SKI*ing* with DOLCE: toward an e-Science Knowledge Infrastructure

Boyan Brodaric[1], Femke Reitsma[2], Yi Qiang[2]

[1]*Geolological Survey of Canada, Ottawa, brodaric@nrcan.gc.ca*

[2]*UK National e-Science Center, University of Edinburgh, femkereitsma@gmail.com, yi_qiang@hotmail.com*

Abstract: An ontology of general science knowledge (SKIo) is developed to enhance machine representation and use of scientific theories in emerging e-Science Knowledge Infrastructures. SKIo specializes the DOLCE foundational ontology with science knowledge primitives, such as science theory, model, data, prediction, and induction. These are arranged to reflect the complex knowledge structures used in science, such as scientific ideas playing different roles within and between theories. SKIo is encoded with OWL-DL, uses the DOLCE Descriptions and Situations module, and provides defining conditions for its primitives to enable an extensible bridge between DOLCE and domain science ontologies. An application to environmental theories is demonstrated, and its utility to other natural sciences is promising.

Keywords: science ontology, scientific reasoning, DOLCE, environmental model

1. Introduction

Ontology-enabled infrastructures in support of cyber-based scientific activity, or e-Science, are being developed and used in many science domains [1]. This is leading to significant scientific and societal benefits, in that faster computations are occurring over more data and the resultant predictive models are providing larger and more accurate scenarios about situations affecting humans and the environment. Although these early e-Science achievements are laudable and significant, they do fall short of a broader e-Science vision in which scientists not only operate over more observed data to make better predictive models, but also directly use e-Science infrastructures to find, generate and test science theories. This broader vision of e-Science requires a Science

Knowledge Infrastructure (SKI [2]) that enables the capture, representation and use of the full spectrum of science knowledge. Using SKIs, scientists should be able to annotate existing resources, such as observed data and predicted models, with respect to potentially competing science theories, in order to enable knowledge search and evaluation, as well as facilitate reproduction of simulations, experiments and results.

The present focus on a fragment of science knowledge has some negative consequences as it limits full scientific discovery and reproducibility in e-Science infrastructures. This occurs because only some aspects are explicitly represented in the infrastructure (e.g. data, models, concepts), while other knowledge (e.g. science theories) is largely implicit as it is buried in scientists' heads and in ancillary resources such as textbooks, papers, reports and maps. An initial challenge then is the development of a computable representation of a wide suite of science knowledge primitives. Foundational ontologies are a good candidate for such representation not only due to their formality, rigor and commitment to internal coherence, but also due to their generality in that, like science knowledge primitives, the contents of foundational ontologies are intended to be re-used across science domains. This contrasts with the numerous ontologies being developed for specific science domains, and is aligned with the few that are being developed as a general science superstructure, but these latter are narrowly focused and do not often utilize a foundational ontology.

In this paper we specialize the DOLCE foundational ontology [3] with a modest number of science knowledge primitives, such as science model, science theory, data, prediction and induction, and test the resultant ontology (SKIo) by representing environmental theories. Section 2 describes a typical use-case scenario in the environmental sciences; Section 3 discusses related work; Section 4 explains our general approach based on computationally inspired renditions of the science knowledge cycle; Section 5 presents SKIo; Section 6 outlines some results from using SKIo to represent environmental theories; and Section 7 concludes with a brief summary.

2. An Environmental Modeling Use-Case Scenario

Problem Scenario: Jane is a scientist who wants to integrate into a global climate scenario some model of Net Primary Production (NPP), which involves the conversion of solar energy, carbon dioxide and water, into biomass. She begins searching using a few keywords in Google as well as in her University library database. She finds a huge number of results, and after much sifting has a large collection of papers that seem relevant but cannot be easily differentiated, largely because they use polysemous terminology. For example the key term "model" is used in several senses in the environmental modeling literature, but only the first two senses are relevant to her:

1. model = a system of equations to support calculations and simulations [4];
2. model = a theory with equations and broader scientific implications [5];
3. model = a simulation software with equations and implications [6];
4. model = the results of a simulation run, or other process, in which some geographically located climatic situation is represented [6].

Proposed Solution: Jane logs on to a web-based SKI and begins searching for a relevant "model" by using a number of concepts she is familiar with in NPP modeling; she expects these concepts to be used as variables in equations. Because the different senses of "model" are well demarcated in the SKIo ontology, and because the SKI's contents are annotated by this ontology, she is able to find candidate "models" corresponding to senses 1 and 2 linked to digital resources. After integrating a newly found model and running her experiment, she creates a web page documenting the process, appends a draft publication, and annotates the resource in the SKI.

