Collaborative Construction of Visual Domain Ontologies Using Metadata Based on Foundational Ontologies

Gabriel M. Torres, Alexandre Lorenzatti, Vitor Rey, Rafael P. da Rocha, Mara Abel

Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

(gmtorres, alorenzatti, vfrey,marabel)@inf.ufrgs.br, rafael.rocha@ufrgs.br

Abstract. *Domain ontologies are widely used to explicit declarative knowledge. However, it is a difficult task to obtain an explicit and shared vocabulary that can be used in computer systems. Besides that, many domains require not only textual data but also visual data to express the meaning of the concepts. Some ontology editors have been developed to support collaboration on the ontology development process, but none of them have support to visual domains or ontological foundation, which is important to build efficient ontologies with semantic validation. In this paper, we introduce an approach to support collaborative construction and evolution of visual domain ontologies using metadata ontologies based on a foundational ontology.*

1. Introduction

In the Knowledge Engineering process, we are usually concerned on collecting knowledge shared by a specific community, storing it as a formal ontology and using this ontology as a reusable artifact for various purposes. As we know so far, an ontology is defined as a formal specification of a *shared* conceptualization (Borst, 1997). Knowledge domains are not static: they evolve when new elements become part of the domain or when elements become obsolete (De León, 2009). These changes need to be adapted to the domain model, updating the ontology by adding or removing components. Therefore, collaboration has become an important part of the ontology development process, helping in making explicit the concept behind vocabulary and evolving the vocabulary to its new meanings.

Several classic ontology editors have been developed over time. Some interesting recent approaches focus on collaboration aspects of ontology development, like NeON Project (Haase et al., 2008), OntoEdit (Sure et al., 2002) and Collaborative Protégé (Tudorache, Noy, & Musen, 2008). Collaborative Protégé has introduced a metadata ontology to support this collaboration, allowing the specification of changes and annotations on the domain ontology. However, we think that the formal language oriented-interface is hard for the domain specialist to deal with, because he/she is not usually familiar with ontology formalization.

Many information domains like Medicine and Geology need visual information to express knowledge. Our tool allows the user to build a conceptual model of visual knowledge, getting advantage of visual representations like icons and images to help in expressing the full meaning of the concepts. On the other hand, in order to capture the correct meaning of a concept, it is necessary to let the user express his/her understanding about that concept through the use of properties that have concrete meaning to him/her. This is the role of foundational ontologies: express the inherent properties that provide identity to the objects in some world. Our proposal consists of developing a collaborative web environment for ontology construction with two main contributions: support to visual domains and support to ontological foundation.

The collaboration is based on metadata information about the ontology components, so that the users can express their understanding about the meaning of the concepts without requiring any formal representation language, but giving mechanisms to manipulate visual information and to express rich semantic models. Metadata information provides the necessary vocabulary and artifacts that can offer the basis for the development of applications for collaborative ontology construction. The metadata and the domain ontology data are generically stored in a database as triples.

This paper is organized as follows: In Section 2, we explain why and how visual domains are important on ontology development. In Section 3, we present the foundational ontology that provides ontological foundation to our metadata models. Section 4 introduces the metadata ontologies that are the main contribution of this work. In Section 6, we conclude and anticipate some future work.

2. Visual Domains

Some information domains require visual knowledge as a crucial part of the problem solving process like Medicine or Geology and most of Natural Sciences. The interpretation process occurs through a visual pattern matching against the domain, capturing the objects that can support the inference path. Some of these visual objects have even barely translation to a propositional description. Therefore, the ontology construction in visual domains requires more than symbolic descriptions to explain the concepts. Besides that, many ontology developers find it easier to provide descriptions of their domain concepts and properties using visual representations rather than only formal descriptions (or pure natural language descriptions).

According to Lorenzatti (2011), a concept can be represented in two different ways: through a symbol from a language or through a pictorial representation (an imagistic representation, like an image or icon). Icons are similar to what they represent, so, their meaning is captured through the same process of perception used to recognize the represented object or event. In other words, its meaning can be understood from the observation of the representation. Images are photographs of concepts, intended to provide examples of instances of a concept, trying to transmit its meaning. Therefore, the concepts can be associated either to a symbol or to a pictorial representation, like an image, a draw, an icon, a chart, etc. An icon can also be associated to each property value. For example, the property *Roundness* can have an associated icon to each of its values: low-rounded, rounded, high-rounded.

