The Evolution of Social Ontologies

Nicolas Weber

Graz University of Technology, Knowledge Management Institute,
Inffeldgasse 21a, 8010 Graz, Austria
nweber@Know-Center.at

Abstract. The idea of representing knowledge in formal models gets more and more popular. Since these models are understandable and thus processable by machines, the retrieval of knowledge is facilitated. Furthermore, ontology based systems use these models to provide meaningful and semantically enriched data. The success of Web 2.0 approaches has shown, that communities contribute collaboratively their knowledge to the web. The approach of this work is to understand and support collaborative modeling of structured knowledge by communities – the Evolution of Social Ontologies.

Keywords: collaborative modeling, ontology evolution

1 Introduction

By definition an ontology as an explicit specification of a conceptualization [1] represents the common understanding[2][3] of a group of people about a certain domain. An ontology describes the result of a process of negotiation and reflects the reality in form of a formal model. Unfortunately this is a theoretical definition and practical reality is different. Ideally the members of a certain community agree on a common perception of a domain. But finally, a person who is enabled to represent the domain in a formal description, the so called ontology engineer, models an ontology assuming that his perception corresponds to the perception of the community. This means that the ontology modeler combines his knowledge about a certain domain with his expertise in formal description languages.
Unlike the ontology modeler, communities do not have the technical ability to formalize their common understanding of a certain domain. Based on the complexity of formal knowledge representation, there is a lack of tools supporting adequately the process of collaborative domain modeling. The domain to be modeled covers the knowledge represented in collective and distributed repositories like wikis, file- and mail-systems. These documents describe the common understanding of a domain in an informal manner. The fact that a collaborative and communicative process is required for establishing and maintaining these ontologies makes them ‘social’.
The objective of this work is to elaborate specific characteristics of social ontologies on the one hand and to support this process of informal collaborative modeling on the other hand. This will be achieved by investigating extracted knowledge structures

(ontologies) on different levels of formalization, e.g. Wikis (informal content and formal markup), tagged entities (Folksonomies) and formal representations of social networks.

Approach

By now, in today's internet there is a lot of structured knowledge available. The results of extracting these structures are formal ontologies on different levels of formalization. The main characteristic of these ontologies is that they are created collaboratively. For instance, social networks represented in networking portals are growing in several dimensions; new members are joining the network, the interconnection increases through establishment of new connections, furthermore the nodes in the network gain additional markup through completion of profile data. Also Wikis grow through creation of new articles and adding further relations (links).
Ontologies that are extracted from highly used systems are characterized by a permanent transformation. The same happens with traditional 'hand made' ontologies – they keep updated through several steps and iterations. This improvement of an ontology over a certain amount of time is called 'the ontology lifecycle'.

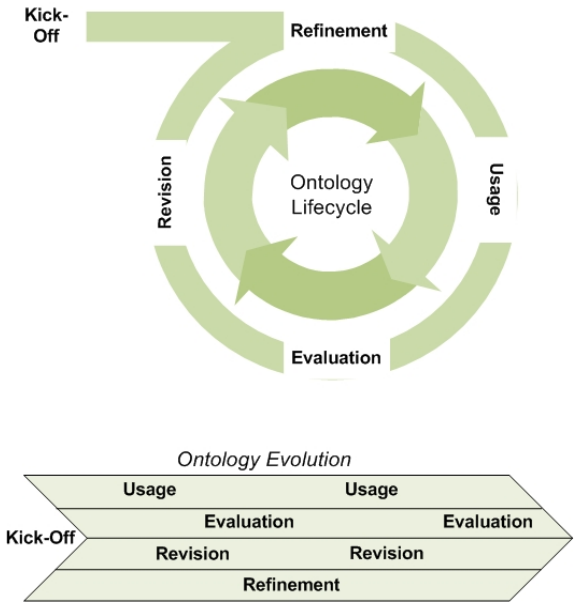

Fig. 1. The Ontology Lifecycle (above) describes the process of iteratively passing through the steps of the lifecycle. Unlike the concept of ontology evolution (below), the steps in the Ontology Lifecycle are arranged in a sequence. Since working with social ontologies is a collaborative process, usage, evaluation and revision of the ontology are simultaneous.