Additional Requirements: SKIs should also help scientists resolve questions such as: who else has solved problem *p*, or a similar problem in another domain? Who is working in the same research field? What results when existing theory x_1 is replaced by a new theory x_2 and tested against data *y* as originally reported in journal paper *z*? What other data satisfy x_2, and to what degree? What other theories are satisfied by *y*, in which papers, and how do these differ from x_1? How was x_1 derived—what observed data, reasoning, and verification procedures were used, as reported in which papers? What other theories is x_1 part of, and what is its role in those theories? What theories have been derived from x_1? What theories could be derived from x_1 that satisfy *y*?

3. Related Work

Although ongoing work on science ontologies is vast, and growing, at present a machine-readable ontology for general science knowledge does not exist. Existing initiatives emphasize the computational representation of the science knowledge cycle, the development of ontologies that span aspects of all sciences, are limited to one science domain, or which incorporate foundational ontologies:

- **The Science Knowledge Cycle:** Several accounts of the science knowledge cycle begin to distill the numerous and complex philosophical approaches into representations amenable to computation. These focus on identifying key primitives in the cycle [7, 8, 9, 10] for incorporation into schemas [2], formal reasoning systems [11], or machine-readable science theories [12], but without explicitly representing the primitives in a machine-readable ontology.

- **Ontologies of Science:** by science ontologies we mean a conceptualization of general science knowledge primitives that can be applied across a wide breadth of science domains and which are well defined and represented in a formal language. Existing science ontologies meet this definition partially because they focus on a fragment of key primitives such as science experiments [13] or publications [14].
- **Domain Science Ontologies**: though ontologies are being developed in numerous science domains [e.g. 15, 16], they cannot serve as a superstructure for science knowledge because the abstractions are not sufficiently general. They are also used primarily for engineering purposes to facilitate data interoperability and workflow operation [17, 18] rather than to annotate and test new science ideas. Many are built bottom-up from existing vocabularies [16] and not around systematic ontological principles such as those utilized by foundational ontologies, resulting in diverse ontological assumptions that are not easily recognized or reconciled.
- **Foundational Ontologies and Science Knowledge:** foundational ontologies provide a superstructure containing the most general abstractions that can be extended to both general science ontologies and domain science ontologies, e.g. DOLCE, BFO, GFO, SUMO, Sowa's [8, 19, 20, 21, 22]. An ideal arrangement would then position a general science ontology as a layer between a foundational ontology and domain science ontologies. With one exception [16], this intermediary layer is at present missing: domain science ontologies directly specialize existing foundational ontologies or related logical theories [23, 24, 25].

4. Approach

SKIo is first represented in UML [26], and then in OWL-DL (http://www.nesc.ac.uk/technical_papers/skio3.owl). Following OWL terminology conventions, the science knowledge primitives are denoted as classes and properties: classes refer to abstractions that can be instantiated in one or more individuals, and properties refer to relations between two classes; individuals are single entities that instantiate a class, i.e. instances. Class and property names are shown in italics herein.

Several principles are followed in the design of SKIo (after [27]): (1) **Modularity**—general science classes and properties are to be added as leaves to the DOLCE hierarchy of classes and properties, such that the original hierarchical structure remains unchanged; (2) **Semantic Grounding**—SKIo is to be founded on recognized accounts of the science knowledge cycle; (3) **Semantic Coverage—**sufficient breadth of the science knowledge cycle is to be encompassed such that SKIo could be specialized by general and domain science ontologies; (4) **Semantic Precision—**sufficient depth is to be attained to enable annotation of science documents through instantiation of SKIo primitives; (5) **Coherence:** these principles are to be formalized to enhance definition and understanding of SKIo components.

5. The SKI ontology

The DOLCE 2.1 (OWL 397) ontology consists of four core classes that categorize particulars (singular entities in the world): *endurant*, *perdurant*, *quality*, and *abstract*. An *endurant* is an object-like entity that is wholly present at any point in time it exists, but whose characteristics can change over time (rock body, building, country); a *perdurant* is a process-like entity that is not wholly present at any point in time it exists, such as a *process*, *event*, or *state* (San Andreas faulting, San Francisco earthquake, being seismically active); a *quality* is a dependent characteristic inherent in an *endurant*, *perdurant*, or *abstract*, such that an *endurant* inheres physical qualities (geospatial position, size, shape, color), a *perdurant* inheres temporal qualities (duration, age), and an *abstract* or *non-physical-endurant* inheres abstract qualities (the value of the Canadian dollar); an *abstract* is an entity that does not possess physical or temporal qualities, and is often the value of a *quality*, or a space containing those values (the number 2, the munsell color space, blue).