3. Foundational Ontology

The aim of a foundational ontology is establishing a basis to obtain coherence in the negotiations of meaning during the collaboration process of individuals to build a conceptual model. Recently, a *Unified Foundational Ontology (UFO)* was proposed (Guizzardi 2005), defining categories that provide ontological foundation in the construction of conceptual models. The UFO is divided in three fragments called UFO-A, UFO-B and UFO-C. We are interested in the theoretical framework of meta-properties and meta-types proposed by Nicola Guarino (Guarino 1995) and Giancarlo Guizzardi, mainly focused in the UFO-A, which is the core of the foundational ontology, consisting of a stable theory that introduces structuring concepts to offer more semantics to conceptual modeling languages. Therefore, we will mention here the notions of *rigid sortals*, *properties*, *quality domains*, *partonomic relations* and *hierarchical relations*. For instance, a rigid sortal is a concept whose definition requires that their instances cannot stop being an instance of this concept in any possible world. This means that if the *essential properties* chosen to define the concept cease to be recognized in the way they were defined, the instance will cease to exist to because it loses its identity criterion. A person is a rigid sortal while a student is not, since there are instances of it that can stop being a student without losing its identity. These are important constructs for ontological models, since they allow producing trustful mappings among different domain ontologies that support interoperability.

The current ontology development tools don't implement this rich formal semantic representation because they are based only in the five basic ontology constructs (concept, property, property value, relation, axiom) that don´t express the differences of objects in reality according to human discrimination. The unified foundational ontology extended these basic primitives, creating several additional constructs that helps in establishing the taxonomic classification and the relationships among concepts. Therefore, ambiguity is reduced and the expressivity of the model is increased.

4. Metadata Ontologies for Collaboration

In this paper, we introduce an upper-level domain independent metadata to specify the structure of the domain ontology components and collaboration events. Using this metadata, the community of users can define concepts, attributes and domain values, making also explicit the intended meaning through the use of primitives of a foundational ontology (assigning values to meta-properties and meta-types of concepts), visual icons and illustrative images. The conflicts about selection of names, attributes, icons and images are solved by the proposition of changes in the models. The changes are justified by the ontological definition and are stored for further reference.

We introduce two metadata ontologies: the Representation Ontology (R.O.), which defines primitives for representing the domain ontology, and the Collaboration Ontology (C.O.), which defines primitives for representing the collaboration events. The domain ontology components are defined as instances of the R.O. concepts and the changes made over the domain ontology are defined as instances of the C.O. concepts. These models are the basis of our environment, structuring the meta-level data and helping the application to deal with the abstract representations of the domain ontology and to track changes involving symbolic or visual representations and foundational artifacts.

4.1. Representation Ontology (R.O.)

When dealing with ontologies, we are commonly focused on its main components: concept, property, property value, relation and axiom. To add more semantics, this meta ontology extends some of the main ontology components by specializing them based on visual and foundational aspects.

In order to provide visual support, we used the concept [Image], which is specialized in two sub-concepts: [Photography] (for representing photos of concept instances) and [Icon] (for representing symbolic pictorial icons). The R.O. contains some relations that link a [OntologyConcept] or [OntologyPropertyValue] to one [Icon] (*hasIcon*) and a [Photography] to a [OntologyConcept] (*photographyOf*). To provide ontological foundation to the model, we have specialized the R.O. concepts using some of the foundation constructs proposed in the UFO-A foundational ontology, enriching the semantics of the model without adding significant complexity. The [OntologyConcept] was specialized to represent Substantial Universals and its subclasses: Sortal, Kind, Mixin, RigidSortal, etc. The [OntologyProperty] concept was specialized to represent the distinct types of property: DataTypeProperty (for properties that point to primitives like *string, int, datetime*, etc.) and QualityUniversalProperty (for properties that have one or more pre-defined values, like color, age interval, etc.). The [OntologyPropertyValue] concept was specialized in Quale. The [OntologyRelation] concept was specialized to represent the different types of relations, allowing the representation of partonomic relations (ExtensionOf, MemberOf, PartOf, SetOf, SubQuantityOf and SubsetOf) and the hierarchical subsumption relation (SubclassOf). These constructs, when used correctly, impose semantic restrictions to the model, which can be analyzed by the knowledge engineer to help the users to detect semantic failures and representational misuses on the domain ontology.