By passing through a sequence of steps several times – creation, refinement, usage, evaluation, adaptation – the ontology improves. In case of social ontologies the process of transformation is different. Since these ontologies are used, improved, corrected etc. at the same time, the model is exposed to permanent change. As shown in Fig.1 the evolution of social ontologies is a simultaneous process of usage, evolution and refinement. This permanent process is more like an evolution [9] than a lifecycle – the Evolution of Social Ontologies.
The objective of my future work is to analyze the evolution of collaboratively created ontologies. Generally the work is based on answering following research questions:

How do social ontologies differ from traditional ontologies?
Compared to traditional ontologies, the process of creation and maintenance of social ontologies is different. As seen in Fig.1, the ontology lifecycle concept is not valid for this type of ontologies. Furthermore, social ontologies differ in level of formalization. Starting from directed graphs (Wikis), over taxonomic models (file systems) and low level ontologies (social networks) up to complex models with a high level of formalization (semantic wikis).

Which characteristics are relevant for the evolution/maturing process of social ontologies?
This aspect focuses on the transformation between two or more snapshots of an ontology over the time. Which characteristics of a social ontologies are relevant for the evaluation? Are there factors to measure the quality of a collaboratively created model?
Collaboratively created models underlie the concept of evolutionary growth[8]. A main point of this concept is that after each seeding and evolutionary growth phase a reseeding phase follows. This phase aims at resolving inconsistencies in the ontologies and the consolidation of similar concepts and attributes (relations). Unlike the process of refinement and correction in Fig.1, where an individual interacts only with one part of the ontology, the task of resolving inconsistencies requires knowledge about the whole ontology. Part of my dissertation will be to find indicators for a needed transition between these phases.

How can the evolution / maturing process of social ontologies be supported?
Semantic Maturing [7] describes the maturing process of semantic models. In order to ensure a permanent improvement of the ontology the maturing process has to be supported. This includes observing the transformation of the ontology and identify indicators in order to provide support to the individual user. This could be for instance a recommendation of relations and semantic markup or tools for clustering and classification in order to conceptualize instances.

The research for this dissertation will be performed within the large-scale integrated project (IP) MATURE. The objective of MATURE is to better understand the process of knowledge maturing and to build tools and services to reduce maturing barriers. Within MATURE the maturing process will be analyzed in three dimensions – content, processes and semantic. In the area of semantic maturing the maturing of

social ontologies will play an important role. This work will benefit from the theoretical and practical work in the project.

Research Methodology

The following steps will be performed during this work:

- Identification of collaboratively created structures (i.e. Wikis, File Systems, Folksonomies)
- Methods for the extraction of meaningful ontologies
- Evaluation of transitions between snapshots
- Identification of meaningful measures describing the evaluation of ontologies
- Supporting the process of ontology evolution

Background

This section highlights shortly the scientific background of the approach. Significant problems as well as the state of existing solutions are discussed in the field of knowledge management, informal modeling and externalization of knowledge.

Knowledge Management

Knowledge management deals with the maintenance of knowledge in organizations and the effort of making knowledge accessible and reusable. In this context, knowledge management distinguishes between tacit and explicit knowledge. Since tacit knowledge is knowledge that people carry in their minds, it is difficult to access and thus to manage. Many approaches in the area of knowledge management try to externalize this knowledge in knowledge artifacts like text documents or wiki articles..

Humans learn from childhood to formulate their knowledge in spoken or written language. But besides the fact that it is often difficult for people to externalize the knowledge, in the meanwhile it turned out that natural language is not always the best way to represent knowledge. Indeed natural language is a non deterministic context-sensitive formal representation since it is restricted by many rules and axioms defined by the grammar of a particular language. However, due to the complexity of its grammar and the multiplicity of exceptions, it is not suitable to be processed by machines. The field of natural language processing tries to extract knowledge from text and models this information in machine readable form. Unlike this work, these approaches are focused on the extraction of formal coherences in informal representations.

A huge amount of tacit and explicit knowledge describes the attributes of entities and the relations between them. Ontologies enable people to formalize these structures and facilitate the representation of the human understanding of a domain in a semantic model. By this reason knowledge management in terms of establishing and maintaining semantic models emerges more and more as crucial task. Software systems like tools for modeling support or wikis, extended with semantic functionality, aim at stimulating the process of externalization of structured knowledge. Traditional semantic systems are built upon ontologies in order to make a common understanding of a community in terms of machine readable background knowledge reusable. This new discipline in knowledge management, the perception of structure, graphs, networks and models as manageable knowledge artifacts, will be supported by the approach of the Evolution of Social Ontologies.

