

# Research on Key Techniques for Ontology Construction System

Xiaoya Deng

School of Computer, Sichuan University of Arts and Science, Dazhou 635000, China

xiaoya@163.com

**Keywords:** Ontology construction, ontology representation, concurrent control, object duration, key techniques.

**Abstract.** The paper proposes the key techniques for ontology construction system. The current situation of tools for ontology construction is analyzed, and then a collaborative ontology construction system framework is given. The persistence layer is introduced between the business logic layer and data layer, which decouples data manipulation from business logic, provides object services for upper layer and interacts with underlying storage by defining the mapping between objects and underlying data. Future more, two key technologies of the system including ontology representation and its mapping to the relational data model and concurrent control method based on lock subdivision are deeply researched. Implementation of object persistence requires defining object structures, storage structures and the mapping files between them, so an intermediate representation model of ontology in the persistence layer is introduced, and corresponding data structures are defined according to this model. Then the relational database model of storing ontology is given, as well as the mapping from relational data model to the intermediate representation model. According to the needs of operating ontology objects collaboratively, a concurrency control method based on lock subdivision is presented to further improve the concurrency ability of the system, which adds different type of locks for objects according to the compatibility between various operations of them. The relevant algorithm is designed at last.

## Introduction

Enterprise applications benefit from semantically enriched enterprise knowledge, and ontology can provide shared conceptual models for various symbols in knowledge management systems and are widely applied. With the increasing of ontology scale, tool support is very important both for the ontology development process and for the ontology usage in applications. However, many existing tools cannot well support collaboration, and the ontology is incompatible because of the different format of ontology they build. [1] So, a collaborative ontology construction system is designed aiming at improving the ability of collaboration and ontology development efficiency.

Ontology, as a tool for knowledge representation, is widespread concerned, which is in-depth researched and applied in many fields of information science, knowledge engineering, artificial intelligence, digital libraries. With the promotion of the depth research and application, more and more ontology need to be built. As building ontology require a lot of ontology resources, which are described in semantic field relationships between things, for ontology retrieving and application.

According to the characteristics of the status quo of the Internet, the vast majority of data is stored in a relational database. However, the ability of describing knowledge or data by using relational data model is limited, [2] because it's unable to meet the users' increasingly requirements of complex information query: But, the relational data model has a few good features, such as the format of the database relational tables is neater than the other data, among the data have semantic information and higher field related degrees, etc., so it's easier to extract ontology based the database and to build ontology than from other data sources. But the database resources are generally massive, which is more difficult to be processed and implemented. In addition, the construction and edit of the ontology needs experts in specific fields to participate in, and during the construction process of an ontology library, it requires a lot of work to do such as manual search, entering, instantiating, which is an arduous and difficult task and can easily lead knowledge acquisition to be a bottleneck problem.

## Overall framework of the system

Since 1990s, ontology has been introduced in Artificial Intelligence, knowledge Engineering and Library & Information sciences. Ontology learning includes the research on knowledge representation and knowledge organization in the domains mentioned above. In Library & Information science, ontology is a vocabulary with the specification of relationships between terms about a special domain. Ontology representation class diagram is shown in Fig. 1 [3].

