International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099

VOL 1 ISSUE 1 (2016) PAGES 29 - 35

Received: 19/07/2016. Published: 27/07/2016

A Survey of Ontology Knowledge for semantic web in Data mining

[a *] B. Kalaiselvi, [b] R. Durairam

[a] Dept. of Computer Science and Engg. Mahendra Engineering College for women

[b] Dept. of Computer Science and Engg.Knowledge Institute of Technology

[a] kalairs2003@gmail.com, [b] slmvj08@gmail.com

Abstract: Ontologies have become common on the World-Wide Web. Research on ontology is becoming increasingly widespread in the computer science community, and its importance is being recognized in a multiplicity of research fields and application areas, including knowledge engineering, database design and integration, information retrieval and extraction. Ontologies are a crucial tool for formally specifying the vocabulary and relationship of concepts used on the Data Mining. We survey the recent development of the field of ontologies in AI. We point to the somewhat different roles ontologies play in information systems, natural language understanding, and knowledge based systems. In this paper, we survey the ontology knowledge and we classify them according to their intent, the way they express the ontology, translation mapping, integration and conceptualization.

Keywords:Ontology, conceptualization, translation, mapping.

1. Introduction

1.1 Overview and Motivation

Using ontologies, web-based agents can treat web documents as sets of assertions, and, in particular, draw inferences from them. More and more ontologies are being developed as formal underpinnings for RDF-based data. An obvious desideratum of this movement is that two ontologies should not cover the same area; instead, those interested in formalizing descriptions in that area should agree on a standard set of concepts. However, this goal cannot always be met, for a variety of reasons. Some standard vocabularies arise in different parts of the world or among different linguistic communities, and attain popularity before their overlap is noticed. Even more likely is that two vocabularies will partially overlap, usually because what is central to one is peripheral to the other. Ontology is a representation vocabulary, often specialized to some domain or subject matter. More precisely, it is not the vocabulary as such that qualifies as an ontology, but the conceptualizations that the terms in the vocabulary are intended to capture. Thus translating the terms in an ontology from one language to another. In the term ontology is sometimes used to refer to a body of knowledge describing some domain, typically a commonsense knowledge domain, using a representation vocabulary. In other words, the representation vocabulary provides a set of terms with which to describe the facts in some domain, while the body of knowledge using that vocabulary is a collection of facts about a domain. At times, theorists use the singular term to refer to a specific set of terms meant to describe the entity and relation-types in some domain. Thus, we might speak of an ontology for “liquids” or for “parts and wholes.” Here, the singular term stands for the entire set of concepts and terms needed to speak about phenomena involving liquids and parts and wholes.

When different theorists make different proposals for an ontology or when we speak about ontology proposals for different domains of knowledge, we would then use the plural term ontologies to refer to them collectively. In AI and information-systems literature, however, there seems to be inconsistency: sometimes we see references to “ontology of domain” and other times to “ontologies of domain,” both referring to the set of conceptualizations for the domain.

1.2 Technology and Use of ontologies

There have been several recent attempts to create engineering frameworks for constructing ontologies. KIF (Knowledge Interchange Format), an enabling technology that facilitates expressing domain factual knowledge using a formalism based on augmented predicate calculus. Predicate calculus facilitates the representation of objects, properties, and relations. Variations such as situational calculus introduce time so as to represent states, events, and processes. If we extend the idea of knowledge to include images and other sense modalities, we might need radically different kinds of representation. For now, predicate calculus provides a good starting point for ontology- sharing technologies. Using a logical notation for writing and sharing ontologies does not imply any commitment to implementing a related knowledge system or a related logic. We are simply taking a knowledge-level5 stance in describing the knowledge system, whatever the means of implementation. In this view, we can ask of any intelligent system, even one implemented as a neural network, “What does the system know?”

