Chapter I

Ontology Extraction Using Views for Semantic Web

Carlo Wouters, La Trobe University, Australia
Rajugan Rajagopalapillai, University of Technology, Sydney, Australia
Tharam S. Dillon, University of Technology, Sydney, Australia
Wenny Rahayu, La Trobe University, Australia

Abstract

The emergence of Semantic Web (SW) and the related technologies promise to make the Web a meaningful experience. Conversely, success of SW and its applications depends largely on utilization and interoperability of well-formulated ontology bases in an automated heterogeneous environment. This creates a need to investigate utilization of an (materialized) ontology view as an alternative version of an ontology. However, high level modeling, design and querying techniques still proves to be a challenging task for SW paradigm, as, unlike classical database systems, ontology view definitions and querying have to be done at high-level abstraction. In order to address such an issue, in this chapter, we describe an abstract view formalism for
SW (SW-view) with conceptual and logical extensions. SW-views provides the needed conceptual and logical semantics to engineer ontology bases using three levels of abstraction, namely (1) conceptual, (2) logical/schema and (3) instance levels. We first provide the view model and its formal properties including a set of conceptual operators, which enable us to do ontology extraction at the conceptual level. Later, we provide a schemata transformation methodology to materialize SW-views under the Ontology Extraction Methodology (OEM) framework.

Introduction

Meaning of data is emerging as the main area of interest in the awake of meaningful Web era, which is the Semantic Web (SW) paradigm (W3C-SW, 2005a). As envisage by Berners-Lee (1998), SW is emerging as the new medium for the decentralized, automated global information sources for the new 21st century information-driven economies (Aberer et al., 2004). This is highly visible in the exponential increase of new research directions in engineering ontologies in a wide spectrum of domains ranging from traditional enterprise data to time-critical medical information and infectious decease databases. For such vast ontology bases to be successful and to support autonomous computing, in a meaningful distributed environment, the preliminary design and engineering of such ontologies should follow strict software engineering disciplines. Furthermore, supporting technologies for ontology engineering such as data extraction, integration and organization have be matured to provide adequate modeling and design mechanism to build, implement and maintain successful techniques. For such purpose, Object-Oriented (OO) paradigm seems to be an ideal choice as it has been proven in many other complex applications and domains (Dillon & Tan, 1993; Graham, Wills, & O’Callaghan, 2001).

OO conceptual models have the power in describing and modeling real-world data semantics and their interrelationships in a form that is precise and comprehensible to users (Dillon & Tan, 1993; Graham et al., 2001). But the existing OO modeling languages (such as UML [OMG-UML™, 2003a]) provide insufficient modeling constructs for engineering SW models and applications. This is mainly due to lack of inherent support for semistructured schema-based data descriptions and constraints in OO modeling languages and the shortcomings of many semistructured data models in providing visual modeling and higher levels of abstraction semantics (such as conceptual models) that are easily understood by humans. Due to this, in the Semantic Web paradigm, most modeling and design constructs are modeled at a lower level of abstraction, namely schema or data description language levels.