Structural Theory of Science as a Systematic Framework for the Design of DL's and CD's for E-Science

Hansje Braam

Utrecht University, Faculty of Humanities
hansje.braam@let.uu.nl

1 Context

The construction of ontologies (DL-instances, Concrete Domains) for E-Science (in my individual case E-Humanities) is quite hard:
i) in many cases we have to cope with ontological divergencies, up to completely disjunct ontologies, as many competing theories (and scientific practices) rely on non-common features, entity-classes and relations;
ii) many scientifically relevant predicates have arity > 2, a potential problem for DL(CD)'s;
iii) in science (natural science, but humanities as well) formalisms are used with high complexity (mathematics, formal grammars, predicate logics, restricted natural language etc.);
iv) higher order predicates occur -especially in the humanities- (f. e. it is possible to reason about certains modes of interpretation). It is a.o. this semantical complexity which distinguishes science from folk cognition.
I propose to use Structural Theory of Science (the classical reference [1], see also [2]) as a common framework for the analysis of scientific theories (and practices) in which the specifications of adequate ontologies can be related and subsequently used in the systematic construction of the DL's and CD's.

2 Structural Theory of Science

Structural theory of science (STS) is basically a semantical approach to scientific theories. It focusses on a model-theoretical exposition of the meaning of theories. Thus, it is not restricted to formal deductive theories. STS has been applied to wide range of scientific theories, including physics, linguistics, psychology, economy and theory of literature. It also facilitates the exact study of theory dynamics and intertheoretical relations. Theories are not approached as deductive systems of statements (axioms, propositions, theorems etc.) but as possibly complex expressions about the relations between certain sets of models. Models are tuples of domains and predicates of the theory. Simply put, four levels are relevant:
i) the set of the intended applications $I_A(T)$ (in most cases, that what the theory

is about: the empirical field);
ii) the set of all possible models for the basic ontology of the theory $M_{pp}(T)$, excluding its theoretical notions (e.g. including only what the theory takes for granted);
iii) the set of all possible models for the ontology of the theory $M_p(T)$, which includes the domains and ranges of its theoretical predicates;
iv) the set of models of the theory itself $M(T)$, the theoretical predicates are specified in full.
An approximating function F relates the set of intended applications to M: the empirical claim of a theory T is that $F(I_A(T)) \subset M(T)$. Generally we have $F(I_A(T)) \subseteq M(T) \subseteq M_p(T)$ and $M_{pp}(T)$ is a projection of $M_p(T)$. In DL-terms one could associate I_A with possible models for the ABox of T and M with all models for its TBox. Example: for a simple theory ILT of interpretation of literary English works we could have $M_{pp}(ILT)$ as a set of tuples $\langle Text, Interpreter \rangle$, $M_p(ILT)$ as tuples from $M_{pp}(ILT)$ extended with the full $Interprete$-relation (i.e. $Interpreter \times Text \times Interpretation$), where $Interpretation$ is a ILT-qualified subset of $Text$. The specification of $M(ILT)$ (by specification of $Interprete$) may well be non-computational. $I_A(ILT)$ covers English literary works with interpretations by certain interpreters. The intertheoretical relations which can be defined on these 4 levels are quite relevant for epistemological and methodological reasons. These relations are therefore also important in the academic educational realm.
STS is expressed using set-theoretical concepts [4] and cannot be used as an implementation language, it is much too strong for that. However, it can be used as a common context for the study of possible DL and CD versions for fragments of scientific theories. In this way we can have a theoretical environment for systematic analysis and development of DL/CD for E-Science. Important issues are how the STS concepts of theoreticity, theory-nets, reduction and emergency [3] can be connected to relations between ontologies in the DL/OWL world. An urgent need is the extension of STS (and DL's) towards the incorporation of empirical experimentation. This will need extensive analysis of experimentation and data gathering itself [5] as well as action logics.

References

1. Balzer, W., Ulises Moulines, C., Sneed, J. D.: *An Architectonic for Science. The Structuralist Program*, Reidel, Dordrecht etc. 1987.
2. Balzer, W., Sneed, J. D., Ulises Moulines, C. (eds.): *Structuralist Knowledge Representation. Paradigmatic Examples*, Rodopi, Amsterdam etc. 2000.
3. Bickle, J.: Concepts Structured through Reduction: a Structuralist Resource Illuminates the Consolidation-Long-Term Potentiation (LTP) Link, *Synthese*, **130**, 2002, 123-133.
4. Suppes, P.: *Representation and Invariance of Scientific Structures*, CSLI, Stanford, 2002.
5. Gooding, D.: *Experiment and the Making of Meaning. Human Agency in Scientific Observation and Experiment*, Kluwer, Dordrecht etc. 1990.