Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
ScienceDirect
Web Semantics: Science, Services and Agents on the World Wide Web 6 (2008) 133–136
JOURNAL OF Web Semantics
www.elsevier.com/locate/websem

Translating the Foundational Model of Anatomy into OWL

Natalya F. Noy *, Daniel L. Rubin

Stanford Center for Biomedical Informatics Research, Stanford University, Stanford, CA 94305, USA

Received 26 February 2007; received in revised form 28 August 2007; accepted 2 December 2007
Available online 8 December 2007

Abstract

The Foundational Model of Anatomy (FMA) represents the result of manual and disciplined modeling of the structural organization of the human body. It is a tremendous resource in bioinformatics that facilitates sharing of information among applications that use anatomy knowledge. The FMA was developed in Protégé and the Protégé frames language is the canonical representation language for the FMA. We present a translation of the original Protégé frame representation of the FMA into OWL. Our effort is complementary to the earlier efforts to represent FMA in OWL and is focused on two main goals: (1) representing *only* the information that is explicitly present in the frames representation of the FMA or that can be directly inferred from the semantics of Protégé frames; (2) representing *all* the information that is present in the frames representation of the FMA, thus producing an OWL representation for the complete FMA. Our complete representation of the FMA in OWL consists of two components: an OWL DL component that contains the FMA constructs that are compatible with OWL DL; and an OWL Full component that imports the OWL DL component and adds the FMA constructs that OWL DL does not allow.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Ontology; OWL; Semantic Web; Foundational Model of Anatomy

1. Motivation and goals

The Foundational Model of Anatomy (FMA) [9] represents the result of manual and disciplined modeling of the structural organization of the human body. Many workers in biomedical informatics consider the FMA to be a tremendous resource that facilitates sharing of information among applications that use anatomy knowledge. The FMA developers used the Protégé ontology editor [2] to create the FMA and the Protégé-frames representation is the canonical representation language for the FMA. With the development and use of biomedical ontologies in the context of the Semantic Web, researchers often want to access the FMA using the tools designed for the Web Ontology Language (OWL) [4]. The FMA Protégé representation relies heavily on some of the features that are common in frame-based representations but are not so common in OWL (and, in particular, in OWL DL) ontologies. These features include metaclasses, classes (rather than instances) as property values, and attributed relationships between classes (see [7] for details of the Protégé frame representation of the FMA).

Previous efforts to represent portions of FMA in OWL [5,3] focused on transforming the FMA into the form that Description Logics reasoners could use efficiently, producing useful inference. Because the original FMA representation was not designed to be used with DL classifiers, such transformation necessarily involves making assumptions about the FMA representation that are not present in the frames version.

Previous efforts also focused on a subset of the FMA. However, anatomy is a broad field and different user communities are interested in different portions of the FMA. Therefore, a complete representation of FMA in OWL is needed.

Our effort is complementary to the previous efforts because the goals of our transformation were different. Specifically, our goals were two-fold:

(1) **Integrity**: Represent *only* the information that is explicitly present in the frames representation of the FMA or that can be directly inferred from the semantics of Protégé frames.

* Corresponding author.
E-mail addresses: noy@stanford.edu (N.F. Noy), rubin@stanford.edu (D.L. Rubin).

1570-8268/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.websem.2007.12.001

Author's personal copy

(2) **Completeness**: Represent *all* the information that is present in the frames representation of the FMA, thus producing an OWL representation for the complete FMA.

Thus, our goal was to provide a complete OWL rendering of the canonical Protégé-frames representation of the FMA. Our goal was *not* to find the most appropriate representation of the FMA in OWL (or OWL DL). Rather, we wanted to stay as close to the structure of the original representation as possible when representing it in OWL and to make as few assumptions as possible about the semantics of the representation. Thus our OWL representation does not contain the types of structures that DL classifiers use to infer new subsumption relationships or inconsistencies in the ontology because the frames representation did not have these constructs.

To the best of our knowledge, our work is the only one at the time of this writing that produced an OWL representation of the complete FMA rather than its fragments.