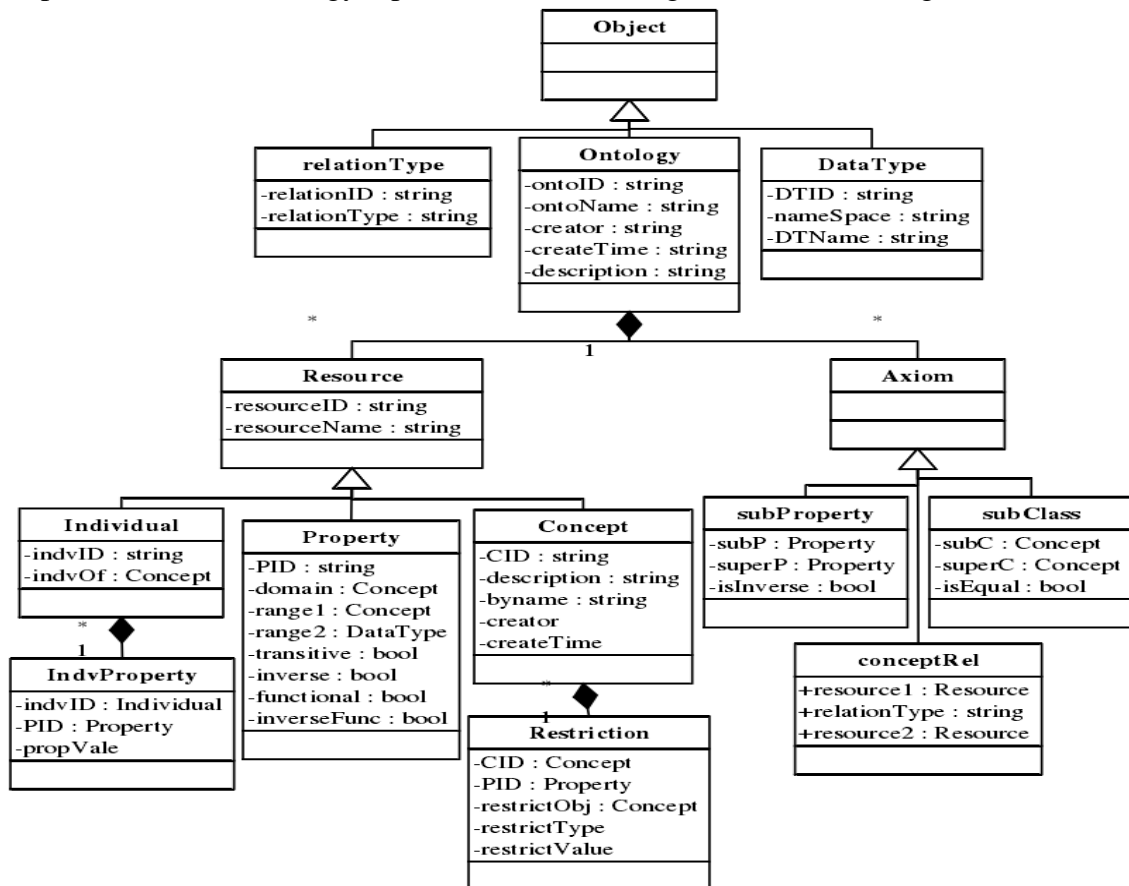


Fig. 1 Ontology representation class diagram

The latest fruits of ontological theory and method have been described systematically on the basis of intensive investigation of large number of documents and web pages on ontology learning in this report since 2004. [4] The fact, trends, applications, main technologies, institutes which majoring in this field have all been discussed in this report. Standardization of ontological technology, semantic annotation based-on ontology, and top ontology transformation in OWL DL, have been studied, tested and analyzed deeply. Ten hot issues in ontology construction have been discovered. "Ontology on ontology" had been provided.

In event ontology, "event" is not merely treated as static concepts or relations between concepts. It is the basic unit of knowledge, which contains information like action, object, environment, as well as time. [5] The concept "event" in event ontology is responsible for representation of relations between different events and relations between people or things involved in event. Furthermore, it should express the roles of event participants and dynamic procedure of event. Based upon six-tuples expression of event, this paper focuses on the key problems in construction of event ontology. Specifically, the Operating compatibility is shown in Table 1.

Table 1 Operating compatibility

Basic operation	Not compatible operating
Addsub (c1/c2)	Addsub delconcept
Delconcept(c)	N(c)
Editconcept	Delconcept editconcept
Queryconcept(c)	Delconcept editconcept
Addcproperty(c,p)	Addproperty queryproperty
Delcproperty(c,p)	Delcproperty delconcept
Addobjprop(p)	Delobjprop addobjprop
Editobjprop(p)	Editobjprop queryobjprop
Editdtprop(p)	Editdtprop querydtprop
Delobjprop(p)	N(n)
Deldtprop(p)	N(n)
Addh(c1,c2)	Queryh delconcept
Addrelation(c1,c2,r)	Addrelation querylation

Based on floricultural literature tentative ontology-based prototype, agricultural thesauri based knowledge base, and opencyc open. source project based top ontology, LODE (Large ontology Development Environment) has been developed. Oracle DB technology and .NET have been used in developing the functions of knowledge entrance and semantic annotation. The general approach of domain ontology construction had been studied and summarized.

### Ontology representation model and database

Current database to ontology transform work is focus on the database semantic excavation. [6] How automatically extracts the semantic relations in the database and carries on the semantic extension, memory, and query are urgently need to solve. This paper proposed a database to ontology automatic extraction and semantic query method. This method is not only extract the database semantic relations automatically but also using Jena and layer reasoning strategy solves magnanimous ontology memory and query difficulty The system's realization proved that this method can realize well on the extracted ontology's query duty.

