A Comparison of Upper Ontologies

Viviana Mascardi, Valentina Cordì
DISI, Università degli Studi di Genova,
Via Dodecaneso 35, 16146, Genova, Italy
E-mail: {cordi,mascardi}@disi.unige.it

Paolo Rosso
DSIC, Universidad Politécnica de Valencia,
Camino de Vera s/n, 46022, Valencia Spain
E-mail: prosso@dsic.upv.es

***Abstract*—Upper Ontologies are quickly becoming a key technology for integrating heterogeneous knowledge coming from different sources. In fact, they may be exploited as a "lingua franca" by intelligent software agents in all those scenarios where it is impossible (or there is no will) for an agent to disclose its own entire ontology to other agent, despite the need to communicate with it. This paper represents the very preliminary step towards the exploitation of Upper Ontologies as bridges for allowing intelligent software agents to align heterogeneous ontologies in an automatic way, where we analyse the most up-to-date state-of-the-art. In this paper we analyse 7 Upper Ontologies, namely BFO, Cyc, DOLCE, GFO, PROTON, Sowa's ontology, and SUMO, according to a set of standard software engineering criteria, and we synthesise our analysis in form of a comparative table. A summary of some existing comparisons drawn among subsets of the 7 Upper Ontologies that we deal with in this document, is also provided.**

I. Introduction

The increasing pressing need that human and software agents have to retrieve and exchange knowledge in a precise and efficient way, have caused ontologies, web services, and the combination of both, i.e., semantic web services, to be more and more exploited for sharing knowledge within and outside the boundaries of companies and other organisations. Intelligent software agents are recognised by both researchers and practitioners from the industry as one of the most suitable means for mediating among the heterogeneity of applications working within open, distributed, concurrent systems, and for this reason they find application in many commercial projects.

However, there are still many unsolved issues for developing and deploying multi-agent systems (MASs) in open, distributed, concurrent scenarios. One of them is how to find mappings between concepts belonging to different ontologies (in technical word, finding an *alignment* between these different ontologies) in an automatic way. We are considering the adoption of "Upper Ontologies" as bridges for making this alignment possible.

Upper ontologies are quickly becoming a key technology for integrating heterogeneous knowledge coming from different sources. In fact, they may be used by different parties involved in a knowledge integration and exchange process as a reference, common model of the reality. In particular, they may be exploited as a "lingua franca" by intelligent software agents in all those scenarios where it is impossible (or there is no will) for an agent to disclose its own entire ontology to other agent, despite the need to communicate with it.

The definition of upper ontology (also named top-level ontology, or foundation ontology) given by Wikipedia [22] is "*an attempt to create an ontology which describes very general concepts that are the same across all domains. The aim is to have a large number on ontologies accessible under this upper ontology*".

This paper represents the very preliminary step towards the exploitation of Upper Ontologies as bridges for allowing intelligent software agents to align heterogeneous ontologies in an automatic way, where we analyse the most up-to-date state-of-the-art. In fact, in this paper we have described 7 upper ontologies along different criteria that include dimension, implementation language(s), modularity, developed applications, alignment with the WordNet lexical resource, and licensing. We have chosen these criteria for three reasons:

- They are software engineering criteria useful for the developer of a knowledge-based system that has to choose the most suitable Upper Ontology for his/her needs, among a set of existing ones. Since all of us have a computer science background, these criteria are more familiar to us than philosophical ones.
- They take into account some of the evaluation questions proposed by the IEEE Standard Upper Ontology Working Group (`http://suo.ieee.org/SUO/Evaluations/`), and they also extend the criteria considered in an existing comparison among SUMO, Cyc, and DOLCE [18], thus allowing us to "reuse", and to be consistent with, the results already obtained there.
- They are not (easily) scientifically falsifiable.

The choice of the 7 upper ontologies we have described, namely BFO, Cyc, DOLCE, GFO, PROTON, Sowa's ontology, and SUMO, is based on how much, to the best of our knowledge, they are visible and used inside the research community. For example, we have discussed all the Upper Ontologies referenced by Wikipedia, apart from WordNet that we consider a lexical resource rather than an Upper Ontology, and from the Global Justice XML Data Model and National Information Exchange Model, that addresses the specific application domain of justice and public safety. We have reported alignments between the Upper Ontologies and WordNet, when existing. To the 5 Upper Ontologies considered by Wikipedia, we have added PROTON and Sowa's ontology. We have also cited three attempts to merge existing Upper Ontologies, namely COSMO, MSO, and OntoMap, although we have not described them in detail since the first two ones are still work in progress, and the last one is over since four years.

The methodology followed to draw this paper consisted in checking the existing literature, producing a first draft of the comparison based on the retrieved literature, submitting it to the attention of the developers of all the 7 upper ontologies under comparison, and integrating the obtained answers and suggestions into the current version of the paper. Due to time constraints, we were not able to experiment with the upper ontologies by our own. This "on the field" experimentation is part of our near future work.

