Characterizing Quality of Knowledge on Semantic Web

Kaustubh Supekar
Department of Computer Science
University of Manchester
Oxford Road, Manchester M13 9PL, UK
ksupekar@cs.man.ac.uk

Chintan Patel, Yugyung Lee
School of Computing and Engineering
University of Missouri-Kansas City
copdk4, leeyu @umkc.edu

Abstract

The Semantic Web is intended for knowledge sharing among agents as well as humans. To achieve this goal, Ontologies, which express knowledge in a certain vitality as well as in a machine interpretable form, were introduced. The growing demand for facilitating deployment and reuse of Ontologies has increased the need to develop adequate criteria to measure the quality of Ontologies conceptualizing a domain. Our research is motivated by the urgent needs of rigorous mechanisms which analyze ontological features from diverse perspectives and determine their quality levels.

This paper presents the methodology representing qualitative and quantitative analysis of Ontologies and their classification. The Ontological tools, which were implemented based on the methodology, are proposed to provide multiple interfaces to humans and agents, thus supporting Ontology Engineering process. The proposed framework has gone through a great deal of testing and evaluation processes in the context of a real application of Ontology analysis and classification.

Introduction

Semantic Web is meant for machines to share knowledge via commonly agreed description of the concepts and their relationships (Berners-Lee, Hendler, & Lassila. 2001), which implies that machines or Software agents are responsible for validating and trusting the information source. In an independent environment such as Web, where there are no restrictions on the information being published, it becomes the liability on the part of the consumer to accurately judge the quality and validity of the information provider. Since the information is provided by many different sources, validity of information becomes an important issue. Currently, there are no appropriate mechanisms to measure the quality of knowledge over the Semantic Web. Hence we believe that there is an urgent need to model the features that characterize validity and quality of the knowledge.

In Semantic Web, the semantics are incorporated on Web using Ontologies. Ontology Engineering (Conceptual Modelling) is an exceedingly intricate and challenging task requiring specialized design skills as well as comprehensive domain knowledge. To ease the burden, Semantic Web allows one to reuse or extend existing Ontologies present on Web. However, Knowledge Engineers are often confronted with multitude of Ontologies for the required domain; hence a methodology for quantitative and qualitative analysis of alternatives is required. Moreover, to solve such a multidimensional problem we also need rigorous schemes to make a correct and optimum choice.

Our hypothesis is that in a cognizant environment agents are more likely to commit to quality knowledge. A cognitive process allows an individual to perceive processes and objects within its environment and to establish relationship between the actually and the formerly perceived things (Foerster 1969). Knowledge is a cognitive model and structure is a byproduct of all cognitive process at a certain point in time. Ontologies (Gruber & Olsen 1994), specification of conceptualization are developed with the motive of sharing knowledge among the agents. Different agents, human or machines have varying conceptualization of their domain of interests. For instance, one ontology may classify a *Vehicle* into *Compact*, *SUV*, *Mid_Size* and *Sports_Car* while the other ontology classifies on the basis of brand name as *Ford*, *BMW*, *Mercedes*. Agents that seek particular information about a *Vehicle* should be guided to choose the *right* ontology.

Determining that a particular Ontology is the *right* one for an application is a critical task. How do we define and agree on the *right* ontology? One could argue that *right* ontology is purely subjective because meaning of quality differs across domain (content), users, and situation (context). For instance, we characterize quality of software different from quality of an aircraft. Related to this issue, we could have further questions on quality of Ontologies: How can we prove that a particular ontology is complete? Can we measure Ontologies? These problems can be attributed to lack of structured definition of quality. These problems become more aggravated as we are characterizing an abstract concept - knowledge. If we presume the existence of a methodology for ontology characterization, it may support to choose one authoritative ontology that can serve impelling purposes, and on which some consensus between different groups can be reached.

In this paper, based on our perception of the quality of Ontological knowledge being useful resources for determining the *right* ontology, we propose a model that characterizes quality of knowledge by defining a *Quality of Knowledge* ontology and compute the *Degree of Quality*. A formal method that determines the level of quality of ontology based on characterization of knowledge has been designed. We have implemented the framework that searches, classifies and analyzes a set of domain Ontologies to quantify the features characterized by the proposed model. We also give experimental results that characterize quality features for tourism Ontologies available over the Web.

Related Work

A variety of Ontologies have been introduced with emphasis on their intended use for knowledge sharing and reuse. While various definitions on ontologies have been introduced across different communities (Guarino 1998), the widely accepted ontology definition is a specification of a conceptualization (Gruber 1993). Several studies have been conducted on knowledge representation languages and ontologies (Maedhe & Staab 2001), design principles and models for Ontologies (Noy & Hafner 1997) and the sharing and reuse of ontology (Benjamin & Fensel 1998). The verification and evaluation of ontology have already been explored: (Kalfoglou & Robertson 1999) presented ontological engineering issue revealing a potential error including the syntax and semantics of the ontology using constraints.

The selection of the right ontology for an application is studied in order to support the prospective user who investigates to what extent the ontology supports the intended task. Specifically, there are some researchers focused on qualitative and quantitative analysis of ontologies. (Noy & Hafner 1997) has described a framework for comparing ontologies, including CYC, UMLS, WordNet etc. Maganaraki et al. (Maganaraki *et al.* 2002) has proposed quantitative criteria about ontological structure which can be distinguished from qualitative comparison criteria. Towards this direction, (Chandrasekaran, Josephson, & Benjamins 1998) has addressed the reusability issue of ontologies in similar applications.