We believe that the resource that we provide has significant utility for the biomedical community. For example:

- The resource is a complete representation of a FMA in a standard ontology language that is endorsed by W3C.
- Researchers can extend the OWL version of the FMA that we publish with additional statements to support different types of reasoning applications. For example, researchers extended an OWL representation of the FMA to create an application to reason about the anatomic consequences of penetrating injury [10].
- Most OWL tools can now load the full FMA (provided the tools can generally handle ontologies of the size of the FMA).
- Other OWL researchers can write conversion scripts against this version to transform it into another representational form that is more suitable for their work. (In fact, since making the resource available, we had several reports of just such use, which was not possible before.)

2. Components of FMA in OWL

Because the FMA uses metaclasses and classes as property values, among other things, we could not use OWL DL for our representation [6], and OWL Full was a natural language of choice for us. However, many researchers need an OWL DL version for compatibility purposes. For example, if they want to extend the FMA and have an OWL DL ontology as the result, they need an OWL DL version of the FMA to start with. Hence, we chose a two-layered approach to the OWL representation of the FMA (Fig. 1): first, we created an OWL DL component that included the part of the FMA that could be directly represented in OWL DL. This part contains all FMA classes and the class hierarchy; all properties; their domains and ranges; specification of functional properties; property restrictions (such as cardinality constraints and local ranges, which became cardinality restrictions and `allValuesFrom`restrictions in OWL). We designated six properties as annotation properties (Fig. 1) and these properties and their values are also in the OWL DL model. Second, we created an OWL Full component that *imports* the OWL DL model using the `owl:imports` construct. The OWL Full constructs refer to classes and properties in the OWL DL model, making additional assertions about them. Note that partonomy and adjacency information in FMA is represented at the level of classes and therefore is part of the OWL Full component in our translation. For details of our translation, please see the corresponding technical report [8].

Researchers are currently developing OWL 1.1 [1]—a new version of the OWL language, which, among other features, will allow some level of meta-modeling within the DL framework. The OWL 1.1 authors suggest the technique of *punning* to enable statements about classes: with punning, the same name can be used for a class and an individual, as long as the use of the name as an individual has no effect on the meaning of the name as a class. It is possible that with the use of punning, all of the information in the FMA can be translated into OWL 1.1 (such

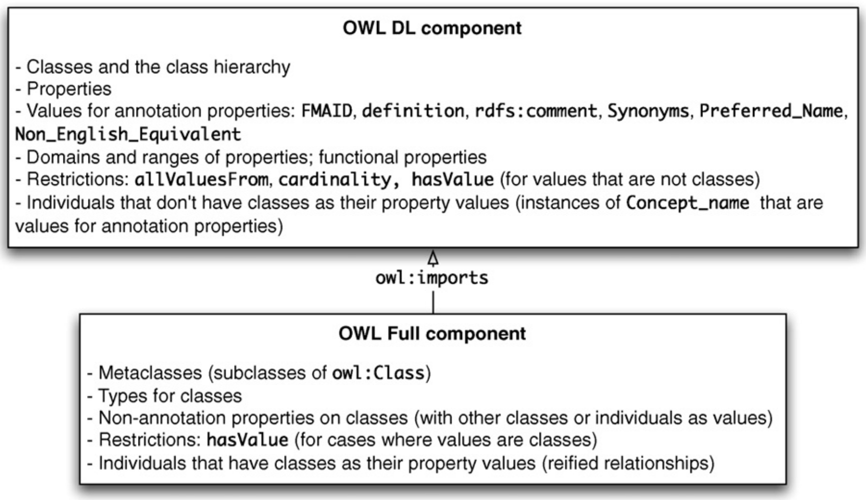

Fig. 1. The two-layered approach to the OWL languages: the OWL DL component contains part of the FMA that directly corresponds to the DL species of OWL. The OWL Full component imports the DL one and makes additional assertions about its classes and properties.

Author's personal copy

transformation is beyond the scope of this work since OWL 1.1 is not currently a standard).