SKIo specializes the DOLCE foundational ontology with a relatively small number of general science knowledge primitives synthesized from computational accounts of the science knowledge cycle (mainly [9, 11]). In this cycle, shown in Figure 1, scientific artifacts are produced by scientific activities: empirical regularities are induced from observed data, hypothetical propositions are abduced from all prior knowledge, predictions about the real world are deduced from empirical regularities or hypothetical propositions, and predictions are verified through further interaction with the world, which involves activities such as data collection, problem finding, and building models of the world. SKIo specializes the *activity* subclass of *perdurant* to include such science activities and it specializes some physical and social subclasses of *endurant* (*description*, *situation*, *concept*, *information-object*, *physical-endurant*) to include various science artifacts. To foster reproducibility and enhance explanation, each science artifact is defined at least in part by the science activity that produces it.

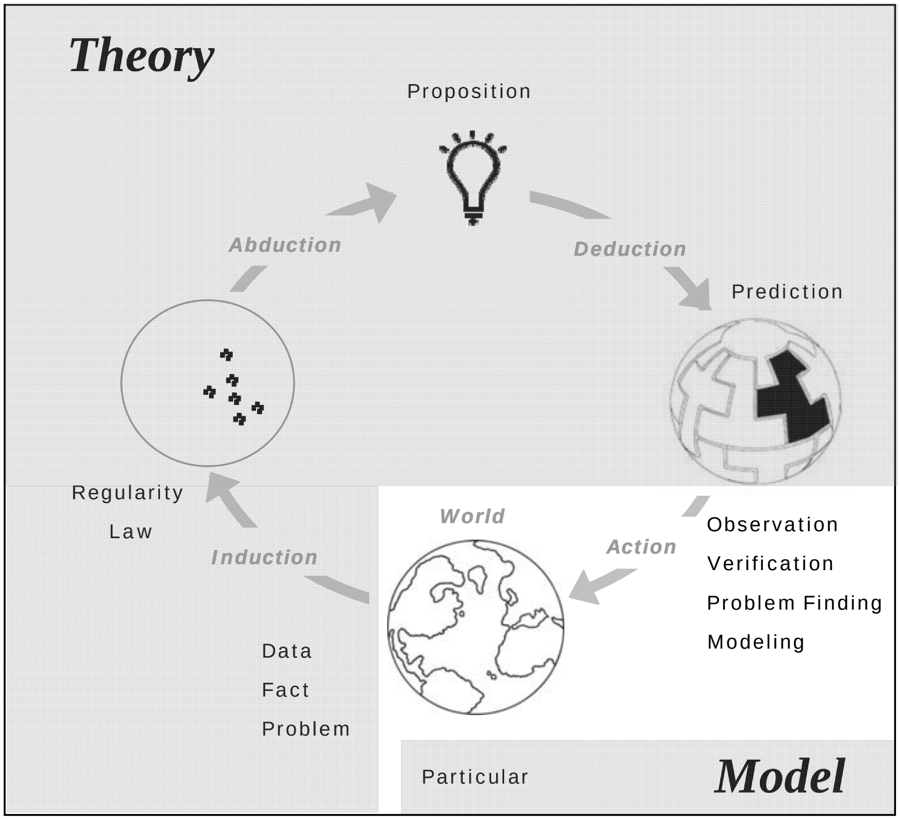

Figure 1: the knowledge cycle in SKI [after 8, 9]; grey areas represent some theory and model contents.

