine that all our temporary truths will some day converge." **W. James** (1907: "Pragmatism's Conception of Truth". Lecture 6 in *Pragmatism: A new name for some old ways of thinking*).

"Pragmatism denies the possibility of getting beyond the Sellarsian notion of "seeing how things hang

together" – which, for the bookish intellectual of recent times, means seeing how all the various vocabularies of all the various epochs and cultures hang together." **R. Rorty** (1982: Consequences of Pragmatism: Essays, 1972-1980, p. XXXVIII)

What is achieved by this approach to truth? In particular, we obtain natural solutions to the "world riddles", for example, to the riddle of "approximate truth". "How true" is the idea of atoms? Do atoms really exist? Were Democritus, Dalton, Thomson, Bohr and (is) the modern physics modeling the same phenomenon? There is a more productive question to ask: *will physicists continue using the idea of atoms in their model building*? They will, and this is the demystified sense in which "atoms really exist".

Any deviations from this (trivial?) approach seem to introduce portions of mystics – they are creating the "world riddles" that remain unsolved for centuries. (Vaihinger was right, *oder*? See Ch. VI, XXI.)

After adopting demystified realism, you will not need to change your beliefs. You only should not try to obtain an *absolutely certain* proof that some of your beliefs are absolutely true. This will fail, and then, only one way will remain how to obtain a satisfaction – the way of mystification. (Thanks to Paulis Kikusts for asking the right question.)

A possible "killer" question: isn't the above definition of truth *trivial*, indeed? What is true, will not change in the future – what does that mean precisely? I will not change it? My community will not change? The human race will not? And how long must something stay unchanged to become acknowledged as truth? Indeed, some truths will persist only for some time, some – only for limited communities, but some – forever and for all of the human race? Which ones? *Do we really need to know that in advance and with absolute certainty?*

P. S. In Chapter 9 of Feynman's (1985: Surely You're Joking, Mr. Feynman!: Adventures of a Curious Character, with contributions by Ralph Leighton, *W. W. Norton & Co*) we can read another demystifying argument:

"The electron is a theory that we use; it is so useful in understanding the way nature works that we can **almost** call it real. … In the case of the brick, my next question was going to be, "What about the inside of the brick?" – and I would then point out that no one has ever seen the inside of a brick. Every time you break the brick, you only see the surface. **That the brick has an inside is a simple theory which helps us understand things better.** The theory of electrons is analogous." (my emphases)