3. Related work and discussion

There have been several earlier efforts to represent the FMA in OWL [5,3]. These efforts pursued a different goal, mainly to transform the FMA into the form that Description Logics (DL) reasoners could use efficiently, producing new subsumption relationships, or finding inconsistencies. Therefore they necessarily had to make assumptions about the FMA components that were not explicitly present in the canonical Protégé frames representation. For instance, in the transformation by Golbreich et al. [5], relationships between classes (where a "literal" transformation would require using classes as property values, which is not allowed in OWL DL) became existential restrictions (`someValuesFrom`) in OWL: such approach is more natural in DLs and allows more inference to be performed by a classifier. Similarly, Dameron et al. [3] introduced disjoint and closure axioms that were not explicitly present in the original representation.

For the biomedical researchers who plan to use DL reasoning with specific portions of the FMA, these transformations may be more appropriate. For the researchers who require the FMA in its entirety and as close to its original Protégé representation as possible and who do not plan to use DL classifiers with the FMA itself (but may want to use it with ontologies they derive from it), the transformation resulting from the effort described here may be more appropriate. In addition, our effort may benefit researchers who need only the information that is contained in our DL version and wish to use it as a foundation upon which they may add additional restrictions appropriate to their specific needs and application domain.

Note, however, that at the time of this writing we are not aware of any successful attempts to use a DL classifier to perform subsumption reasoning on all of the FMA. First, the sheer size of the FMA poses a considerable challenge: classifiers simply run out of memory when given such a large OWL ontology to process [12]. In fact, none of the classifiers that we tried on a machine with 4G of memory was able to classify or validate the OWL DL file before we removed values for all annotation properties (thus reducing the size of the OWL DL file by 80%). This limitation makes the current active area of research on ontology modularization [11] all the more pertinent. While we may not yet be able to classify ontologies of the size of the FMA as a whole, we may be able to classify subsets of it. Note that we can load both the OWL DL and OWL Full version of the FMA into Protégé 3.3 for browsing and editing. Second, the FMA does not have any necessary and sufficient conditions defined for its classes. Zhang et al. report on using one of the patronomic relations (`constitutional_part_of`) to define a class [12]. However, their experiments demonstrated that using this property as a definitional one was inappropriate. It is not clear if there is any property or a set of properties that could be used automatically for necessary and sufficient conditions because the FMA simply was not designed with these types of definitions in mind.

While we tried to introduce as few assumptions as possible into the semantics of frames, there are some fundamental differences between the semantics of frames and OWL that we had to keep in mind. One of the key differences is the use of the closed-world assumption in frames and the open-world assumptions in OWL. Thus, it is impossible to achieve the identical semantics of the two representations without making additional assumptions that are not in the original representation. Consider, for example, the notion of allowed classes for slots in Protégé frames, which we translated into the `allValuesFrom`restrictions in OWL. In the former case, we are saying that if there is an instance where a slot value is not a member of the allowed class then that instance violates the constraints on the model. In the second case, if there is an instance where slot value is not known to be a member of the class in `allValuesFrom`, that value is assumed to be in that class. "Closing" the OWL model in this case would require making assumptions about disjointness of classes in the FMA. Thus, the semantics are different, but such difference is inevitable due to the fundamental differences between the languages.

In many cases though we were able to get the correct semantics even though the "natural" semantics of frames and OWL differ. For instance, several domain statements for a slot in Protégé mean that the slot is valid for instances of *any* of the classes among the domains. In OWL, if there are multiple domain statements for a property, then any individual with this property must be an instance of *each* of the classes in the declared domains. Thus, when a slot in Protégé had multiple domains, we set the domain of the OWL property to be the union of these. The same is true for ranges of properties (allowed classes for slots in Protégé).

Even though we tried to add as few assumptions about the semantics of the FMA representation as possible, the conversion script that we wrote is not a generic conversion script from Protégé frames to OWL. First, we have omitted from consideration any Protégé-frames features that are not used in the FMA representation (e.g., numeric slots and their ranges, subslots, standalone individuals that are not values for slots on classes, etc.). Some of these features do not have a straightforward representation in OWL (such as numeric minimum and maximum), but we did not have to worry about it because our goal was strictly to convert the FMA.