The paper is organised in the following way: Section II provides a description of the 7 upper ontologies, and Section III surveys some existing, partial comparisons drawn in the past years among subsets of the Upper Ontologies that we describe in Section II, and provides a synthesis of the results of our comparison among them.

II. DESCRIPTION

a) Basic Formal Ontology (BFO):

- **Status of this description.** Validated by H. Stenzhorn, research associate and doctoral student at the IFOMIS and the University Hospital Freiburg - Medical Informatics Department, and one of BFO's developers.
- **Home page.** http://www.ifomis.org/bfo.
- **Developers.** B. Smith, P. Grenon, H. Stenzhorn, A. Spear (IFOMIS, Saarland University).
- **Description.** BFO consists in two sub-ontologies: SNAP – a series of snapshot ontologies (O_{ti}), indexed by times – and SPAN – a single videoscopic ontology (O_v). An O_{ti} is an inventory of all entities existing at a time, while an O_v is an inventory of all processes unfolding through time. Both types of ontology serve as basis for a series of sub-ontologies, each of which can be conceived as a window on a certain portion of reality at a given level of granularity.
- **History.** The theory behind BFO has been developed and formulated by Smith and Grenon in a series of publications starting in 1998. Its current implementation in OWL has been developed by Stenzhorn with contributions from Spear.
- **Dimensions.** BFO contains 1 top connecting class ("Entity"), 18 SNAP classes, and 17 SPAN classes for a total of 36 classes which are, in version 1.0 of the implementation, connected via the *is_a* relation. The forthcoming version of BFO will incorporate relations between classes too.
- **Implementation language(s).** OWL [21].
- **Modularity.** BFO is divided into the SNAP and SPAN modules.
- **Applications.** BFO has been applied to the biomedical domain [8] and it is currently used in building an ontology for clinic-genomic trials on cancer (http://www.acgt-eu.org).
- **Alignment with WordNet.** Not supported.
- **Licensing.** BFO is freely available; its OWL implementation may be downloaded from http://www.ifomis.org/bfo/1.0.

b) Cyc:

- **Status of this description.** Validated by L. Lefkowitz, executive director for business solutions at Cycorp.
- **Home page.** http://www.cyc.com/.
- **Developers.** Cycorp.
- **Description.** The Cyc Knowledge Base (KB) is a formalised representation of facts, rules of thumb, and heuristics for reasoning about the objects and events of everyday life. The KB consists of terms and assertions which relate those terms. These assertions include both simple ground assertions and rules. The Cyc KB is divided into thousands of "microtheories" focused on a particular domain of knowledge, a particular level of detail, a particular interval in time, etc.
- **History.** The Cyc project was founded in 1984 by D. Leant as a lead project in the Microelectronics and Computer Technology Corporation (MCC). In 1994, Cycorp was founded to further develop, commercialize, and apply the Cyc technology. Cycorp offers a no-cost license to its semantic technologies development toolkit to the research community (ResearchCyc). Additionally, it has placed the core Cyc ontology (OpenCyc) into the public domain.
- **Dimensions.** The Cyc KB (including Cyc's microtheories) contains more than 300,000 concepts and nearly 3,000,000 assertions (facts and rules), using more than 15,000 relations.
- **Implementation language(s).** Cyc is represented in the CycL formal language (http://www.cyc.com/cycdoc/ref/cycl-syntax.html). The latest release of Cyc includes an Ontology Exporter that allows to export specified portions of Cyc to OWL files.
- **Modularity.** The "microtheory" approach supports modularity.
- **Applications.** Cyc has been used in the domains of natural language processing, in particular for the tasks of word sense disambiguation [4] and question answering [5], of network risk assessment [19], and of representation of terrorism-related knowledge [6].
- **Alignment with WordNet.** The last release of Cyc (as well as of OpenCyc and ResearchCyc) includes links between Cyc concepts and about 12,000 WordNet synsets.
- **Licensing.** Cyc is a commercial product, but Cycorp also released OpenCyc (http://www.opencyc.org/), the open source version of the Cyc technology, and ResearchCyc (http://research.cyc.com/), namely the Full Cyc ontology, but with a research-only license.

c) DOLCE (a Descriptive Ontology for Linguistic and Cognitive Engineering):

- **Status of this description.** Not validated by the ontology developer(s).
- **Home page.** http://www.loa-cnr.it/DOLCE.html.

