Semantic Web Rules
Use Cases and Requirements for Health Care and Life Sciences

Christine Golbreich
Christine.Golbreich@univ-rennes1.fr
Draft
6 December 2005

This paper suggests some possible Use Cases and Requirements in the domain of *Health Care and Life Sciences*. Most of them are based on our experience with real applications in the biomedical domain. The list below gives a short description of the use cases and their implications in terms of requirements for a Web rule language for the W3C RIF WG. The first section (§1) concerns the basic requirement of *compatibility* with OWL. Section 2 lists different types of rule uses (§2). The use cases description follows in sections (§3- 4 -5). While Section 3 mainly focuses on the requirement of *close and safe integration*, the following cases illustrate other uses of rules in the medical domain. The title of each use case indicates its main implications in terms of a Web rule language.

1. Compatibility

1.1. [Use case] CG-1: A Web rule extension compatible with OWL DL

- Outline: library of huge biomedical ontologies in OWL e.g. the FMA [17].

Life Sciences have a long tradition of controlled vocabularies. Large-scale terminologies, classifications and ontologies have been developed for many years in various biomedical domains. These resources have the potential to contribute to the Semantic Web for the Life Sciences. Several actual biomedical ontologies, e.g. The Foundational Model of Anatomy (FMA), the Medical Subject Headings (MeSH), the Gene Ontology™, the National Cancer Institute Thesaurus are converted to OWL. The conversion of other ontologies to OWL is also been investigated, including the UMLS® Metathesaurus® and Semantic Network and other ones e.g., SNOMED-CT, GALEN are represented in other DLs. New biomedical ontologies are now directly developed in OWL e.g.; BioPax an ontology for biological pathway information. For example, the FMA is the most complete ontology of human canonical anatomy. It contains more than *72,000 concepts and more than two million instantiations of 150 relations. 2/3* of the FMA (i.e. a subset of 40000 concepts) including about 40000 subclass axioms, with existential and union in rhs and other OWL constructors has been represented in OWL DL [17]. The FMA is used in several actual applications e.g. the Virtual Soldier project Virtual Soldier project (http://www.virtualsoldier.net/) [16]. This use case is connected to [CG7] (5.4) and [CG3] (3.2)

Implications:

- DLs more suited for structural knowledge (ontology) than LP language
- Impossible to use LP rules for representing large-scale biomedical ontologies
- Interoperability between ontologies through OWL standard is necessary to allow reasoning across connected domains e.g. pathology, genes, anatomy.
- As most of the ontologies use union or existential in rhs, OWL-DLP expressiveness is not enough

- OWL DL expressiveness is required.
- OWL DL reasoining services (consistency checking, classification) are crucial for the *quality assurance* of such large-scale biomedical ontologies.

In conclusion, many biomedical ontologies are represented in OWL DL, a Web rule language should be compatible with OWL DL.

2. Different uses of rules

Rules are required for different tasks [1]:

(1) "deductive rules" are needed for inferences based on dependencies between some ontology properties, such as the transfer of properties from parts to wholes (e.g. a disease located in an organ part, is located in the organ), or dependencies between topological and mereological properties in the brain-cortex [4] http://idm.univ-rennes1.fr/~obierlai/anatomy/annexes/annexes.pdf ([CG-3] 3.2) For a long time, rule-based expert systems have shown the usefulness of deductive rules in health care e.g. for diagnosis ([CG-6]- 5.2), guidelines (CG-7] 5.3) etc.

(2) "meta-reasoning rules" are needed for facilitating meta-reasoning on ontology in control or knowledge engineering tasks e.g., acquisition, validation, maintenance of an ontology [1] [10] ([CG-8] 5.4)

(3) "connecting rules" between ontologies are required for reasoning across several domains such as Genomics, Proteonomics, Pathology, for example when searching for correlations between diseases and the abnormality of a function of a protein coded by a human gene.