5.1. Descriptions and Situations

DOLCE's descriptions and situations are designed to represent socially constructed contexts and the states-of-affairs interpreted by those contexts, respectively [28]. Although descriptions and situations have been applied to the biomedical domain [24], they are specialized by SKIo into general science classes to provide scientific constraints on their meaning. In SKIo, a description is a science idea that is syntactically expressed as a *ScienceStatement* (a DOLCE *information-object*), such as a text body, figure, or web site, and which is contained in a *SciencePublication* as well as physically manifest in multiple forms such as in hardcopy or computer memory. A science theory is then a science idea comprised of one or more coherent descriptions that characterize the structure or behavior of some aspect of reality in sufficient generality to satisfy a wide number of science models and be used to predict the particulars in the models [2]. Specifically, a *ScienceTheory* contains descriptions which are satisfied by (can be used to scientifically predict) the particulars in the science models, and conversely a *ScienceModel* contains scientifically discovered particulars which satisfy (are scientifically predictable by) some *ScienceTheory*. Specializations include *GeoScienceTheory* and *GeoScienceModel* consisting of geoscience particulars.

For example, if a geoscience theory consists of equations then satisfaction implies that model members can be calculated from the application of data to the equations. A1 states this in a first order logic reference statement via the *predictable-by* property, which takes as its range a *Prediction* (the result of a *Deduction*), indicating a particular is forecast by the prediction. Formally, given *ScienceModel* (M), *ScienceTheory* (T), *Prediction* (p), *particular* (m) a member of model (M), and a *particular* (y) then:

(A1) $\forall (m \in M)\, \exists T\, [\mathit{satisfies}\,(M,T) \leftrightarrow \exists p\, [\mathit{predictable\text{-}by}\,(m, p, T) \rightarrow \exists y\, \mathit{Deduction}\,(T, y, p)]]$

When the description itself is a concept definition, then any satisfying model is part of the extension of the concept, because each particular in the model can be deductively classified from the concept definition. This includes the case where some extension members might be designated as prototypes for the concept. In SKIo, a *Definition* is a canonical idea included mainly for explicitness, as DOLCE descriptions can be definitions implicitly. *Data* is also a science idea, one that results from the observation or inference of some quality. Specifically, *ObservedData* results from the observation of physical or temporal qualities, as it is assumed that abstract qualities are not observable but are inferred. The syntactic expression of some data is a *DataSet*.

5.2. Concepts

The DOLCE *concept* class is a contextualized socio-cognitive artifact, one used to classify a particular within a situation to enable it to satisfy a description [29]. DOLCE provides three types of concepts: *role*, *course*, and *parameter*, for classifying endurants, perdurants and quality regions, respectively. Concepts are related to descriptions in four ways in SKIo: (1) a concept can be defined by a description; (2) a description (e.g. a theory part) uses concepts in its body to describe an idea; (3) the *role* concept is specialized to *ScienceRole* to represent the science function performed by a theory or its part, because a science idea can be expressed by many statements and can play different roles within and between theories; and (4) a science theory can maintain an index of its component concepts such as parameters or science roles. For example, the theory of special relativity has as a part the idea $e = mc^2$ that: defines the *parameter* concept *energy*; *d-uses parameter* concepts *energy* (e), *mass* (m), and *constant speed of light* (c); *plays* the *ScienceRole* of *Proposition* within the theory; and is indexed by the theory for its *ScienceRole* and *parameter* concepts which become components of the theory. In subsequent theories the same idea plays the role of *Assumption* or *ScienceProblem* [30]. Key science roles in SKIo include:

- *Assumption*: is defined by an originating *Assertion* and is considered primitive such that the related idea is not necessarily empirically supported or inferred.
- *ScienceProblem*: is defined as the product of *ProblemIdentification* in which a theory part is identified as problematic because it exhibits inconsistencies (initially obtained via *Verification*) with observed data, or inferred theory or models.
- *Fact*: is the incorporation of some *Data* into some theory. Because of this dependence facts are always 'theory-laden'. Facts also support specification of the scientific discovery of a particular, in that a fact can indicate that data leads to the identification of a particular, e.g. a magnetic measurement indicates a rock body.
- *EmpiricalRegularity*: is defined as a situational empirical pattern produced by *Induction*. Situational refers to the case where the regularity is satisfied by some but not all of the possibly valid science models. For example, if the regularity is expressed as a relation amongst concepts, then the relation is present only in some but not all situations in which particulars classified by the concepts are jointly present. The regularity might not be present universally because of insufficient verification or because the pattern is dependent on certain historical conditions that are temporary and change in time. This is implemented in SKIo by requiring the regularity's existence to be dependent on one or more endurants or perdurants, likely some subset of those involved in its original induction.
- *ScienceLaw*: is defined as an universal empirical pattern produced by *Induction*, i.e. its existence is not dependent on any specific endurants or perdurants. Empirical regularities and science laws can evolve toward each other by the logic of induction: additional data might suggest that a situational pattern is universal, and conversely more data might contradict a science law and reveal it to be situational.
- *Prediction*: is defined as a conjecture about individuals produced via *Deduction* that can be empirically verified. Because only physical and temporal qualities are observable it follows only these qualities are predictable, but SKIo does not impose this constraint as it is convenient to also predict individuals.
- *Proposition*: is a best-guess conjecture produced via *Abduction*, which can be situational or universal, and about individuals or theories.