Informal Modeling Approach

As described above, the process of externalizing tacit structural knowledge can be done by ontology modelers. In this case knowledge about the domain and expertise in the task of formalization is required to represent the knowledge in a declarative language. Former approaches have shown that the bottleneck is often ontology engineering. The formal representation of a certain domain does not always reflect the common understanding of the community. The level of granularity of formalized facts does not always match the required complexity for the purpose of the ontology. These and other problems led to several approaches of (semi) automatic ontology learning from text. This discipline tries to use the externalized knowledge in form of documents to extract concepts and their relations and attributes. Some of these approaches provided useful results in the past.
Some examples show, that the approaches of ontology mining from text provide sometimes better results than ontologies modeled by an ontology engineer (regarding number of found concepts and some special relations/axioms). However, there are points that make the automatic ontology learning from text quite difficult. On the one hand the complexity of natural language allows only a narrow amount of extractable information. On the other hand, faults, ambiguity and lack of quality result in a strong noise. Thus the extraction of structured facts from natural language is complicated. In order to avoid the challenge of modeling a common understanding by an individual and also to avoid the problem of learning ontologies from text, some approaches enable a community to model their knowledge in a semiformal manner. An example is given by Web 2.0 techniques. Through tagging of informal knowledge artifacts, the user is enabled to enhance unstructured information with structural knowledge. By adding tags to all kinds of resources, these objects are provided with meaningful semantic markup. In addition, by setting relations between objects with identical (similiar) tags, graphs of objects are established. Thus folksonomies can be seen as a form of community modeled ontologies [4] with a low level of formalization. Today's Web 2.0 portals provide depending on the chosen object a lot of recommended objects related through automatic evaluation of user given tags. Objects like videos on YouTube pictures on Flickr or links on del.icio.us get a relation through having the same (or similar) tags[4].

Externalization of Knowledge

Ideally the externalized knowledge is modeled according to the structures of the mind. The prevailing perception of the human mind is that knowledge is represented in structures [9]. Concepts and individuals (instances) are associated with each other. Searching for a certain piece of information –- i.e. recalling something -– is, from the point of view of many researchers, very similar to the concept of spreading activation The fact that spreading activation algorithms are used in cognitive psychology [5] as well as in information retrieval[6] proves the fact that the mind and associative networks (models) are very similar. Recalling or finding the required information implies two preconditions: (1)the concept / object is existent in the model, (2)there is a valid path from the start concept

This work does not include research on the cognition of humans; this work is rather driven through the insight that the representation of knowledge in structures is natural. This clarifies the origin of the trend of structuring and modeling knowledge. Indeed, today's Web 2.0 approaches enable the user to provide unstructured information like text, audio or video but on the other hand the user is enabled to relate the content with other objects. Mostly the content is related to the profile of an author. By means of tags, content is related to other content –- clicking on a tag leads to other content annotated with the same tag. In fact, today's web provides tools and interfaces for collaborative modeling of ontologies, but in most cases only in combination with unstructured knowledge artifacts and on a low level of formalization.

Conclusion and Outlook

Collaborative modeling using Web 2.0 techniques facilitates the externalization of structured knowledge and enables communities to develop meaningful ontologies. These ontologies differ from traditional ontologies in many ways. Part of future work will be to explore the characteristics and features of ontologies modeled by communities. Therefore the transformation of a social ontology over a certain amount of time will be analyzed. The main goal of this work is to apply these findings for supporting informal modeling approaches in collaborative environments,

Acknowledgement

MATURE is partially funded under the FP7 of the European Commission within the unit for Technology-Enhanced Learning (TEL)..

The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

References

1. Gruber T. R.. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2):199-220, 1993
2. Fensel D,. Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. Springer 2001
3. Uschold M,. Grninger M. Ontologies: principles, methods, and applicaTions. Knowledge Engineering Review, S. 93155, Vol. 11, Nr. 2.1996
4. Gruber T.R.,. Ontology of Folksonomy: A Mash-up of Apples and Oranges. Int'l Journal on Semantic Web and Information Systems, 3(2), 2007
5. Anderson J. R., . A spreading activation theory of memory. Journal of Verbal Learning and Verbal Behavior 22, 261-295. 1983
6. Crestani F.,. Application of Spreading Activation Techniques in Information Retrieval Artificial Intelligence Review, 1997
7. Braun S,. Schmidt A,: Walter A., Zacharias V.. The Ontology Maturing Approach to Collaborative and Work-Integrated Ontology Development: Evaluation Results and Future Directions International Workshop on Emergent Semantics and Ontology Evolution (ESOE), ISWC 2007, Busan/Korea, 2007.
8. Fischer G., McCall R., Ostwald J. L., Reeves B., Shipman III F.M.: Seeding, evolutionary growth and reseeding: supporting the incremental development of design environments. CHI Conference Companion: 220. 1994
9. Anderson J. R., Bothell D., Byrne M. D., Douglass S., Lebiere C., & Qin Y. An integrated theory of the mind. Psychological Review 111, (4). 1036-1060. 2004