How can scientific methods provide guidance for Semantic Web Research and Development?

Kjetil Kjernsmo[1,χ]

[1]Department of Informatics, University of Oslo, Norway
[χ]kjetil@kjernsmo.net

Abstract. This essay initiates an epistemological discussion on how current methodology is insufficient to gain knowledge on how to engineer the Semantic Web. It argues that the rate of progress is not up to par with that of natural sciences, that the community should think in terms of hypotheses and how their external validity must be established within larger theories, that also needs to be formulated. It then initiates a discussion that testing against a reality that does not yet exist is difficult, but for that reason, work on epistemology is needed. To motivate this work, it reviews some prominent directions in philosophy of science.

1   Introduction

The naive depiction of science is that it rapidly advances human knowledge, using clear methodology, by the cumulative addition of knowledge.

Let us as a starting point assume that science has indeed a remarkable rate of progress because scientists use methodologies that are uniquely suited to enhance the common understanding of phenomena, and thereby create new knowledge. This is an important motivation for research, since using science may accelerate the progress of the Semantic Web.

If this is true, it means that we cannot content ourselves with testing individual hypotheses, we must have methodologies to understand how a hypothesis would fit more generally into the Semantic Web. Then, we have to understand the nature of knowledge to understand the methodology that is used.

So, what does it mean to *know* something? That is one of the questions epistemologists concern themselves with.

Then, the question is if we can improve the rate of progress towards a Semantic Web by developing methodologies based on an improved understanding of epistemology.

The problem is not only epistemological, but also ontological: There isn't a clear, shared idea of what the Semantic Web is, and possibly, there shouldn't be. A highly influential vision was the article in Scientific American from 2001 [1], but this and other visions cannot define the Semantic Web operationally for science. Also, Frank van Harmelen, in his extensive keynote on searching for universal patterns [12], also discusses solutions to the heterogeneity problem. This underlines that a definition must encompass social, cultural and economical aspects. This can also be understood when considering the impact even simpler visions of the Semantic Web would have on society, it would be major revolution in fields of e.g. politics and economy. The impact of a full-scale successful Semantic Web would be such that e.g. workloads and data profiles would be very different from the modest successes of Linked Open Data or the use of large ontologies today. For the purpose of this discussion, the most important trait of the Semantic Web is that it doesn't yet exist.

The rate of progress has not been what I expected when I first got interested in the Semantic Web around the time when the Semantic Web Roadmap [10] was written. Can this lack of progress be explained in part by a lack of epistemological clarity, resulting in that the sizeable academic Semantic Web community has failed to provide scientific guidance to the development? And if so, why?

2 Philosophy of Science

A standard introduction to the philosophy of science is [2]. The entire book shows how difficult it is to find a firm philosophical foundation of science.

Karl Popper struggled with the problem of demarcation, i.e., a clear criterion that could tell science apart from pseudo-science. Popper argued that the method of science is one where *falsifiability* is of central importance, since a single counterexample can prove a hypothesis wrong. Therefore, Popper argued, scientific theories are those that are falsifiable, but haven't been proven wrong.

For my master's thesis work [3], falsificationism provided the epistemological framework I needed. I worked within two large theories, Theory of General Relativity, and Quasar Theory. Even within these theories, it was easy for a student like myself to find an interesting point to attack, and I found that there were certain parts of the parameter space where my work would not only be falsifiable, but could point out fatal flaws in Quasar Theory, and therefore were knowable. Other parts could not do so, and therefore, were unknowable.

Abraham Bernstein and Natasha Noy have published a report [4] with suggested research practice in our field. They take a falsificationist view of science, and their advice is insightful and well written, but I think it falls short of providing sufficient guidance in most cases, and for the reasons that Popper has been criticised for by more recent philosophers. The key problem with falsificationism is that the whole construction around testing theories or even individual hypothesis usually is a very complex one. If an experiment fails, one cannot tell if the hypothesis that is being tested has actually failed, or if there's something wrong with auxiliary hypothesis, or if the experimental setup is to blame. If you test a SPARQL engine for performance, you cannot tell if your implementation is flawed, or if something is wrong with the benchmark. Moreover, theories sometimes take a long time to develop, and therefore, falsificationism may therefore cause premature rejection.

The above argument was a central tenet of the writings of Paul Feyerabend, who attacked most aspects of the understanding that scientist seemed to have of themselves. Feyerabend rejected the idea that there is a scientific method, that knowledge is cumulative, that the rate of progress is measurably faster, and so on. Importantly, he accused science of having become as authoritarian as religions had been in the past. He quipped that falsificationism would destroy all of science for the reasons outlined above.