5.3. Activities

DOLCE activities are non-atomic perdurants that follow some plan, sequence some tasks, can produce some endurants, and are performed by some agents. In SKIo the plan is likely some research project containing tasks performed by scientists, or some computational methods containing procedural tasks syntactically expressed in SKIo *Software* and performed by computers. SKIo activities include observation, inference, assertion, verification, problem finding, science modeling, and doing research. These produce key science artifacts such as science models, science roles, and information objects. Of particular importance are the *Inference* activities as they bind together much of the knowledge cycle (after [8, 9, 11]):

- *Induction*: involves finding a pattern in data (logical induction), or dis/confirming a pattern, or instance, via data collection and evaluation (pragmatic induction). In logical induction: given data $\{(a_1, b_1), (a_2, b_2), (a_3, b_3)\}$ then infer T (A, B), where T is some theoretical relation over classes A, B, and a_i and b_i are their respective instances. In pragmatic induction: given T (A, B), note in newly observed data that $\{(a_1, b_1), (a_2, b_2), (a_3, b_3)\}$ and infer T (A, B) $\models$ TRUE (or FALSE if disconfirmed), or given a_1 note in data $\{a_1, a_1,\ldots\}$ and infer a_1 = TRUE (or FALSE if disconfirmed). SKIo *Induction* refers to logical induction performed on *Data* where the induced T is an *EmpiricalRegularity* or *ScienceLaw*, while *Verification* encompasses pragmatic induction resulting in dis/confirmation of a science role when tested against data, or in/coherency when tested against theoretical relations.
- *Deduction*: involves generating a prediction about the world using existing theory and data. Logically, given theory $T{:}A \rightarrow B$ (A and B are classes) and instance a_1 (of A), then b_1 (of B) is deduced: $T \wedge a_1 \vdash b_1$. In SKIo *Deduction*, T is realized as a theoretical science role (*Proposition*, *EmpicalRegularity*, *ScienceLaw*), a_1 as a science role about an observed or inferred instance (*Fact*, *Prediction*, *Proposition*), and b_1 as the resultant *Prediction*.
- *Abduction*: involves guessing a *Proposition* to enable coherence of discordant science knowledge. Logical abduction is reverse deduction: in a deduction of the form $T \wedge a_1 \vdash b_1$, as above, where T and b_1 are known, then guess the missing instance a_1. Pragmatic abduction, on the other hand, is more concerned with guessing theory T via transformation from radically different prior knowledge (often via analogy ~): e.g. given $T_1{:}C \rightarrow D$, and $\sim(A,C) \wedge \sim(B,D)$, then $T_2{:}A \rightarrow B$. SKIo *Abduction* encompasses both logical and pragmatic abduction, such that the resultant *Proposition* refers to either a_1 (an instance) or T_2 (a theoretical relation).

6. Application of SKIo to Environmental Modeling

To exemplify SKIo, we represent NPP theories (i.e. “models” [4]) such as those sought by Jane in Section 2. The theories are expressed primarily as systems of equations containing (1) input variables and (2) fixed values segregated into general constants (e.g. atmospheric pressure) and specific parameters (e.g. temperature). The theories can be categorized into biogeographic or biochemical, where the former are empirically inferred from data and the latter are derived from existing theories and express biochemical processes more explicitly. Some of these theories share a few ancillary theories which are obtained from other sources such as books or web pages. SKIo representation of all these elements involves the following SKIo classes, which are applied to the BIOME3 Model [31] in Figure 2:

- Each paper, book, and web site that expresses the theories is represented as an instance of *SciencePublication* containing *ScienceStatement* instances.
- Each equation, or other theory part such as a table or figure, is expressed as a distinct instance of *ScienceStatement*.
- The description of each theory is represented as a whole *GeoScienceTheory* instance, and is expressed by the relevant *SciencePublication* instances.
- The description of each equation, or other theory part, is represented as a *GeoScienceTheory* part instance, and is expressed by *ScienceStatement* instances.
- The variables (e.g. temperature) and fixed values in each equation are represented as DOLCE *parameter* classes such that an instance of each parameter is *d-used* by each equation, and each parameter instance classifies (is *valued-by*) a region instance containing the value for the parameter. The variables are also realized as qualities of geoscience particulars in associated models.
- Each part of each empirically inferred theory is represented as an *EmpiricalRegularity* instance, because these are largely induced from empirical data and are local to Earth situations (the fixed values are calibrated to the Earth).
- The parts of the theories that were derived from other theories play a *Proposition* science role, because these are theoretically postulated (hence via abduction).
- Some theories share common parts which play different roles amongst the theories, e.g. a theory developed by [32] adopts a *Proposition* from [33] as an *Assumption.*
- Many of the theories, such as the Miami Model [34] or BIOME3 Model [31], are satisfied by well known *GeoScienceModel* instances which are either based on empirical facts (contain particulars generated from observed data) or predictions (contain particulars generated from inferences).

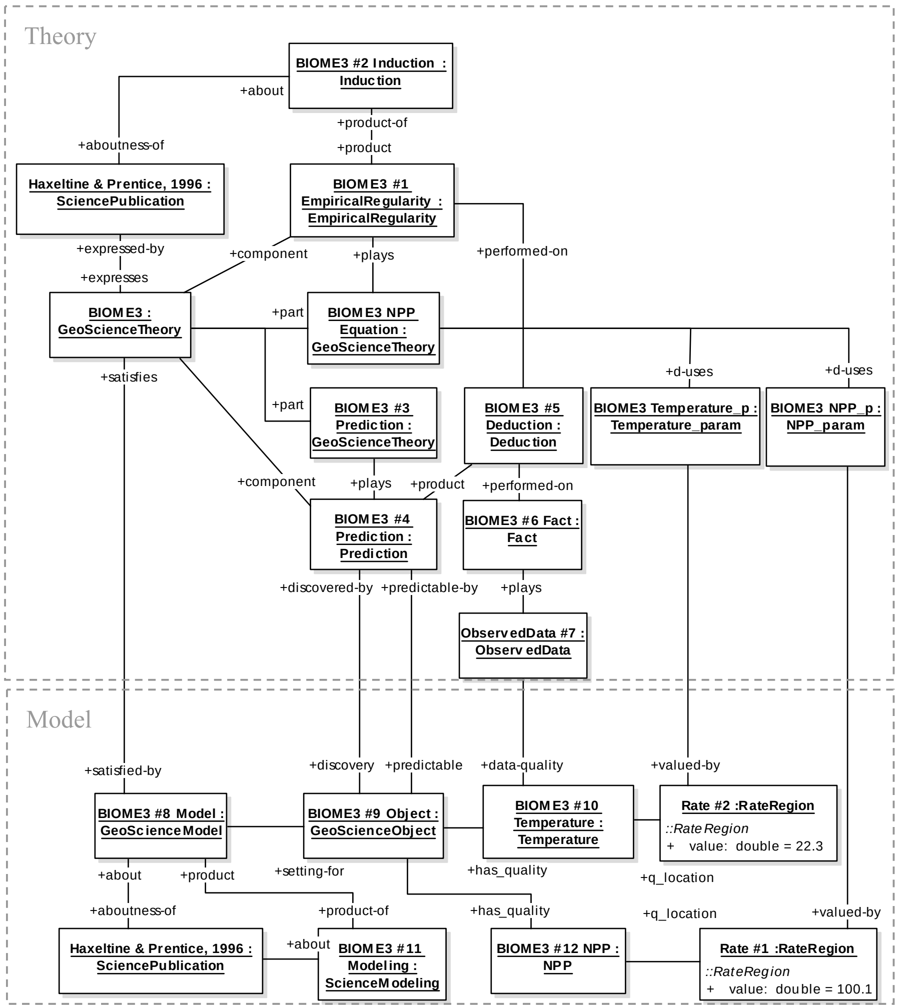

Figure 2: partial SKIo representation of the BIOME3 NPP theory [31] and a related model; in the boxes, the text before the colon denotes the instance name, and the text after the colon denotes the SKIo class name.