(4) "mapping rules" for mapping ontologies in data integration, and querying heterogeneous sources e.g., patient data scattered in many Hospital Information Systems [11 -13-14-15] ([CG-5] 5.1)

(5) "querying rules" expressing complex queries upon the Web or sources ([CG-5] 5.1) etc.

The Use Cases description and their requirements follow (more detailed description are given for each use case in the referred papers). Whatever the type of rules concerned, all the use cases need reasoning on ontology and rules. Most of them require a Web rule language allowing a close integration between ontology and rules.

3. Integration

[Use Case] CG2 - CG3: Close integration between the ontology and rule language allowing safe reasoning

The first use-case illustrates a "mapping" approach between the SWRL extension and the LP language Jess on a simple example: the Family use-case, and its limitation. The second use-case illustrates the requirements of a close integration between the ontology and the rule language on a real application: Brain Anatomy digitally processed images.

3.1. [Use Case] CG-2 limitations of a "mapping" approach

- **Outline: The Family rules example and the SWRLJessTab Protégé plugin [2] [3]**

The Family rules example is a simple example including an ontology representing the usual family relationships, and a rule base representing their dependencies.

The SWRLJessTab Protégé plugin presented at the 7th Protégé 2004 Conference and the RuleML WS [2] [3] illustrates on the family example how some reasoning support might be provided to interoperate between SWRL and OWL, and why such loose interoperation between ontology and rules is not satisfying. SWRLJessTab, was intended to bridge between OWL, SWRL, Racer and Jess, for supporting reasoning with SWRL. The approach is illustrated on the 'family' SWRL rules (Rbox) and ontology (Tbox and Abox). Concepts individuals and role instances of the Abox are *mapped* to Jess facts. SWRL rules are *mapped* to Jess rules. After the Jess engine execution, the new Jess facts derived from Jess knowledge are converted into assertions on concepts and roles in Racer Abox by an inverse mapping process. Iterative calls of Racer and Jess are done until an inconsistency is detected or no new fact is inferred. The family use-case illustrates that, if reasoning is performed *separately* by an OWL reasoner and a rule engine, some inferences are obviously missed. Some results can even be false, when the OWL ontology and SWRL rules components are not closely integrated. The Protégé environment including a SWRL Editor and a plugin mechanism for integrating third party rule follows a similar "mapping" approach. [7] [6]. However, for the same reasons, this loose integration has similar limitations. First, the rule inferences are based only on the rule component, since the OWL ontology (Tbox) component is not integrated into the rule (Jess) knowledge base. As Jess knowledge base is incomplete, some inferences are inevitably missed. In fact, as DL and LP languages have basically different expressiveness and different semantics, a DL ontology e.g., the family ontology, *cannot* be converted into an equivalent Jess knowledge base, unless serious restrictions of the ontology sublanguage e.g., to DLP fragment of OWL [8]. Second, even under DLP restrictions, the reasoning method consisting in deriving all consequences of a DL reasoner e.g., Racer for the (DLP) OWL component, all consequences of a rule engine e.g., Jess for the (DLP) SWRL component separately, and iterating until saturation or inconsistency, may still be not satisfying. One reason is that the ontology and rule components have incompatible semantics [9]: the OWL-DLP ontology being a fragment of OWL (the intersection of OWL with Horn clauses [8]) has open world first order semantics, while the DLP rule components has Datalog semantics (based on closed world assumption and Herbrand models). The soundness and completeness of the basic reasoning tasks (satisfiability and subsumption) should be carefully investigated. Although this simple approach is quite attractive, such loose integration have some limitations. The risks run in processing so have to be clearly identified and notified.

Implications :

- A loose interoperation of SWRL rules with existing rules engine is not satisfying, interoperating between ontology and rules requires a close integration
- A Web rule language should have a clear semantics that enables OWL DL and rules to safely interoperate.