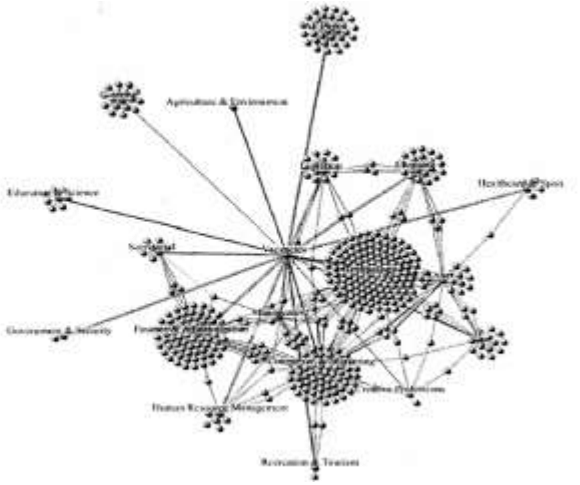

Fig.1.2.1 Data-set organized by economic sector

Received: 19/07/2016. Published: 27/07/2016

In AI, knowledge in computer systems is thought of as something that is explicitly represented and operated on by inference processes. However, that is an overly narrow view. All information systems traffic in knowledge. Any software that does anything useful cannot be written without a commitment to a model of the relevant world to entities, properties, and relations in that world. Data structures and procedures implicitly or explicitly make commitments to a domain ontology. It is common to ask whether a payroll system "knows" about the new tax law, or whether a database system "knows" about employee salaries. Information-retrieval systems, digital libraries, integration of heterogeneous information sources, and Internet search engines need domain ontologies to organize information and direct the search processes. For example, a search engine has categories and subcategories that help organize the search. The search-engine community commonly refers to these categories and subcategories as ontologies.Object-oriented design of software systems similarly depends on an appropriate domain ontology. Objects, their attributes, and their procedures more or less mirror aspects of the domain that are relevant to the application. Object systems representing a useful analysis of a domain can often be reused for a different application program. Object systems and ontologies emphasize different aspects, but we anticipate that over time convergence between these technologies will increase. As information systems model large knowledge domains, domain ontologies will become as important in general software systems as in many areas of AI. Ontologies are useful in NLU in two ways. First, domain knowledge often plays a crucial role in disambiguation. A well designed domain ontology provides the basics for domain knowledge representation. In addition, ontology of a domain helps identify the semantic categories that are involved in understanding discourse in that domain. For this use, the ontology plays the role of a concept dictionary. In general, for NLU, we need both a general-purpose upper ontology and a domain-specific ontology that focuses on the domain of discourse (such as military communications or business stories). CYC, Word net, 8 and Sensus15 are examples of sharable ontologies that have been used for language understanding. Information systems and NLU systems need factual knowledge about their domains of discourse. The inferences they make are usually simple. Problem-solving systems, in contrast, engage in complex sequences of inferences to achieve their goals. Such systems need to have reasoning strategies that enable them to choose among alternative reasoning paths. Ontology specification in knowledge systems has two dimensions:

•Domain factual knowledge provides knowledge about the objective realities in the domain of interest (objects, relations, events, states, causal relations, and so forth).

•Problem-solving knowledge provides knowledge about how to achieve various goals. A piece of this knowledge might be in the form of a problem-solving method specifying in a domain-independent manner how to accomplish a class of goals.

2. Ontology and ontologies

"Ontology is a fascinating discipline" and "ontology" (with the lowercase "o"), as in the expressions "Aristotle's ontology" or "CYC's ontology". The same term has an uncountable reading in the former case, and a countable reading in the latter. While the former reading seems to be reasonably clear (as referring to a particular philosophical discipline), two different senses are assumed by the philosophical community and the Artificial Intelligence community (and, in general, the whole computer science community) for the latter term. An ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words. This set of assumptions has usually the form of a first-order logical theory, where vocabulary words appear as unary or binary predicate names, respectively called concepts and relations. In the simplest case,an ontology describes a hierarchy of concepts related by subsumption relationships; in more sophisticated cases, suitable axioms are added in order to express other relationships between concepts and to constrain their intended interpretation.

2.1 What is a conceptualization

A conceptualization has been defined in a well-known AI textbook [17] as a structure <D,R>, where D is a domain and R is a set or relevant relations on D6. This definition has been then used by Tom Gruber, who defined an ontology as “a specification of a conceptualization”. The notion of conceptualization making clear the relationship between an ontology, its intended models, and a conceptualization. The problem with notion of conceptualization is that it refers to ordinary mathematical relations on D, i.e. extensional relations. These relations reflect a particular state of affairs: for instance, in the blocks world, they may reflect a particular arrangement of blocks on the table. We need instead to focus on the meaning of these relations, independently of a state of affairs: for instance, the meaning of the above relation

lies in the way it refers to certain couples of blocks according to their spatial arrangement reserving the simple term relation to ordinary mathematical relations.