Thomas Kuhn also challenged falsificationism and the view that science is accumulating knowledge. He also insisted that philosophy of science must be based on the history of science. He did see science as a critical dialogue and emphasised the social aspects of science, but found no evidence that falsificationism had played an important role historically. He also gave us the terminological scaffolding of "paradigms", "normal science", "scientific revolutions" etc., where scientists in the "normal science" mode of operation are concerned with solving puzzles within the constraints of current theories. This goes on to a point where too many anomalies are accumulated for the theory to remain tenable, at which point a scientific revolution starts. Kuhn examines several scientific revolutions in his "Structure of scientific revolutions" [5] and attempts an extensive review of the revolution that happened in the transition from a geocentric to a heliocentric world view, in the "Copernican revolution" [6]. In spite of his insistence on the historical record as a basis of philosophy of science, he based important parts of it on a common myth: The myth that to make theory fit better to observation, astronomers had to add many "epicycles" (i.e. add more circles that travelled on the circles that described the planetary motions) to the astronomical model of the time. It has been showed in [7] that this did not happen. I appreciate Kuhn's argument against the cumulative nature of science, accept his emphasis on social processes and his scaffolding as useful. Though I agree with his view on the importance of history, his philosophy is otherwise hard to accept as his historical narrative is wrong.

After this review, I find very little that can bring epistemological clarity or

help establish a method that I can use in practical research.

3   A Theory of the Semantic Web

The rejection of the cumulative nature of science is common amongst recent philosophers. Is the realisation that well-established theories may later be rejected the motivation behind the title of this essay? Without going into subtleties, the answer is no.

To further understand my motivation, we have to discuss the term "theory". I have found surprisingly little guidance amongst philosophers as to what a theory is. Moreover, it is rare to see it used in our field the way it is used for e.g. "Theory of General Relativity", even rare to see the word "hypothesis" used. My own understanding of the term "theory" would be something like:

Theory

A theory is a coherent set of hypotheses that are held as true after having been subjected to vigorous testing by a scientific community.

This acknowledges empirical testing as a criterion for knowledge, and also Kuhn's emphasis on science as a social endeavour. The emphasis on testing is needed, because coherence (which not only requires consistency, but generally applies e.g. Occam's Razor [11]) is in my view not sufficient, allthough I acknowledge that it is useful to develop large formal frameworks, like string theory in physics, or indeed much of the work on logic in our community. As the Wikipedia page referenced above indicates, coherentism is a major direction, and even though we are perhaps not in the position to appreciate the depth of the philosopher's discussions, I mention this to say that it should not deter us from having our own.

By this definition, for something to be a science, it is not sufficient to just devise any test, the scientific community must be capable of vigorous testing.

To be able to test a hypothesis in the context of the Semantic Web, attention should be given to formulate a (or possibly several) Theory of the Semantic Web. In that context the search for universal patterns in [12] is interesting, as the discussion of what comes first, theory or facts, is a difficult one as the first few chapters of [2] show. A purely retrospective and descriptive view is however, not sufficient to test new hypotheses.

Now, assuming that we are able to formulate a Theory of the Semantic Web, the challenge becomes to test against it. This is relevant for the entire community, because even though those amongst us who are working with formal methods are highly adept at validating their research hypotheses with formal tests, they are still working with small part of a Theory of the Semantic Web.

I have previously argued that empirical research in our field is lacking [8],

but in principle we can test individual hypotheses vigorously. The external validity with respect to the "Theory of the Semantic Web" can nevertheless always be contested, based on the premise that the Semantic Web does not yet exist. Once it comes into existence, it will have characteristics (e.g. workloads. data profiles, etc.), very different from what it has now. It may be permissible to extrapolate such characteristics from what is presently deployed to test hypotheses, but the Semantic Web is currently so small compared to Web that I would contest the validity (relative to a large-scale Semantic Web) of any study based on that at present.

Say that I successfully test a hypothesis empirically with sound statistical methods. Yet, this does not answer key questions with respect to The Theory of the Semantic Web. First, if my hypothesis was that we have a performance gain in spite of a slightly longer connection time, this will be incoherent with a hypothesis that the connection time must always be kept to a minimum. Both these hypotheses may have been verified independently, but it is only with respect to the Theory that they will find a meaningful context.

Unfortunately, that context is something that does not exist, we can only speculate about it. Therein lies the fundamental problem: Whereas other sciences are testing against a reality that does exist, but with unknown properties, we are obliged to test against something that doesn't exist. Therefore, our results will always be inconclusive.