3.2. [Use case] CG-3: Safe integration of OWL DL with rules with clear semantics

- **Outline: Brain Anatomy digitally processed images**

This use case on Brain Anatomy digitally processed images[1] (MRI) reported at Washington 2005 W3C Workshop on Rule Languages for Interoperability [4] and at the OWL Experiences and Directions Workshop collocated with the International Conference on Rule Markup Languages for the Semantic Web [5] presents some requirements of a Web Rule language expected for the Semantic Web Health Care and Life Sciences[2].
The general framework is sharing anatomical knowledge (ontology and rules) and tools (services) needed in the context of neuroimaging, applied both to medical practice, i.e. decision support in neurology and neurosurgery, and research about neurological pathology

[1] Slides available at http://www.w3.org/2004/12/rules-ws/slides/christinegolbreich.pdf
[2] Slides available http://www.med.univ-rennes1.fr/~cgolb/Slides/OWLED-Rules-CG.pdf

such as epilepsy, dementia, etc. The application aims at developing new methods for assisting the labeling of the brain cortex structures in MRI images. The approach proposed relies on a brain ontology storing the a priori "canonical" knowledge about the most important brain cortex structures, combined with rules describing the dependencies between their properties. A simplified example is provided to illustrate the need for supplementing OWL with rules, for reasoning over the ontology complemented with rules. This example illustrates that solutions are missed if the Web ontology language and the rule languages are not closely integrated. Although it should be more carefully investigated, the language required for the rules might be a FOL extension with function-free Horn rules (with negation as failure).

Implications

- A language extending OWL DL expressiveness with rules is required for Health Care and Life Sciences applications
- Some reasoning support is required to reason over a knowledge base composed of an ontology and rules
- The rule language to be devised should have a clear semantics that enables OWL DL and rules to safely interoperate (decidable, sound and complete reasoning)

In conclusion of these two use-cases, a close integration between ontology and rules allowing safe reasoning is required.

4. Fuzzy rules

4.1. [Use case] CG-4: Fuzzy Brain Anatomy

The Brain Anatomy use case has been extended for fuzzy reasoning, see use-case "Fuzzy Reasoning with Brain Anatomical Structures", presented by Giorgos Stamou et al.

5. Other use cases with ontologies and rules

Several other uses cases can be provided in Health Care and Life Sciences, illustrating different uses of rules in that field.

5.1. [Use case] CG-5: a language combining ontology and rules for semantic integration of heterogeneous innformation

- **Outline: integration of dialysis and transplantation data for strategic decisions in Health care** [11 -13-14-15]

Semantic integration is now crucial in many biomedical domains where better patient care, as well as better understanding of diseases and sound decision making in public health requires accessing large amounts of data from heterogeneous resources. For example, strategic decisions in the field of organ failure public health policies for end-stage renal disease, dialysis or transplantation requires accessing patient data scattered in many Hospital Information Systems. A simple scenario can be provided, based on a first case study that was achieved two years ago with the National French Biomedicine Agency (Agence de Biomedecine). The goal of the use-case might be to answer queries from 3 local databases where the dialysis and transplant data are stored and a (partial) dialysis and transplantation ontology in OWL. The ontology (still under construction) is issued from the Biomedicine Agency terminological server which was originally built in integrating several existing terminologies, e.g. the French Thesaurus of Nephrology and the International Classification of Diseases (ICD). A first prototype [11] achieving this simple scenario was achieved using PICSEL mediator, according to the LAV approach based on CARIN language.

Implications

- OWL DL is required for the dialysis and transplantation ontology
- A rule language for "mapping rules" and for expressing the "queries"
- A combination of OWL DL and rules that allows representing the ontology in OWL DL, expressing mapping rules between the local and global ontologies, formulating queries, so as to answer the user queries is desirable:

5.2. [Use case] CG-6 Interoperating between ontology and rules

- **Outline: Diagnosis rules in Health Care**

This use-case is based on the famous Clancey's Heuristic Classification method [18]. According to it, Diagnosis involves three main steps: data abstraction, heuristic matching, and refinement.