Let **C** = <D, W, Â> be a conceptualization. For each possible world *w*ÎW, the intended structure of w according to **C** is the structure **S***w***C** = <D, **R***w*C>, where **R***w***C**={r(*w*) | rÎÂ }is the set of extensions (relative to w) of the elements of Â. The set {**S***w***C** | *w*ÎW} all the intended world structures of **C**. Let us consider now a logical language **L**, with a vocabulary V. Rearranging the standard definition, we can define a model for **L** as a structure <**S**, I>, where **S** = <D, **R**> is a world structure and I: V®DÈ**R** is an interpretation function assigning elements of D to constant symbols of V, and elements of **R** to predicate symbols of V. As well known, a model fixes therefore a particular extensional interpretation of the language. we can fix an in tensional interpretation by means of a structure <**C**, Á>, where **C** = <D, W, Â> is a conceptualization and Á: V®DÈÂ is a function assigning elements of D to constant symbols of V, and elements of Â to predicate symbols of V. We shall call this in tensional interpretation an ontological commitment for L. If **K** = <**C**, Á> is an ontological commitment for **L**, we say that **L** commits to **C** by means of **K**, while **C** is the underlying conceptualization of **K**9. A set of intended models is therefore only a weak characterization of a conceptualization: it just excludes some absurd interpretations, without really describing the “meaning” of the vocabulary.

2.2 What is an ontology

We can now clarify the role of an ontology, considered as a set of logical axioms designed to account for the intended meaning of a vocabulary. Given a language **L** with ontological commitment **K**, an ontology for **L** is a set of axioms designed in a way such that the set of its models approximates as best as possible the set of intended models of **L** according to **K.** Therefore, an ontology can specify a conceptualization only in a very indirect way, since it can only approximate a set of intended models and such a set of intended models is only a weak characterization of a conceptualization. An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models.

The relationships between vocabulary, conceptualization, ontological commitment and ontology are illustrated in Fig.2.2 It is important to stress that an ontology is language dependent, while a conceptualization is language-independent. In its de facto use in AI, the term “ontology” collapses the two aspects, but a clear separation between them becomes essential to address the issues related to ontology sharing, fusion, and translation, which in general imply multiple vocabularies and multiple conceptualizations.

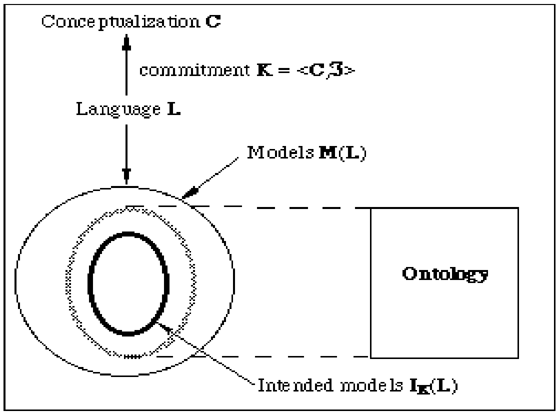

Fig.2.2.1 An ontology indirectly reflects this commitment by approximating this set of intended models.

2.3 Ontology Integration Problem

Assuming that each system has its own conceptualization, a necessary condition in order to make an agreement possible is that the intended models of the original conceptualizations overlap. Supposing now that these two sets of intended models are approximated by two different ontologies, it may be the case that the two ontologies overlap while the intended models do not. This means that a bottom-up approach to systems integration based on the integration of multiple local ontologies may not work, especially if the local ontologies are only focused on the

Received: 19/07/2016. Published: 27/07/2016

conceptual relations relevant to a specific context, and therefore they are only weak and ad hoc approximations of the intended models. Hence, it seems more convenient to agree on a single top-level ontology rather than relying on agreements based on the intersection of different ontologies.