4   Discussion of consequences

4.1   Consequences for science

Certainly, fundamental research in informatics doesn't need to be tested against the Theory of the Semantic Web, for the same reasons that research on astrophysics doesn't need to justify its existence based on applications. It is beyond the scope of this essay to discuss such justifications.

However, a scientific work that makes the assertion that it forwards the Semantic Web certainly has the burden of proving that it does. Such papers must address the epistemological problems discussed in this essay. This requirement is equally strong regardless of whether the investigation uses formal or empirical methods.

I consider this fair: For example, a triple store, as a specialisation of a database management system, represents a distraction unless it can be argued that this particular specialisation is fruitful in a larger context.

Moreover, we cannot make a "Theory of the Semantic Web" falsifiable, if one so desired. If we are challenged by outsiders who say that "The Semantic Web isn't going to work", an assertion that certainly is common, can we address that scientifically?

4.2   Consequences for development

This inconclusiveness causes the breakdown of the motivation that we should use science to drive development, a motivation I have retained because I think a remarkable rate of progress in the sciences can be demonstrated, in spite of Feyerabend's attacks. This happens because we don't have a way to test hypotheses without for example a Semantic Web workload, and without science, we cannot help build that Semantic Web to get that workload.

I have always maintained that science and engineering must be seen as one, when working with assertions on specific technologies (except in some cases, like implementing most standards, which is pure engineering), to ensure that the extrapolation between known characteristics and future characteristics are kept at a minimum, but I'm not sure this is enough anymore, the problem appears fundamental.

I therefore write this essay to start a discussion around methodology, a discussion that I think must be rooted in epistemology, and what we can hope achieving when trying to accelerate the Semantic Web. A new direction, that I find appealing is known as "The New Experimentalism", see e.g. [9]. This direction focuses on experimentation, instrumentation, laboratory practices, etc. They argue that it is not sufficient to test a hypothesis, you need to prove that the test is severe. Then, they go on to formalize what severe means in statistical terms. Since I used similar statistical methods in [8] it seems clear that that formalism could be extended to severe testing. However, despite their efforts, I don't see how it applies beyond that type of study. Moreover, I don't see any immediate contribution towards the understanding of the fundamental problem of testing against something that doesn't exist, nor a thorough rooting in the history of science, but this is likely due to insufficient time on my part.

Acknowledgements

References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web, Scientific American, 2001, http://www.cs.umd.edu/~golbeck/LBSC690/SemanticWeb.html
2. Chalmers, A.: What Is This Thing Called Science? University of Queensland Press, Open University press, 3rd revised edition, Hackett, 1999.
3. Kjernsmo, Kjetil. Gravitational Microlensing of Quasar Clouds. Oslo: Digital publications of University of Oslo 2002. http://urn.nb.no/URN:NBN:no-2021
4. Abraham Bernstein, Natasha Noy, Is This Really Science? The Semantic Webber's Guide to Evaluating Research Contributions, Version: 1, 2014. (Technical Report) http://www.merlin.uzh.ch/contributionDocument/download/6915

5. Kuhn, T. S.: The Structure of Scientific Revolutions. University of Chicago Press 1962
6. Kuhn, T. S. The Copernican Revolution. Harvard University Press, 1957. ISBN 0-674-17103-9.
7. Gingerich, Owen. "Crisis" versus aesthetic in the Copernican revolution. Vistas in Astronomy 17(1): 85-95, 1975. http://nrs.harvard.edu/urn-3:HUL.InstRepos:4258973
8. Kjernsmo, K.: Introducing Statistical Design of Experiments to SPARQL Endpoint Evaluation, The Semantic Web – ISWC 2013 Lecture Notes in Computer Science Volume 8219, 2013, pp 360-375. http://folk.uio.no/kjekje/#doe-sparql1
9. Mayo, D. G., Evidence as Passing Severe Tests: Highly Probable versus Highly Probed Hypotheses. In P. Achinstein (ed.), Scientific Evidence: Philosophical Theories & Applications. The Johns Hopkins University Press. 95--128 (2005) http://www.phil.vt.edu/dmayo/personal_website /(2005)%20Highly%20Probable%20vversus%20Highly%20Probed%20Hypotheses%20 F2.pdf
10. Berners-Lee, T.: Semantic Web Roadmap, 1998 http://www.w3.org/DesignIssues /Semantic.html
11. Wikpedia: Coherentism, https://en.wikipedia.org/wiki/Coherentism, accessed on 2015-03-18.
12. Frank van Harmelen: 10 Years of Semantic Web research: Searching for universal patterns http://videolectures.net/iswc2011_van_harmelen_universal/, accessed on 2015-05-26.