1. Some (provocatively over-simplified) history

The early AI success stories of the 1970s were based in small 'worlds' with carefully bounded semantic domains: Winograd's SHRDLU (WIN72) is perhaps the canonical example. The rapid growth of efforts to found the next generation of systems on general-purpose knowledge representation languages (I'm thinking of several varieties of semantic nets, from plain to partitioned, as well as KRL, KL-ONE and their successors, ending (not yet, of course) with CYC (See BRA08 for all these) stumbled to a halt once their failure to provide any viable inference mechanism was recognised as incorrigible.

Progress continued on two fronts, however:

- Inference engines for less elaborate KR languages have advanced from resolution theorem provers through a number of stages to the current proliferation of a range of Description Logic 'reasoners';
- As the scale and complexity of the 'worlds' of interest grew, so did the need to manage the impact of change and conflict: enter 'truth maintenance', subsequently renamed 'reason maintenance'.

But outflanking these 'normal science' advances of AI, the paradigm shifters were coming up fast on the outside: over the last ten years machine learning has spread from small specialist niches such as speech recognition to become the driving force for success across a wide range of problem domains, from machine translation to expert systems.

2. An aside on rationalism vs. empiricism

It's tempting to see the above-noted paradigm shift as a shift from rationalism to empiricism. It is certainly seen that way in the natural language community. Up through the mid-1970s (and later, the further up the speech chain you went), virtually all forms of computational engagement with natural language were completely invested in the Chomskian rationalist tradition. This was manifested in a corresponding commitment to formal systems, representationalist theories of mind and, usually, so-called 'strong AI'.

But starting in the late 1970s, in the research community centred around the (D)ARPA-funded Speech Understanding Research effort, with its emphasis on evaluation and measurable progress, things began to change. (D)ARPA funding significantly expanded the amount of digitised and transcribed speech data available to the research community. Instead of systems whose architecture and vocabulary were based on linguistic theory (in this case acoustic phonetics), new approaches based on statistical modelling and Bayesian probability emerged and quickly spread. "Every time I fire a linguist my system's performance improves" (Fred Jelinek, head of speech recognition at IBM, c. 1980, latterly repudiated by Fred but widely attested).

Today progress in speech recognition and many other branches of perceptual-interface aspects of AI depend to a considerable extent on only two things: larger datasets for training, and better machine learning techniques. As more and more problems are re-conceived as instances of the noisy channel model, the empiricist paradigm continually broadens its reach.

And rationalism? Whereas in the 1970s and 1980s there was real energy and optimism at the interface between computational and theoretical linguistics, the overwhelming success of the empiricist programme in the applied domain have separated them once again. While still using some of the terminology of linguistic theory, computational linguistics practitioners are increasingly detached from theory itself, which has suffered a, perhaps connected, loss of energy and sense of progress. Within cognitive psychology, there is significant energy going in to erecting a theoretical stance consistent with at least some of the new empiricist perspective. But the criticism voiced 25 years ago by Herb Clark, who described cognitive psychology as "a methodology in search of a theory", remains pretty accurate. And within computer science in general, and Artificial Intelligence in particular, the interest in "probably nearly correct" solutions, as opposed to constructively true ones, is dominant.

3. Let me recite what history teaches:

"History teaches", as Gertrude Stein says. Over the last ten years the Semantic Web project has recapitulated, with all too painful precision in some cases (I have in mind the recurring enthusiasm for Intelligent Agents), the history of AI as outlined above. We started with static representations, which any fool could see contained a wealth of information. We're really only just starting to get to grips with the infer-
ence problem, since it turns out that computers aren’t just any fool, but particularly obdurate ones. Reason mainte-
nance is . . . , well, yes, trust and provenance and attribution 
are important, and we’ll need to tackle them Real Soon Now.

But machine learning? Almost not at all. The SemWeb 
community still seems in the grip of good old-fashioned 
logic-grounded AI: sentences represent states of affairs, 
they are true or false, and inference engines which accu-
rately reflect the relevant proof theory can tell us which. 
The prospect of abandoning certainty in favour of probably 
early is apparently abhorrent.

4. Not a threat, an opportunity

Maybe the shift from talk about 'the Semantic Web' to
talk about 'Open Data' heralds the advent of the empiricist 
hordes at, or even inside, the gates. But I don’t see much sign 
of a change in attitude amongst the SemWeb 'old guard'.
I recently heard Web Science described, by an outsider, as 
the empirical study of what actually happens on the Web:
the correlation between TLDs in search-result clickthrough 
with time of day, for example. That’s empiricism alright, but 
not obviously related to the matter at hand in any interesting 
way. More to the point is the work of Halpin (HAL09) and 
Fernandez (FERN09), treating aggregates of RDF triples as 
grist for the machine learning mill in the context of state-of-
the-art information retrieval. The Statistical Semantic Web?
To exploit the reality of the kind of linked data which is actu-
ally rapidly appearing, i.e. data which is not consistent with 
regard to ontology, these techniques are essential.

5. The end of the story?

A lot of value will come from statistically-based exploitation 
of SemWeb data, I have no doubt. But that value will not 
realise the original "look Ma, no humans" Semantic Web 
vision. The hard problems which have beset AI for the last 
forty years are still largely unsolved. Although RDF’s use of 
URIs distinguishes it from the historical KR languages, and 
may yet provide real benefit, it seems implausible at best that 
that difference alone will give us the breakthrough we need.

References

Brachman, Ron and B. C. Smith, eds, (1980). Why is 
a Knowledge Representation System Like a Piano  ACM 
SIGART Bulletin Special Issue 70, Association for Com-
puting Machinery, New York, NY, USA.

Fernandez, Miriam et al., "Using TREC for cross-
comparison between classic IR and ontology-based 
search models at a Web scale", 2009. Presented 
at Semantic Search 2009 Workshop, WWW2009, 
Madrid, Spain. Available online as http://km.aifb.uni-
karlsruhe.de/ws/semsearch09/semse2009_18.pdf.

Halpin, Harry and Victor Lavrenko, "Relevance Feedback 
Between Hypertext and Semantic Search", 2009. Pre-
sented at Semantic Search 2009 Workshop, WWW2009, 
Madrid, Spain. Available online as http://km.aifb.uni-
karlsruhe.de/ws/semsearch09/semse2009_27.pdf.

Winograd, Terry, Understanding Natural Language, 1972. 