7. Conclusions

The SKIo ontology meets the representation and evaluation requirements outlined in Sections 2 and 4. The four senses of "model" from Section 2 are disambiguated as *ScienceTheory* (senses 1 and 2), *ScienceModeling* (sense 3), and *ScienceModel* (sense 4). The evaluation criteria from Section 4 are satisfied via: (1) modular specialization of DOLCE; (2) grounding in primitives adopted from the science knowledge cycle; (3) specialization of domain ontologies via representation of geoscience classes such as *GeoscienceModel* and *GeoScienceTheory* as well as *parameter* classes such as NPP and temperature; (4) representation of environmental knowledge from several peer reviewed papers; and (5) definition and OWL formalization of the science knowledge primitives. Future coupling of SKIo to an operational SKI should then facilitate the search, retrieval and use of basic science knowledge. For example, Jane could obtain knowledge about NPP by searching for geoscience theories that use the NPP concept. For the correlation between annual average temperature and NPP, Jane could retrieve the Miami Model [34] and the theory of [32], including a predicted model of world NPP. Consequently, we assume SKIo can be applied in other sciences and are encouraged that this will advance next-generation e-Science by facilitating not only computing operations in existing infrastructures, but also the discovery and testing of science artifacts in emerging SKIs. Limitations of SKIo include incompleteness in representing elements such as science methods and instruments, representing science knowledge change, and guides for specializing social versus physical DOLCE classes.

Acknowledgements

The authors gratefully acknowledge support of the U.K. e-Science Institute under theme 4 “Spatial Semantics for Automating Geographical Information Processes”, and the valuable comments of the three reviewers as well as F. Probst and M. Gahegan.

References

[1] Hendler, J. (2004). Science and the Semantic Web. Science 299(5606):520-521.

[2] Hars, A. (2001). Designing Scientific Knowledge Infrastructures: The Contribution of Epistemology. Information Systems Frontiers, 3(1):63-73.

[3] Gangemi, A., Guarino, N., Masolo C., Oltramari, A. (2003). Sweetening WordNet with DOLCE, AI Magazine, 24(3):13-24.

[4] Adams, B., A. White, and T. M. Lenton. (2004). An analysis of some diverse approaches to modelling terrestrial net primary productivity. *Ecological Modelling* 177 (3-4):353-391.

[5] Baldocchi, D., Finnigan, J., Wilson K., Paw U K.T., Falge E. (2000). On Measuring Net Ecosystem Carbon Exchange Over Tall Vegetation on Complex Terrain. Boundary-Layer Meteorology, 96(1-2), August 2000, pp. 257-291(35).

[6] Phillips, T. J., AchutaRao, K., Bader, D., Covey, C., Doutriaux, C. M., Fiorino, M., Gleckler, P. J., Sperber, K. R., Taylor, K. E. (2006). Coupled climate model appraisal, a benchmark for future studies. EOS Trans. AGU 87(19): 185.

[7] Shrager, J., Langley, P. (1990). Computational Models of Scientific Discovery and Theory Formation. Morgan Kaufman, San Mateo.

[8] Sowa, J. F. (2000). Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks/Cole: New York, 594 pp.

[9] Sowa, J. F. (2006). The Challenge of Knowledge Soup. In: J. Ramadas, S. Chunawala, (Eds.), Research Trends in Science, Technology and Mathematics Education, Homi Bhabha Centre, Mumbai, 2006, pp. 55-90.

[10] Thagard, P. (1988). Computational philosophy of science. MIT Press, Cambridge, MA.

[11] Ray, O. (2005). Hybrid Abductive Inductive Learning. PhD Thesis, Imperial College, University of London, http://www.bcs.org/upload/pdf/oray.pdf.

[12] Poole, D., Smyth, C., Sharma, R. (2008). Semantic Science and Machine-Accessible Scientific Theories. In: McGuinness, D.L., Fox, P., Brodaric, B. (Eds.), Semantic Scientific Knowledge Integration, AAAI Spring Symposium, Technical Report SS-08-05, AAAI Press, Menlo Park, CA, 81-86.

[13] Soldatova, L.N., King, R.D. (2006). An ontology of scientific experiments. Journal of the Royal Society Interface, 3(100):795-803.

[14] Benjamins, V. R. Fensel, D., Gómez Pérez, A. (1998). Knowledge Management through Ontologies. In Proceedings, Knowledge Engineering: Principles and Methods, PAKM-98, pp. 5.1-5.12, 1998.

[15] McCray, A. (2003). An upper-level ontology for the biomedical domain. Comparative and Functional Genomics, 4:80–84.