- **Developers.** Researchers from the Laboratory for Applied Ontology, headed by N. Guarino.
- **Description.** DOLCE is the first module of the WonderWeb Foundational Ontologies Library. DOLCE has a clear cognitive bias, in the sense that it aims at capturing the ontological categories underlying natural language and human commonsense. According to DOLCE, different entities can be co-located in the same space-time. DOLCE is described by its authors as an "ontology of particulars", which the authors explain as meaning an ontology of instances, rather than an ontology of universals or properties. The taxonomy of the most basic categories of particulars assumed in DOLCE includes, for example, abstract quality, abstract region, agentive physical object, amount of matter, non-agentive physical object, physical quality, physical region, process, temporal quality, temporal region.
- **History.** DOLCE has been developed as part of WonderWeb, a project funded as a shared-cost RTD under the European Commission information society technologies (IST) programme. WonderWeb started in 2002 and ended in 2004. Although the project has already ended, DOLCE is actively maintained and used.
- **Dimensions.** Around one hundred of terms, and a similar number of axioms.
- **Implementation language(s).** First Order Logic, KIF [1], OWL.
- **Modularity.** The intended use of DOLCE is within a modular library of foundational ontologies, but it is not currently divided into modules.
- **Applications.** According to the "DOLCE around the world" web page (http://www.loa-cnr.it/dolcevar.html), there are many projects that use DOLCE, including the LOIS Project – an international research project on multilingual information retrieval from legal databases –, SmartWeb – a centre of excellence in research on intelligent computing technologies and their application to web-based systems and services –, Language Technology for eLearning – a project funded by the EC, and using multilingual language technology tools and semantic web techniques for improving the retrieval of learning material –, AsIsKnown – a semantic-based knowledge flow system for the European home textiles industry, also funded by the EC –, and the Projects of the Laboratory for Applied Ontology.
- **Alignment with WordNet.** The OntoWordNet Project aims at aligning the top-level of WordNet to DOLCE, in order to obtain an "ontologically sweetened" lexical resource, meant to be conceptually more rigorous, cognitively transparent, and efficiently exploitable in several applications. The beta version (v0.72) of the OWL alignment of WordNet 1.6 Noun Synsets to the DOLCE-Lite-Plus ontology library consists of an alignment between DOLCE-Lite-Plus and about 100 Wordnet sysnsets, and can be downloaded from http://www.loa-cnr.it/ontologies/OWN/OWN.owl.
- **Licensing.** The OWL version of DOLCE can be freely downloaded from http://www.loa-cnr.it/ontologies/DLP3971.zip.

d) GFO (General Formal Ontology):

- **Status of this description.** Validated by F. Loebe, PhD student at the University of Leipzig under the supervision of H. Herre and M. Löffler, members of the scientific board of Onto-Med.
- **Home page.** http://www.onto-med.de/ontologies/gfo.html.
- **Developers.** The Onto-Med Research Group (http://www.onto-med.de/).
- **Description.** GFO includes elaborations of categories like objects, processes, time and space, properties, relations, roles, functions, facts, and situations. Work is in progress on an integration with the notion of levels of reality in order to more appropriately capture entities in the material, mental, and social areas.
- **History.** Work on GFO has started in 1999 in the context of the GOL project (General Ontological Language). Meanwhile, several directions of research have been recognised and divided the initial project, such that GFO is now one component of a larger framework. Work on GFO remains in progress, because the development of top-level ontologies is a long-term research effort.
- **Dimensions.** The OWL version of GFO consists of 79 classes, 97 subclass-relations, and 67 properties.
- **Implementation language(s).** The FOL axiomatization of GFO and a KIF implementation of it are forthcoming. An OWL-DL version also exists.
- **Modularity.** GFO exhibits a three-layered meta-ontological architecture consisting of an abstract top level, an abstract core level, and a basic level. The foundational ontology GFO is structured into several ontological modules including a module for functions and a module for roles.
- **Applications.** One of the aims of the group Onto-Med is the application of the GFO in the field of biomedical science. GFO has been used to represent knowledge about biological functions in the Gene Ontology, the Celltype Ontology, and the Chemical Entities of Biological Interest (ChEBI) Ontology [2], and GFO-Bio (http://onto.eva.mpg.de/gfo-bio.html) is based on GFO and is a core ontology for biology. Another area of application is the ontological foundation of conceptual modelling. First examples of applying GFO to UML are demonstrated in [9].
- **Alignment with WordNet.** Not supported.
- **Licensing.** The OWL version of GFO is released under the modified BSD Licence (http://www.opensource.org/licenses/bsd-license.php) and can be downloaded from http://www.onto-med.de/ontologies/gfo.owl.

e) PROTON (PROTo ONtology):

- **Status of this description.** Validated by A. Kiryakov, head of Ontotext Lab, member of the board.
- **Home page.** http://proton.semanticweb.org/