- data abstraction transforms the data (e.g., finding such as temperature value of 39° C) into data abstractions (e.g., high fever), usually using abstraction rules
- heuristic matching associates the previous data to a generic hypothesis (e.g. disease), using heuristic matching rules
- refinement allows to specialize the hypothesis into more refined hypothesis, based on an ontology (e.g. ontology of diseases such as SNOMED-CT)

This use-case scenario may come with various scenarios, for example including fuzzy reasoning e.g., for breast cancer diagnosis based on hybrid reasoning combining fuzzy results issued from digital image analysis with reasoning based on the ACR classification (from the American College of Radiology) of mammography images, or with the TNM Breast Tumors classification.

Implications

- Interoperating between ontology (Tbox) and rules (Rbox)
- possibly fuzzy reasoning

5.3. [Use case] CG-7 Combining ontologies and rules

- **Outline: Guidelines assisting decision making in Health Care**

Another use-case and scenario might be provided in the domain of Recommendations for Chronic Diseases (guidelines), more specially for Type 2 Diabetes. Guidelines. Therapeutic recommendations in the guidelines can be considered as rules composed of body and head [19]. For chronic diseases, body are usually expressed as combinations of clinical and therapeutic criteria. Therapeutic criteria include patient's past or ongoing treatments, i.e. earlier treatment that has been prescribed and its outcome in terms of efficacy and tolerance. Body are sets of therapeutic options, generally expressed in terms of therapeutic classes, but sometimes expressed otherwise, as a particular type or group of therapeutic agents. Here some examples of rules:

If oral monotherapy with maximal doses of sulfamide or metformin associated with lifestyle changes is not effective, then the monotherapy should be replaced by oral biotherapy

If a drug may interact with patient's medication or other conditions e.g., contraindications do not prescribe this treatment.

Suggesting therapeutic recommendations require to combine the knowledge issued from several ontologies, e.g., an ontology of drugs, an ontology of diseases, an ontology of food with the rules.

Implications

- Combining ontology (Tbox) and rules (Rbox)

5.4. [Use case] CG-8: Reasoning with rules for building and validating ontologies

- Outline: meta-reasoning rules for ontologies

The goal is to use a set of rules for building or validating ontologies such as the FMA (1.1) or the Brain Anatomy (3.2). For example, the rule below generates 221 relations between the classes of the the Brain Ontology [1] [10] such as CentralSulcus separates FrontalLobe and ParietalLobe.

IF Y is part of X and Z is not part of Y and T separates X and Z THEN T separates Y and Z

```
AE(?x) ∧ AE(?y) ∧ AE(?z) ∧ AE(?t) ∧ part-of(?y,?x) ∧ not
part-of(?z,?y) ∧ separates(?t,?x,?z) → separates(?t,?y,?z)
```

Similary, the rule below expressing that if an entity has laterality, then its parts have the same laterality, enables to verify whether laterality constraints are respected in the FMA or Brain ontologies.

If X is part of Y and Y has side Z then X has the same side

```
isPartOf(?x, ?y) ∧  hasSide(?y, ?z) → hasSide(?x,?z)
```