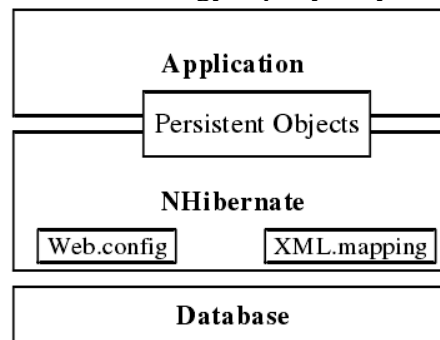


Fig. 2 The architecture of NH

The architecture of NH is shown in Fig. 2. The design follows as:

(1) Reuse of entities from existing ontology and representation of event's dynamic procedure based on concept algebra. In event, both object-element and environment-element are entities of existing ontology. To reuse those existing entities when constructing event ontology, the paper extends operations of concept algebra via adding "time flag" and "information quantity". Representation of event's dynamic procedure can be considered as representation of a series of event's states changing along with different time points, and then assertion-element can essentially express event's states. By taking advantage of extended operations (inheritance, tailoring, extension and substitute) in concept algebra, we can change event's states along with different time points. In short, the application of extended concept algebra in cognitive informatics greatly raises the efficiency when constructing event ontology.

(2) Frame-based representation model of event and analysis of conceptual operations. This paper proposes a novel frame-based representation model of event on the basis of six-tuples expression,

analyzes several conceptual operations, and defines the upper classification of events. Based on Nilsson's concept algebra, the frame-based representation model is able to express both event class and instance. Furthermore, it can express dynamic procedure of event and the relations between events and their elements. The frame-based representation model is flexible, effective and widely applicable. It is appropriate for representation of event-based knowledge and extremely provides theoretical support for applications like knowledge storage and reasoning.

(3) Event-based tagging and analytic technique for Chinese text. Building corpora of domain knowledge plays an important part in construction of ontology. Similarly, event-oriented corpora tagging is essential for construction of event ontology. In order to overcome shortcomings of existing Chinese analytic technique and facilitate the build of event ontology, the paper presents an event-based tagging and analytic technique for processing Chinese text. In comparison with the tagging method of CEC (Chinese Event Corpus), it enhances the tagging ability and provides high quality corpora for event auto-recognition, auto-classification and retrieval of event elements or relations.

### **System design and implementation**

Ontology construction and ontology application is a pair of contradictory unity based on the discovery of 10 hot issues in ontological theory and practice.

Cycle ontology is a more logical top ontology than SUMO based on the transformation and comparison.

LODE aims at collaborative development. The functions of knowledge entrance and semantic annotation have been developed. The searching function is under developing. Though manual work is still in need, the research in this field has really been promoted.

Finally, the collaborative ontology construction system is implemented, and is applied to a specific project, completing the construction of the ontology about knowledge of truck designing.

### **Conclusions**

Currently works on ontology is still attracting research experts and scholars. This thesis' work is expected to play in this field to promote the development of the role. In this paper, the proposed method has a high theoretical and practical value. Hopefully, it will benefit the future research work on ontology engineering research area.

### **References**

- [1] Alicia Diaz, Guillermo Baldo and Gerome Canals. Co-Protege: Collaborative Ontology Building with Divergences. 17th International Conference on Database and Expert Systems Applications (DEXA'06), 2006: 156-160.
- [2] York Sure, Michael Erdmann, Juergen Angele, Steffen Staab and Dirk Wenke. OntoEdit: Collaborative Ontology Development for the Semantic Web. Lecture Notes in Computer Science, 2012, 2342: 221-235.
- [3] Agrawal R, Somani A, Xu Y. Storage and querying of ecommerce data. In: Peter M G A Paolo A, eds. Proc of the 27th VLDB. Roma: Morgan Kaufman Publishers Inc, 2011: 149-158.
- [4] S.L.Li. Ontology Learning for Chinese Documents Based on SVD and Concept Clustering. Journal of Beijing Institute of Technology. 2013, 12: 178-183.
- [5] Liu Wei, Xu Wenjie, Fu Jianfeng, et al., An extended description logic for event ontology. Proceedings of the 5th International Conference on Grid and Pervasive Computing (LNCS), Springer, 2010: 471-451.
- [6] Cimiano P., Hotho A., Staab S., Learning concept hierarchies from text corpora using formal concept analysis. Journal of Artificial Intelligence Research, 2005, 24(1): 305-339.