- **Developers.** Ontotext Lab, Sirma (http://www.ontotext.com/).
- **Description.** PROTON (PROTo ONtology) is a basic upper-level ontology providing coverage of the general concepts necessary for a wide range of tasks, including semantic annotation, indexing, and retrieval of documents. The design principles can be summarized as follows (i) domain-independence; (ii) light-weight logical definitions; (iii) alignment with popular standards; (iv) good coverage of named entities and concrete domains (i.e. people, organizations, locations, numbers, dates, addresses).
- **History.** PROTON has been developed in the scope of SEKT, a project co-funded by the EU 6th Framework programme. SEKT started the 1st of January, 2004 and will conclude at the end of 2006. PROTON is a development of the KIMO ontology, which had been created and used in the scope of the KIM platform for semantic annotation, indexing, and retrieval (http://www.ontotext.com/kim/). Currently, KIMO does not exist anymore; it is replaced by PROTON, KIMLO (http://www.ontotext.com/kim/2005/04/kimlo#) and KIMSO (http://www.ontotext.com/kim/2005/04/kimso#).
- **Dimensions.** PROTON contains about 300 classes and 100 properties.
- **Implementation language(s)** A fragment of OWL Lite.
- **Modularity.** PROTON is organized in three levels including four modules.
  The System module ontology module occupies the first, basic layer. It defines several notions and concepts of a technical nature that are substantial for the operation of any ontology-based software, such as semantic annotation and knowledge access tools. The Top ontology module occupies the second layer and includes basic philosophically-reasoned distinctions between entity types, such as Object, Happening, Abstract. Further uplevel, PROTON extends into its third layer, where either of two independent ontologies, which defines much more specific classes, can be used: PROTON Upper module or PROTON KM (Knowledge Management) module. Examples of concepts belonging to these modules are Mountain, as a specific type of Location, and ResourceCollection as a sub-class of InformationResource.
- **Applications.** As witnessed by a large number of publications (http://www.ontotext.com/publications/), PROTON has been used in different domains and for different purposes, including semantic annotation within the KIM platform, and knowledge management systems in legal and telecommunications domain [3]. It has also been used as a basis for a domain ontologies in media research and analysis (project MediaCampaign) and research intelligence (project IST World), and a basis for Business Data Ontology for Semantic Web Services [13].
- **Alignment with WordNet.** Not supported.
- **Licensing.** The four modules that compose PROTON are freely accessible via Web: System module (http://proton.semanticweb.org/2005/04/protons); Top module (http://proton.semanticweb.org/2005/04/protont); Upper module (http://proton.semanticweb.org/2005/04/protonu); Knowledge Management module (http://proton.semanticweb.org/2005/04/protonkm).

f) Sowa's Ontology:

- **Status of this description.** Not validated by the ontology developer(s).
- **Home page.** http://www.jfsowa.com/ontology/.
- **Developers.** J. F. Sowa.
- **Description.** Sowa's ontology is based on [20]. The basic categories and distinctions have been derived from a variety of sources in logic, linguistics, philosophy, and artificial intelligence. To keep the system open-ended, Sowa's ontology is not based on a fixed hierarchy of categories, but on a framework of distinctions, from which the hierarchy is generated automatically. For any particular application, the categories are not defined by drawing lines on a chart, but by selecting an appropriate set of distinctions. These categories include Object, Process, Schema, Script, Juncture, Participation, Description, History, Structure, Situation, Reason, and Purpose. Each of these categories may be either Physical or Abstract (and in both cases, it can be either Continuant or Occurrent), and it may also be either Independent, Relative, or Mediating. For example, Process is Physical, Occurrent and Independent.
- **History.** Sowa's ontology dates back to 1999. The two major influences on it are the semiotics of C. Sanders Peirce and the categories of existence of A. North Whitehead.
- **Dimensions.** The KIF encoding of Sowa's ontology contains about 30 classes, 5 relationships among classes, and among classes and instances (has, instance-of, subclass-of, temp-part-of, spatial-part-of), about 30 axioms.
- **Implementation language(s).** Sowa's ontology uses a first-order modal language, i.e., a first-order language with the modal operators "nec" and "poss". A version written in KIF also exists.
- **Modularity.** Sowa's ontology is not explicitly divided into modules, although each of the top level categories can be intended as a module by its own, connected to the other ones by means of relations.
- **Applications.** Sowa's ontology inspired many existing implemented upper ontologies, and thus its exploitation in the development of "second-generation" upper ontologies may be seen as one, and probably the most relevant, of its practical applications.
- **Alignment with WordNet.** Not supported.
- **Licensing.** The KIF encoding of Sowa's upper ontology can be freely downloaded from http://suo.ieee.org/SUO/ontologies/Sowa.txt.

g) SUMO (Suggested Upper Merged Ontology):

- **Status of this description.** Validated by A. Pease, current Technical Editor of SUMO.