To develop different kinds of ontology according to level of generality, • Top-level ontologies describe very general concepts like space, time, matter, object, event, action, etc., which are independent of a particular problem or domain: it seems therefore reasonable, at least in theory, to have unified top-level ontologies for large communities of users.

Domain ontologies and task ontologies describe, respectively, the vocabulary related to a generic domain (like medicine, or automobiles) or a generic task or activity (like diagnosing or selling), by specializing the terms introduced in the top-level ontology.

Application ontologies describe concepts depending both on a particular domain and task, which are often specializations of both the related ontologies. These concepts often correspond to roles played by domain entities while performing a certain activity, like replaceable unit or spare component.

3. Ontology Translation and Ontology

Mapping

The process of finding correspondence between the concepts of two ontologies. If two concepts correspond, they mean the same thing, or closely related things. Obviously, finding such mappings can be a valuable preprocessing step in solving the ontology-translation problem for the two ontologies. Automating the process of ontology mapping is an active area of research. However, the emphasis on finding mappings has led, we believe, to a distorted view of the translation problem. Suppose one starts by assuming a particular notation for ontologies, such as OWL. These notations are often represented visually as graph structures. Then it is natural to express a mapping between two ontologies as a network of “meta-links” that join nodes and links in one ontology graph with nodes and links in the other . Each such link can be annotated with labels specifying whether the two constructs it joins mean exactly the same thing, or which of them covers a larger set of objects or relationships. It seems almost inevitable at that point to think of translation as a process of substitution of labels. One might identify a role R1 in one ontology with a role R2 in another. (In the terminology of relational databases, the columns R1 and R2 would be labelled as equivalent.) Translation then becomes a matter of relabelling data. Complexities arise when the labels aren’t exactly equivalent. Two classes might be connected by a subclass link instead of being exactly equivalent. This process is usually thought of as translating the query rather than translating the data used to answer it. The rules required to translate a query are essentially logical axioms (or logic-programming rules) of the form A ← B, where A is in the query vocabulary and B is in the vocabulary of the remote database. in a forward direction. From this point of view, translation rules are just axioms. The purpose of ontology mapping should be to find these axioms.

Received: 19/07/2016. Published: 27/07/2016

4. Conclusion and further study

This study has introduced a technology, use and kinds of ontology, conceptualization, as well as problem of ontology integration, translation and mapping the ontology. The revisitation of the recent definitions concerning the notion of ontology was not planned in advance. we will give some conclusions for our work so far and discuss our future plans for developing interactive tools for ontology merging based on our recent work on integrating different neuronal databases.

References

1. Casati, R., Smith, B., and Varzi, A. “Ontological Tools for Geographic Representation and Formal Ontology in Information Systems” IOS PressI. 1998

2. Artale, A., Franconi, E., Guarino, N., Pazzi, L ”Part-Whole Relations in Object-Centered Systems: an Overview. Data and Knowledge Engineering”, 20(3): pp 347-383, 1996.

3. Bateman, J. A”On the Relationship Between Ontology Construction and Natural Language: a Socio-Semiotic View” International Journal of Human-Computer Studies, vol:43 pp:929-944, 1995.

4. D. McDermott, M. Burstein, and D. Smith ”Overcoming ontology mismatches in transactions with self-describing agents” In Proc. Semantic Web Working Symposium, pages 285–302, 2001.

5. P. Mitra, G. Wiederhold, M. Kersten ”A graph-oriented model for articulation of ontology interdependencies” In Proceedings of Conference on Extending Database Technology (EDBT 2000), 2000.

6. T.R. Gruber, “A Translation Approach to Portable Ontology Specifications,” Knowledge Acquisition, Vol. 5, pp. 199–220, 1993.

7. D. Fensel et al., “Using Ontologies for Defining Tasks, Problem-Solving Methods and Their Mappings” Knowledge Acquisition, Modeling and Management, pp. 113–128, 1997.

8. S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure. Knowledge processes and ontologies. IEEE Intelligent Systems, 16(1), 2001.

9. A. Doan, J. Madhavan, P. Domingos, A. Halevy ”Learning to map between ontologies on the semantic web” In Proceedings of the World-Wide Web Conference ,(WWW-2002), 2002.