4.2. Collaboration Ontology (C.O.)

The Collaboration Ontology (C.O.) defines which collaboration activities can be done on the domain ontology. In a simplified way, the C.O. instances are the changes related to what has been represented by the R.O. ontology, which is the domain ontology. The collaboration process is focused on the proposal and storage of changes made over concepts, properties or relations and also on annotations that can be attached to any domain ontology component. The specialists can make directly changes or annotations in the domain model by adding, changing or removing ontology components. Collaboratively, they can see each other change history and discuss about it, possibly making new modifications until a consolidated domain model is obtained.

The C.O. concepts define the set of possible ontology changes that can be done in the system. A change event has some properties that give meaning to it: *domainComponent* relates the change to one domain ontology component; *author* stores who made the change; *date* stores the date and time when the changed occurred, to help in tracking the evolution of ontology; *value* stores the new value of the change. The C.O. describes not only common changes (ConceptCreated, PropertyRemoved) but also visual changes (ConceptIconChange, ConceptPhotographyChange) and foundational changes (MemberOfRelationCreated, SubsetOfRelationCreated, QualeCreated, QualeIconChanged, etc.). Therefore, we can change the semantics of the concepts, adding more information to the domain ontology model than the current approaches offer. For example, if the user once created a concept by instantiating the R.O. concept [OntologyConcept], and now he/she wants to change its stereotype because it is, in fact, a RigidSortal, it can be done by creating an instance of the C.O. [ConceptTypeChange] concept and setting its *value* property to “RigidSortal”. The concept type can be changed unlimited times by the community, to reach the correct consensual semantics.

The C.O. also allows the collaboration on visual domains. Icons are an unique alternative representation of the concept, based in the visual perception, that helps in avoiding the excessive use of propositional interfaces in ontology-based systems. When a user changes a concept icon, an instance of [ConceptIconChange] is created and associated both to the concept and to the icon image uploaded. Using the same procedure, a concept can be associated to one or more photographs by creating instances of the C.O. concept [ConceptPhotographyCreated]. If a user deletes a photograph, an instance of [ConceptPhotographyRemoved] is created. A property value can also have an associated icon by creating an instance of [QualeIconChanged]. Further, the users can make comments specifically about the icons or photos and change these artifacts later on until they reach the correct consensual visual representation. In Lorenzatti (2011), a library of icons related to the Sedimentary Geology domain was developed with cognitive analysis support. We are currently using this library to validate our environment. An example of the metadata ontology interaction and the collaboration history generated when changing visual and foundational aspects of the domain ontology is shown in Figure 1.