- **Home page.** http://www.ontologyportal.org/.
- **Developers.** The SUMO starter document was created at Teknowledge by I. Niles and A. Pease, with a contribution by C. Menzel.
- **Description.** SUMO and its domain ontologies [14] form one of the largest formal public ontology in existence today. They are being used for research and applications in search, linguistics and reasoning. SUMO is extended with many domain ontologies and a complete set of links to WordNet, and is freely available.
- **History.** SUMO was first released in December 2000. It was created at Teknowledge Corporation and it was proposed as a starter document for the Standard Upper Ontology Working Group (http://suo.ieee.org/), an IEEE-sanctioned working group of collaborators from the fields of engineering, philosophy, and information science. SUMO was created by merging publicly available ontological content into a single, comprehensive, and cohesive structure. This content included the ontologies available on the Ontolingua server (http://www.ksl.stanford.edu/software/ontolingua/), Sowa's upper level ontology, and various mereotopological theories, among other sources.
- **Dimensions.** SUMO contains about 1000 terms and 4000 axioms; if we consider also the terms and axioms of its domain ontologies, however, it reaches the dimension of 20,000 terms and 60,000 axioms.
- **Implementation language(s).** The first-order logic language SUO-KIF (http://suo.ieee.org/SUO/KIF/suo-kif.html), OWL.
- **Modularity.** SUMO consists of SUMO itself (the official latest version on the IEEE web site can be downloaded from http://suo.ieee.org/SUO/SUMO/SUMO_173.kif), the MId-Level Ontology (MILO), and ontologies of Communications, Countries and Regions, Distributed Computing, Economy, Finance, Engineering Components, Geography, Government, Military, North American Industrial Classification System, People, Physical Elements, Transnational Issues, Transportation, Viruses, World Airports. Additional ontologies of terrorism are available on request.
- **Applications.** The applications of SUMO are documented by the almost one hundred published papers describing its use (http://www.ontologyportal.org/Pubs.html). The largest user community is in linguistics, but other classes of applications include "pure" representation, and reasoning. Applications range from academic to government, to industrial ones.
- **Alignment with WordNet.** SUMO has been mapped to all of Wordnet v2.1 by hand. The mappings can be downloaded from `http://sigmakee.cvs.source-forge.net/sigmakee/KBs/WordNetMappings/`.
- **Licensing.** SUMO is free and owned by the IEEE. Its SUO-KIF implementation can be downloaded from `http://sigmakee.cvs.sourceforge.net/*check-out*/sigmakee/KBs/Merge.kif`, while the OWL implementation can be downloaded from http://www.ontologyportal.org/translations/SUMO.owl.txt. The ontologies that extend SUMO are available under GNU General Public License.

h) Merging Upper Level Ontologies.: Three attempts to merge some of the upper level ontologies, thus leading to an "upper-upper level ontology", are COSMO (COmmon Semantic MOdel, `http://colab.cim3.net/cgi-bin/wiki.pl?CosmoWG/-TopLevel`), MSO (Multi-Source Ontology, http://www.webkb.org/doc/MSO.html), and the OntoMap Project [11].

COSMO results from the efforts of the COSMO working group (COSMO-WG) and its parent group, the Ontology and Taxonomy Coordinating Working Group (ONTACWG). COSMO is viewed as consisting of a lattice of ontologies which will serve as a set of basic logically-specified concepts (classes, relations, functions, instances) with which the meanings of all terms and concepts in domain ontologies can be specified. The use of a common set of defining concepts will permit accurate interoperability of knowledge-based systems using the logical relations of their ontologies as the basis for reasoning in the system. Currently, COSMO integrates concepts from the OpenCyc and SUMO ontologies, with some classes from DOLCE and BFO. The work on COSMO is in progress.

MSO is the Multi-Source Ontology of WebKB-2, a knowledge server that permits Web users to browse and update private knowledge bases on their machines, or alternatively, a large shared knowledge base on the server machine. The ontology of the shared knowledge base is currently an integration of various top-level ontologies and a lexical ontology derived from an extension and correction of the noun-related part of WordNet 1.7. The semantics of some categories from WordNet has been modified in order to fix inconsistencies, while the semantics of categories from other sources (e.g. Sowa, DOLCE) has been kept. Also the work regarding the MSO is still in progress. In particular, the integration of the SUMO is still far from being complete. This integration links the SUMO categories to the existing categories of the MSO, adds some structure when needed, adds equivalent categories the names of which are better suited for knowledge representation conventions that are "common" in the communities using graph-based or frame-based notations, and finally translates the axioms from KIF to more intuitive notations that permit people to more easily understand the meanings of the categories and their relationships.

Finally, OntoMap was a project with the goal to facilitate the access, understanding, and reuse of such resources. A semantic framework on conceptual level was implemented that was small and easy enough to be learned on-the-fly. Technically, OntoMap was implemented as a web-site providing access to several upper-level ontologies and manual mapping between them. OntoMap was similar in spirit to COSMO and MSO, but only the very top concepts of each of the Upper Ontologies considered there were aligned. Unfortunately, OntoMap was over 4 years ago, and no maintenance was guaranteed to it. The web-portal which was allowing online browsing is no longer available, but the stand-alone viewer may be downloaded from http://www.ontotext.com/projects/OntoMapViewer/install.htm.