[16] Raskin, R.G., Pan, M.J. (2005). Knowledge representation in the Semantic Web for Earth and Environmental Terminology (SWEET). Computers and Geosciences, 31(9):1119-1125.

[17] Ludäscher, B., K. Lin, S. Bowers, E. Jaeger-Frank, B. Brodaric, C. Baru, C. (2006). Managing scientific data: From data integration to scientific workflows. In: Sinha, A.K. (Ed.) Geoinformatics: Data to Knowledge, GSA Special Paper 397, 109-130.

[18] McGuinness, D. L., Fox, P., Cinquini, L., West, P., Garcia, J., Benedict, J. L., Middleton, D. (2006). Enabling Scientific Research using an Interdisciplinary Virtual Observatory: The Virtual Solar-Terrestrial Observatory Example. AI Magazine, 29(1):65-76.

[19] Grenon, P., Smith, B. (2004). SNAP and SPAN: Towards Dynamic Spatial Ontology. Spatial Cognition and Computation, 4(1).

[20] Herre, H., Heller, B., Burek, P., Hoehndorf, R.Loebe, F., Michalek, H. (2006). General formal ontology (GFO) – part I: Basic principles. Technical Report 8, Onto-Med, University of Leipzig, 2006.

[21] Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. (2003). WonderWeb DeliverableD18, Ontology Library (final). Laboratory For Applied Ontology, http://www.loa-cnr.it/Papers/D18.pdf.

[22] Pease, A., (2006). Formal representation of concepts: The Suggested Upper Merged Ontology and its use in linguistics. In: Schalley, A. C., Zaefferer, D. (Eds.), Ontolinguistics. How Ontological Status Shapes the Linguistic Coding of Concepts. Vorbereitung Berlin:New York (Mouton de Gruyter).

[23] Bittner, T. (2007). From top-level to domain ontologies: Ecosystem classifications as a case study. In M. Duckham, B. Kuipers, L. Kulik, and S. Winter, editors, Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference, COSIT 2007, 61-77.

[24] Gangemi, A., Catenacci, C., Battaglia, M. (2004). Inflammation Ontology Design Pattern: an Exercise in Building a Core Biomedical Ontology with Descriptions and Situations. In D.M. Pisanelli (ed.) Ontolgies in Medicine, IOS Press, Amsterdam, Berlin, 2004, pp. 64-80.

[25] Grenon, G., Smith, B., Goldberg, L. (2004). Biodynamic Ontology:Applying BFO in the Biomedical Domain. In: Pisanelli, D.M. (ed.), Ontologies in Medicine, Amsterdam: IOS Press, 2004, 20–38.

[26] Brodaric, B. (2008). Science Knowledge Infrastructure Ontology 3.0. U.K. e-Science Technical Report Series, Report UKeS-2008-03, 40 pp., http://www.nesc.ac.uk/technical_papers/UKeS-2008-03.pdf.

[27] Gruber, T.R. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies. 43(5-6): 907-928.

[28] Gangemi, A., Mika, P. (2003). Understanding the Semantic Web through Descriptions and Situations. In: Meersman R, *et al*. (eds.), International Conference on Ontologies, Databases and Applications of Semantics (ODBASE 2003), Catania, (Italy), November 3-7, 2003.

[29] Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N. (2004). Social Roles and their Descriptions. In: D. Dubois, C. Welty, M.A. Williams (eds.), Proceedings of the Ninth International Conference on the Principles of Knowledge Representation and Reasoning (KR2004), Whistler, Canada, June 2-5, 2004, pp. 267-277.

[30] Kaku, M. (2004). Einstein's Cosmos: How Albert Einstein's Vision Transformed Our Understanding of Space and Time. Norton, New York.

[31] Haxeltine, A., Prentice, I.C. (1996). BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability and competition among plant functional types. Global Biogeochem. Cycles 10:4, 693–70.

[32] King, A. W., W. R. William, and W. M. Post. (1995). In search of missing carbon sink: a model of terrestrial biospheric response to land-use charge and atmosphere CO2. *Tellus. Series B* 47:501-519.

[33] Polglase, P. J., and Y. P. Wang. (1992). Potential CO2-enhanced carbon storage by the terrestrial biosphere. *Australian Journal of Botany* 40 (641-656).

[34] Leith, H. (1975) Modeling the primary productivity of the world. In *Primary Productivity of the Biosphere*, ed. H. Leith, Whittaker, R.H. New York: Springer-Verlag.