Implications

- Interoperating between ontology and rules

6. References

1. Golbreich C, Dameron O, Gibaud B, Burgun A, Web ontology language requirements w.r.t expressiveness of taxononomy and axioms in medicine, 2nd International Semantic Web Conference, ISWC 2003, Sanibel Island, Florida, October 2003, LNCS 2870, Springer. 2003.
2. Golbreich C., Combining Rule and Ontology Reasoners for the Semantic Web, Invited talk, Rules and Rule Markup Languages for the Semantic Web, Hiroshima, Japan, G. Antoniou, H. Boley Editors, LNCS 3323, Springer, 2004.
3. Golbreich, C., Imai, A. Combining SWRL rules and OWL ontologies with Protégé OWL Plugin, Jess, and Racer. 7th International Protégé Conference, Bethesda, MD, 2004.
4. Golbreich, C. Bierlaire, O. Dameron, O. and B. Gibaud. Use case: Ontology with rules for identifying brain anatomical structures. W3C Workshop on Rule Languages for Interoperability, 2005.
5. Golbreich, C. Bierlaire, O. Dameron, O. and B. Gibaud. What reasoning support for Ontology and Rules? the brain anatomy case study, OWL Experiences and Directions Workshop, collocated with the International Conference on Rule Markup Languages for the Semantic Web, Galway, Ireland, 2005
6. Martin O'Connor, Holger Knublauch, Samson Tu, Mark Musen, Writing Rules for the Semantic Web Using SWRL and Jess, Protégé With Rules workshop collocated with 8th International Protégé Conference, Bethesda, MD, Madrid, 2005.
7. M. J. O'Connor, H. Knublauch, S. W. Tu, B. Grossof, M. Dean, W. E. Grosso, M. A. Musen. Supporting Rule System Interoperability on the Semantic Web with SWRL. Fourth International Semantic Web Conference (ISWC2005), Galway, Ireland, 2005

8. Grosof B., Horrocks I., Volz R., and Decker S., Description Logic Programs: Combining Logic Programs with Description Logic In: Proc. 12th Intl. Conf. on the World Wide Web (WWW-2003), Budapest, Hungary, May 20-23, 2003.
9. Ian Horrocks, Bijan Parsia, Peter Patel-Schneider, and James Hendler. Semantic web architecture: Stack or two towers? In Francois Fages and Sylvain Soliman, editors, Principles and Practice of Semantic Web Reasoning (PPSWR 2005), number 3703 in LNCS, pages 37-41. SV, 2005.
10. Dameron O., Gibaud B., Musen M. Using semantic dependencies for consistency management of an ontology of brain-cortex anatomy, KR-MED 2004.
11. Golbreich C., Mercier S., Construction of the dialysis and transplantation ontology: advantages, limits, and questions about Protege OWL, Workshop on Medical Applications of Protégé, 7th International Protégé Conference, Bethesda, 2004 (Master report: data integration for Health Care decision support in dialysis and transplant, in French).
12. Golbreich C., Burgun, Biomedical Information Integration, a hot issue, Poster at MEDINFO 2004, San Francisco, 2004.
13. Golbreich C, Burgun A, Challenges for Biomedical Information Integration, Position Statement, in Proceedings of the Semantic Integration Workshop, Sanibel Island, Florida, Edited by AnHai Doan, University of Illinois at Urbana-Champaign, Alon Halevy, University of Washington, Natasha Noy, Stanford University, CEUR Vol 82, 2003.
14. Maquet G., Golbreich C., Burgun A., From an ontology-based search engine towards a more flexible integration for medical and biological information, In Proceedings of the Semantic Integration Workshop, Sanibel Island, Florida, Edited by AnHai Doan, University of Illinois at Urbana-Champaign, Alon Halevy, University of Washington, Natasha Noy, Stanford University, CEUR Vol 82, 2003
15. Golbreich C., Burgun, Biomedical Information Integration, a hot issue, Poster at MEDINFO 2004, San Francisco, 2004
16. Rubin DL, Dameron O, Bashir Y, Grossman D, Dev P and Musen MA, Using ontologies linked with geometric models to reason about penetrating injuries. American Med Info J, 2005 (AMIA)
17. C. Golbreich, S. Zhang, O. Bodenreider Foundational Model of Anatomy in OWL: experience and perspectives, OWL Experiences and Directions Workshop, collocated with the International Conference on Rule Markup Languages for the Semantic Web, Galway, Ireland, 2005 (extended version submitted)
18. W. Clancey. Heuristic classification. Artificial Intelligence, 27:289--350, 1985.
19. Vahid Ebrahiminiaa, Catherine Duclosa, Regis Cohenb, Alain Venota, Representing the Patient's Therapeutic History in Medical Records and in Guideline Recommendations for Chronic Diseases Using a Unique Model, MIE 2005