III. Comparison

Some partial comparisons exist among subsets of the Upper Ontologies that we have considered in Section II. In the next paragraphs, we have summarised them in the most faithful way. The interested reader should go to the source, always cited, in order to have a complete picture of the conclusions reached by the comparisons' authors. The last paragraph, instead, provides a synthesis of the description we have given in Section II.

i) Pease's comparison of DOLCE and SUMO.: In [15] and [16], Pease draws a comparison between DOLCE and SUMO. His conclusions are that DOLCE has a similar purpose and business process to SUMO in that it is a free research project for use in both natural language tasks and inference. DOLCE has been carefully crafted with respect to strong principles. DOLCE is an "ontology of particulars"; it does have universals (classes and properties), but the claim is that they are only employed in the service of describing particulars. In contrast, SUMO could be described as an ontology of both particulars and universals. It has a hierarchy of properties as well as classes. This is a very important feature for practical knowledge engineering, as it allows common features like transitivity to be applied to a set of properties, with an axiom that is written once and inherited by those properties, rather than having to be rewritten, specific to each property. Other differences include DOLCE's use of a set of meta-properties as a guiding methodology, as opposed to SUMO's use and formal definition of such meta-properties directly in the ontology itself. With respect to SUMO, DOLCE does not include such items as a hierarchy of process types, physical objects, organisms, units and measures, and event roles.

j) Onto-Med's comparison of GFO, DOLCE, and Sowa's ontology.: In [10], informal mappings from GFO to DOLCE and from GFO to Sowa's ontology, and viceversa, are specified. The authors of the comparison observe that all of Sowa's categories except for three can be reinterpreted in GFO. However, mapping in the opposite direction seems to be more problematic. For many of GFO categories, the corresponding notions in Sowa's ontology has not been found. Neither a space-time model nor a property model is included in Sowa's ontology, and the construction method of GFO is not as strictly combinatorial as is Sowa's ontology. In DOLCE, levels of reality are not introduced explicitly, while in GFO the authors explicitly distinguish three levels of reality. Universals are excluded from DOLCE, which supports neither the distinctions provided in GFO concerning sets and items, nor concerning the typology of categories. A time or a space model is not built directly into DOLCE. Instead, the representation of various models of space and time is permitted, which can be introduced by means of qualities and their associated "qualia" (the latter are similar to GFO's quality values). In the GFO, spatial location is modelled in terms of spatial regions and relations, like occupation and location; temporal location is based on time regions and projection relations. In addition, presently the GFO provides a model for time and space. The DOLCE distinction between endurant and perdurant is based on the behavior of entities in time. Endurants are entities that can change in time, are wholly present at any time of their existence, and have no temporal parts but their parts are time-indexed, and participate in perdurants. GFO distinguishes between persistence through time and being wholly present at a time-boundary. This has produced two GFO categories instead of endurant alone: persistants and presentials. GFO persistant refers to the idea of persistence through time as attributed to DOLCE's endurant, although persistants are not considered in GFO as individuals but as universals[1]. GFO presentials can be generally interpreted as DOLCE endurants, but without temporal extension. Intuitively, DOLCE notion of perdurant corresponds to GFO notion of occurrent. Moreover, it seems that the GFO notions of process, state and change can be interpreted in DOLCE as stative, state and event, respectively. Finally, the GFO categories that concern properties and their values correspond rather well to DOLCE qualities, qualia and quality spaces.

k) MITRE's comparison of SUMO, Upper Cyc, and DOLCE.: In [18], Semy, Pulvermacher and Obrst compare SUMO, Upper Cyc, and DOLCE according to the existence of an open license, modularity and evidence of use. We have adopted these criteria inside our analysis, which thus subsumes Semy, Pulvermacher and Obrst's one.

l) Grenon's comparison of DOLCE and BFO.: Grenon made a careful comparison between DOLCE and BFO [7]. The conclusion is that both ontologies contain a category of endurants and perdurants and an eternalist stance, and that the theory of parthood and the theory of dependence are similar in the two ontologies. Despite these similarities, there are also many differences, including:

- DOLCE is methodologically fundamentally conceptualist while BFO is methodologically fundamentally realist;
- DOLCE seems to be oriented toward commonsense, and BFO's naïve realism is in the same spirit. However, DOLCE distinguishes between abstract and concrete entities, and it includes agents and intentionality. BFO is deliberately not committed to these distinctions. In particular, the physical / non-physical endurants distinction in DOLCE is absent in BFO.
- As already mentioned, DOLCE is intended as an ontology of particulars. BFO is intended to be an ontology of both universals and particulars.
- In DOLCE, qualities are abstract entities which may not be found in space or time, and do not have parts. For BFO, the proxies of DOLCE's qualities ("tropes") are located in space and exist at a time in the very same way that the entities in which they inhere.

[1] The forthcoming release of GFO, expected by early 2007, will include some refinements of the notion of persistence which will make this statement no longer valid.