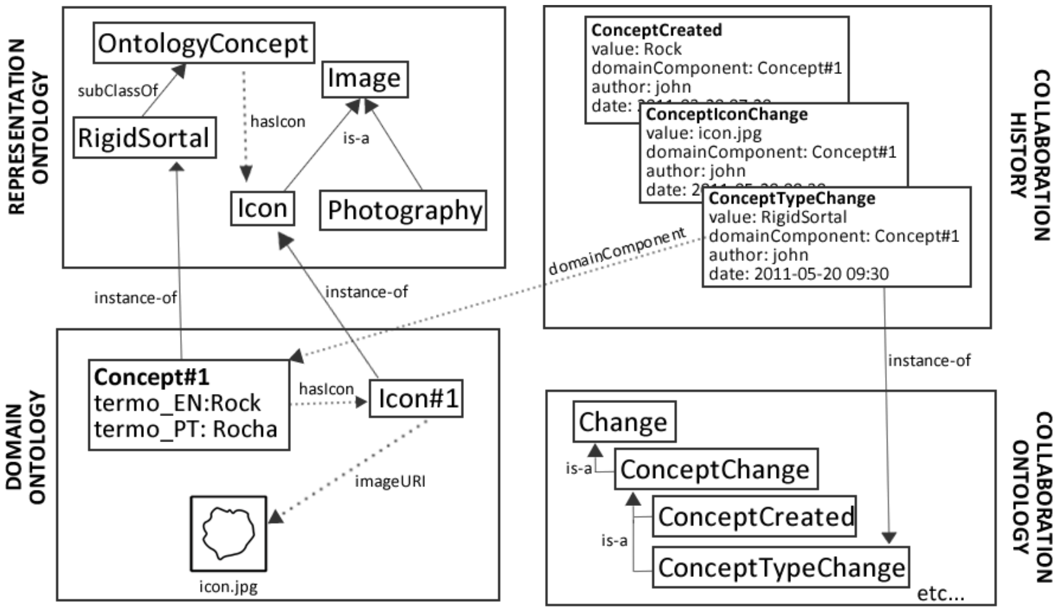

Figure 1 Collaboration structure with visual and foundation constructs

6. Conclusions and Future Work

Domain ontologies are explicit conceptual models of a shared knowledge, focused on a specific domain of common interest. Some domains need more than textual representations for concepts of reality, requiring visual representations as well. The foundational concepts introduced in the UFO-A foundational ontology helps the construction of semantic models by helping to establish the essential properties that express the identity of the concepts. Besides that, the foundational data helps the knowledge engineer to understand the domain and to interact with the specialists. We consider that supporting the user to recognize these properties would help in doing ontological choices in associating the concepts with the representation constructs. This would lead to the development of better quality domain ontologies, in terms of lucidity and laconicity, avoiding ambiguity and redundancy. In order to provide the adequate support, we develop a framework based on metadata ontologies that allow the domain experts in developing and evolving domain ontologies using textual and visual representations and ontological foundational data.

The metadata ontologies introduced in this paper and its usage are focused on providing a basis for the collaborative construction of rich language-independent domain ontologies. The generated collaboration events form an important collaboration and evolution history that can be used to analyze critical points of the ontology or to track changes. Our approach is experimentally available as a web-based environment for the collaborative construction of the Sedimentary Geology domain ontology at the address *http://obaita.inf.ufrgs.br/*, which is currently under constant development.

7. Acknowledgements

This work had the scholarships supported by CAPES and the financial support of CTPETRO Government program and ENDEEPER Co.

8. References

BORST, W. N. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse.** Twente: University of Twente, 1997. Doctoral thesis.

DE LEÓN, P.. **Ontology metadata management in distributed environment**. Madrid:Universidad Politécnica de Madrid, 2009. Doctoral thesis.

GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. **Int. Journal of Human Computer Studies**, v.43, n.5, p. 625–640,1995.

GUIZZARDI, G. Ontological Foundations for Structural Conceptual Models. **Enschede**, The Netherlands: Universal Press, v.5, n.74. 2005. 410 p. (CTIT PhD Thesis Series).

HAASE, P. et al. The Neon Ontology Engineering Toolkit. In International World Wide Web Conference, 17, Beijing, China, April, 2008, **Proceedings**…, 2008 (Developers Track.)

LORENZATTI, A.;ABEL, M.; FIORINI, S.;BERNARDES, A.;DOS SANTOS SCHERER, C..Ontological Primitives for Visual Knowledge. In Brazilian Conference on Advances in artificial intelligence,São Bernardo do Campo, 2010. **Proceedings**… Berlin / Heidelberg:Springer, , 2011. (Lectures Notes in Artificial Intelligence v.6404,)

SURE, Y. et ali.OntoEdit: Collaborative ontology development for the semantic web. In The International Semantic Web Conference,1, Sardinia, Italy, June 2002. **Proceedings**…. 2002. p.221-235.

TUDORACHE, T.; NOY, N. F.; MUSEN, M. A.Collaborative Protege: Enabling community-based authoring of ontologies. In International Semantic Web Conference, 7, Karlsrhue, Germany, 2008. **Proceedings**… 2008. (Posters & Demos).