Second, some of our decisions on the OWL counterparts to the frames constructs are specific to the FMA representation and cannot be generalized. The main example is the class-valued slots. In Protégé frames, one can specify that values of a particular slot are limited to classes that are subclasses of the *allowed parents* for the slot. There is no similar construct in OWL. In general, it is incorrect to assume that a class-valued slot with allowed parent `A` is equivalent to the same slot with range (allowed class) `A`. However, in the specific structure of the FMA, it is indeed correct to make such assumption: in the FMA, classes are both subclasses and instances of their superclasses Similarly, we did not represent default values for slots (i.e., as the information on the slots) because they were already "materialized" in the FMA representation: Protégé assigns default values for slots at the time it creates an instance. Therefore, all the

Author's personal copy

slots in the FMA representation that we were translating already had the values assigned. Thus, simply by translating the values of these slots we got the "materialized" defaults in the OWL version.

4. Availability, versions, statistics

We used FMA version 2.0, accessing it from a mySQL database. We used the Protege 3.3 release. The FMA in OWL contains 78,988 classes; 3309 restrictions; 110 object properties; 54 datatype properties; 6 annotation properties.

The FMA in OWL is available at the National Center for Biomedical Ontology: http://bioontology.org/projects/ontologies/fma/. The conversion software is available at http://www.bioontology.org/wiki/index.php/FMAInOwl.

Acknowledgements

This work was supported by the National Center for Biomedical Ontology, under roadmap-initiative grant U54 HG004028 from the National Institutes of Health. We are grateful to Olivier Bodenreider, Todd Detwiler, Christine Golbreich, Onard Mejino, and Songmao Zhang for fruitful discussions. We thank anonymous reviewers of an earlier draft for their insightful comments and suggestions.

References

[1] OWL 1.1 Web Ontology Language, http://www.webont.org/owl/1.1/, 2007.
[2] The Protégé project, http://protege.stanford.edu, 2007.
[3] O. Dameron, D.L. Rubin, M.A. Musen, Challenges in converting frame-based ontology into OWL: the Foundational Model of Anatomy case-study, in: AMIA Annual Symposium, Washington DC, 181-185, 2005.
[4] M. Dean, et al., Web Ontology Language (OWL) reference version 1.0, http://www.w3.org/tr/owl-guide/, 2002.
[5] C. Golbreich, S. Zhang, O. Bodenreider, The Foundational Model of Anatomy in OWL: experience and perspectives, J. Web Semantics 4 (3) (2006) 181–195.
[6] N.F. Noy, Representing classes as property values on the Semantic Web, Technical Report, W3C Working Group Note, 2005.
[7] N.F. Noy, M.A. Musen, J.L.V. Mejino, C. Rosse Jr., Pushing the envelope: challenges in a frame-based representation of human anatomy, Data Knowl. Eng. 48 (3) (2004) 335–359.
[8] N.F. Noy, D.L. Rubin, Translating the Foundational Model of Anatomy into OWL, Technical Report SMI-2007-1234, SMI, 2007.
[9] C. Rosse, J.L.V. Mejino, A reference ontology for bioinformatics: the Foundational Model of Anatomy, J. Biomed. Inform. (2004).
[10] D.L. Rubin, O. Dameron, M.A. Musen, Use of description logic classification to reason about consequences of penetrating injuries, in: AMIA Annual Symposium, 2005.
[11] J. Seidenberg, A. Rector, Web ontology segmentation: analysis, classification and use, in: 15th International World Wide Web Conference, Edinburgh, Scotland, 2006.
[12] S. Zhang, O. Bodenreider, C. Golbreich, Experience in reasoning with the Foundational Model of Anatomy in OWL DL, in: Pacific Symposium on Biocomputing (PSB 2006), Maui, Hawaii, 2006.