Another source of information about the similarities and differences between DOLCE and BFO is [12], where Masolo, Borgo, Gangemi, Guarino, and Oltramari of the Laboratory For Applied Ontology (LOA) compare DOLCE and BFO (besides the OCHRE object-centered ontology, [17], that we did not consider in our analysis) by representing the assertion "A statute of clay exists for a period of time going from t_1 to t_2. Between t_2 and t_3, the statue is crashed and so ceases to exists although the clay is still there." in both of them.

m) Other existing sources of comparison.: Evaluations of three Candidate Common Upper Ontologies, including SUMO and MSO, can be found at http://suo.ieee.org/SUO/Evaluations/. The criteria considered there include maturity, robustness, potential for broad acceptance, language flexibility, ownership/cost, and domain friendliness. These evaluations are not comparative: each Upper Ontology is evaluated (usually, by its creator) according to the above metrics.

n) Our comparison.: The description of the 7 Upper Ontologies given in Section II is synthesised here in Tables I and II.

IV. Conclusions

This paper represents a preliminary step towards the exploitation of upper ontologies as the means for allowing intelligent software agents to integrate heterogeneous sources of information, respecting privacy issues that are more and more commong in many scenarios, such as virtual enterprises and e-commerce. In fact, this paper provides an original and unpublished analysis of the state-of-the-art in the field of upper ontologies. This analysis is a necessary activity before starting to think how upper ontologies may be actually exploited as a bridge among two or more ontologies to be integrated. If the original ontologies cannot be disclosed for privacy issues, each agent involved in the application and responsible for accessing and integrating one ontology, may "align" (i.e. find a mapping between concepts) its own, private ontology, with the upper ontology, and refer to the latter one in all its communicative acts. At the time of writing, the design of an algorithm for aligning ontologies using upper ontologies as a bridge is under way. As soon as we will be able to implement and test it, we will obtain results that will give us an important help in understanding under which conditions the exploitation of upper ontologies is feasible, and which upper ontologies are better for being used as a bridge in the alignment process. Our current and future work is entirely aimed at completing the design and implementation of the algorithm and systematically describing our experimental results.

Acknowledgments

We want to acknowledge all the researchers that helped in drawing this comparison with their constructive comments and useful advices. In particular, many thanks go to J. Euzenat, A. Kiryakov, L. Lefkowitz, F. Loebe, A. Pease, J. Schoening, P. Shvaiko, and H. Stenzhorn.

We also acknowledge the research projects TIN2006-15265-C06-04 and "Iniziativa Software" CINI-FINMECCANICA that partially funded this work.

References

[1] American National Standard. KIF Knowledge Interchange Format – draft proposed American National Standard (dpANS) NCITS.T2/98-004, 1998.

[2] P. Burek, R. Hoehndorf, F. Loebe, J. Visagie, H. Herre, and J. Kelso. A top-level ontology of functions and its application in the open biomedical ontologies. In *ISMB (Supplement of Bioinformatics)*, volume 22, pages 66–73, 2006.

[3] N. Casellas, M. Blázquez, A. Kiryakov, P. Casanovas, M. Poblet, and R. Benjamins. OPJK into PROTON: Legal domain ontology integration into an upper-level ontology. In R. Meersman and et al., editors, *Proceedings of the 3rd International Workshop on Regulatory Ontologies (WORM 2005)*, volume 3762 of *Lecture Notes in Computer Science*, pages 846–855. Springer, 2005.

[4] J. Curtis, D. Baxter, and J. Cabral. On the application of the Cyc ontology to word sense disambiguation. In *Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference*, pages 652–657, 2006.

[5] J. Curtis, G. Matthews, and D. Baxter. On the effective use of Cyc in a question answering system. In *Proceedings of the IJCAI Workshop on Knowledge and Reasoning for Answering Questions (KRAQ'05)*, 2005.

[6] C. Deaton, B. Shepard, C. Klein, C. Mayans, B. Summers, A. Brusseau, and M. Witbrock. The comprehensive terrorism knowledge base in Cyc. In *Proceedings of the 2005 International Conference on Intelligence Analysis*, 2005.

[7] P. Grenon. BFO in a nutshell: A bi-categorial axiomatization of BFO and comparison with DOLCE. Technical Report 06/2003, IFOMIS, University of Leipzig, 2003.

[8] P. Grenon, B. Smith, and L. Goldberg. Biodynamic ontology: Applying BFO in the biomedical domain. In D. M. Pisanelli, editor, *Ontologies in Medicine*, volume 102 of *Studies in Health Technology and Informatics*, pages 20–38. IOS Press, 2004.

[9] G. Guizzardi, H. Herre, and G. Wagner. On the general ontological foundations of conceptual modeling. In S. Spaccapietra, S. T. March, and Y. Kambayashi, editors, *Proceedings of the 21st International Conference on Conceptual Modeling (ER 2002)*, volume 2503 of *Lecture Notes in Computer Science*, pages 65–78. Springer, 2002.

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

[11] A. Kiryakov, K. Ivanov Simov, and M. Dimitrov. OntoMap: portal for upper-level ontologies. In *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS 2001)*, pages 47–58. ACM, 2001.

[12] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. Ontology library (final). IST Project 2001-

33052 WonderWeb Deliverable D18, 2003.

[13] G. Nagypál and J. Lemcke. A business data ontology. Data, Information and Process Integration with Semantic Web Services Project, FP6 Ű 507483, Deliverable D3.3, 2005.

[14] I. Niles and A. Pease. Towards a standard upper ontology. In C. Welty and B. Smith, editors, *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)*, pages 2–9. ACM Press, 2001.

[15] A. Pease. Formal representation of concepts: The Suggested Upper Merged Ontology and its use in linguistics. In A. C. Schalley and D. Zaefferer, editors, *Ontolinguistics. How Ontological Status Shapes the Linguistic Coding of Concepts*. Mouton de Gruyter, 2006.

[16] A. Pease and C. Fellbaum. Formal ontology as interlingua: The SUMO and WordNet linking project and GlobalWordNet. To appear.

[17] L. Schneider. How to build a foundational ontology: The object-centered high-level reference ontology OCHRE. In A. Günter, R. Kruse, and B. Neumann, editors, *Proceedings of the 26th Annual German Conference on AI, KI 2003: Advances in Artificial Intelligence*, volume 2821 of *Lecture Notes in Computer Science*, pages 120–134. Springer, 2003.

[18] S. K. Semy, M. K. Pulvermacher, and L. J. Obrst. Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Technical Report MTR 04B0000063, The MITRE Corporation, 2004.

[19] B. Shepard, C. Matuszek, C. B. Fraser, W. Wechtenhiser, D. Crabbe, Z. Güngördü, J. Jantos, T. Hughes, L. Lefkowitz, M. J. Witbrock, D. B. Lenat, and E. Larson. A knowledge-based approach to network security: Applying Cyc in the domain of network risk assessment. In M. M. Veloso and S. Kambhampati, editors, *Proceedings of the 20th National Conference on Artificial Intelligence and the 17th Innovative Applications of Artificial Intelligence Conference*, pages 1563–1568. AAAI Press AAAI Press / The MIT Press, 2005.

[20] J. F. Sowa. In *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing, 1999.

[21] W3C. OWL Web Ontology Language Overview – W3C Recommendation 10 February 2004 , 2004.

[22] Wikipedia. Upper ontology – wikipedia, the free encyclopedia, 2006. [Online; accessed 15-December-2006].

	Home page	Developers	Dimensions	Language(s)
BFO	http://www.ifomis.org/bfo	Smith, Grenon, Stenzhorn, Spear (IFOMIS)	36 classes related via the is_a relation	OWL
Cyc	http://www.cyc.com/	Cycorp	About 300,000 concepts, 3,000,000 assertions (facts and rules), 15,000 relations (these numbers include microtheories)	CycL, OWL
DOLCE	http://www.loa-cnr.it/DOLCE.html	Guarino and other researchers of the LOA	About 100 concepts and 100 axioms	First Order Logic, KIF, OWL
GFO	http://www.onto-med.de/ontologies/gfo.html	The Onto-Med Research Group	79 classes, 97 subclass-relations, 67 properties	First Order Logic and KIF (forthcoming); OWL
PROTON	http://proton.semanticweb.org/	Ontotext Lab, Sirma	300 concepts and 100 properties	OWL Lite
Sowa's	http://www.jfsowa.com/ontology/	Sowa	30 classes, 5 relationships, 30 axioms	First Order Modal Language, KIF
SUMO	http://www.ontologyportal.org/	Niles, Pease, and Menzel	20,000 terms and 60,000 axioms (including domain ontologies)	SUO-KIF, OWL

Table I
COMPARISON, PART I

	Modularity	**Applications**	**Alignment with WordNet**	**Licensing**
BFO	SNAP and SPAN modules	Mainly in the biomedical domain	Not supported	Freely available
Cyc	“Microtheory” modules	Natural language processing, network risk assessment, terrorism management	Cyc is mapped to about 12,000 WordNet synsets	Commercial product; ResearchCyc and OpenCyc are instead freely available (ResearchCyc for research purposes only)
DOLCE	Not divided into modules	Multilingual information retrieval, web-based systems and services, e-learning	DOLCE-Lite-Plus has been aligned with about 100 Wordnet sysnsets	Freely available
GFO	Abstract top level, abstract core level, basic level	Mainly in the biomedical domain	Not supported	Released under the modified BSD Licence
PROTON	Three levels including four modules	Semantic annotation within the KIM platform, knowledge management systems in legal and telecommunications domain, media research and analysis, research intelligence, Business Data Ontology for Semantic Web Services.	Not supported	Freely available
Sowa’s	Not divided into modules	No documented applications have been developed, but Sowa’s ontology inspired the creation of many implemented Upper Ontologies	Not supported	Freely available
SUMO	Divided into SUMO itself, MILO, and domain ontologies	Linguistics, representation, reasoning	SUMO has been mapped to all of Wordnet v2.1 by hand	Freely available

Table II
COMPARISON, PART II