Growth of Semantic Web has led to massive growth in use and development of Ontologies over Web. The central idea of Ontology Engineering in Semantic Web focused on extensive reuse of existing Ontologies. Recently, several tools for Ontology engineering have been developed; Protégé-2000[1], OntoEdit[2], OilEd[3]. Other related tools have been built for Ontology merging-PROMPT (Noy & Musen 2000), Ontology Management and Querying KAON[4]. In contrast to the earlier work, with the new paradigm of Ontology Engineering over Web, our contribution is a model to assist Knowledge Engineers and Software Agents in selecting the *right* ontology. In a highly distributed environment such as Web, a critical task is to search and retrieve the required information. Analogous to currently successful search techniques (e.g., Google), our tool, OntoKhoj (K. Supekar, C. Patel, & Y. Lee 2003) searches and ranks Ontologies (RDF, DAML+OIL, OWL) over web. We also use the existing Classification algorithms (KNN, Naïve Bayes etc) to classify the crawled Ontologies into respective domain hierarchy (derived from Open Directory Project[5])

[1]http://protege.stanford.edu/

[2]http://www.ontoknowledge.org/tools/ontoedit.shtml

[3]http://oiled.man.ac.uk/

[4]http://kaon.semanticweb.org/

Characterizing Ontological Features

Knowledge Quality Features

In this section we describe our model that characterizes quality of knowledge, which ultimately forms the basis of determining correctness of ontology. The term *quality* is defined as a distinguishing characteristic or a property. As mentioned previously, meaning of quality differs across users, domain and situation. The problem becomes more difficult when an abstract concept - knowledge need to be characterized. We believe that a structured definition of quality may help to resolve this issue. Thus, we propose an ontology of features (Figure 1) that characterizes quality of knowledge on Semantic Web. The feature set is broadly classified into *quantifiable* and *non-quantifiable* knowledge characteristics that determine quality of ontology. The limitation of subset of features that can only be evaluated through empirical ways prompted this sub-classification.

First of all, we identify two knowledge features as non-quantifiable features. **Cognitive Adequacy** is how well the ontology reflects the domain of interest. Analyzing this feature requires determining how fit the ontology is for the purpose. Assessment can only be done through empirical ways of testing thee effectiveness of applications based on the ontology. **Context** refers background information about the domain of interest that an agent needs to know before they seek knowledge about the domain. If one incorporates more background information into the ontology it increases the complexity, making it general - less specific, which is not desired.

Secondly, we identified four quantifiable features (1) Veracity, (2) Complexity, (3) Practical Use, (4) Specificity and Reference. **Veracity** is a measure of correctness of the ontology with respect to the domain of discourse. **Syntactic Correctness** is a dimension to check if the ontology is syntactically correct. If the ontology contains syntactical errors then the parser supporting the associated knowledge representation language would point out errors. A related measure is richness of language or expressiveness of knowledge, **Semantic Expressibility**. Knowledge Representation techniques incorporate semantics in the representation model that renders sophisticated inferencing. Semantic Web lacks rich representation models, but new applications and requirements demand more expressibility and reasoning capabilities. Hence, many changes were made by W3C in the RDFS (RDF Schema) model, that resulted in a richer RDF vocabulary model, Web Ontology Language (OWL). The expressiveness of the language contributes more to the syntactic dimension of ontology against the correctness feature. Another issue in ontology engineering is sub-classification. A well-designed Taxonomy should

[5]http://www.dmoz.org/

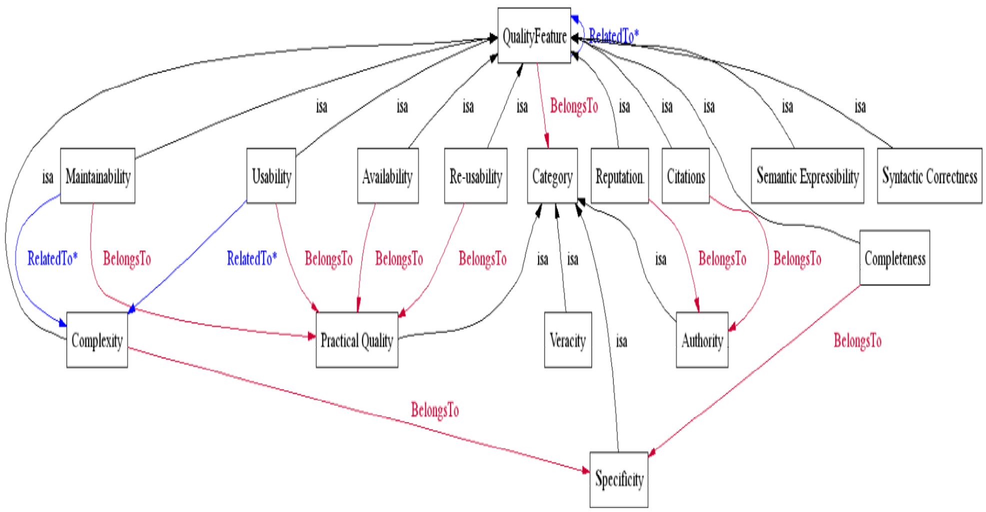

Figure 1: The Quality of Knowledge Ontology

contain all relevant domain concepts i.e. it should be **complete**, subclasses should be disjoint and exhaustive. A cognitive modelling approach for sub-classification that produces coherent balanced knowledge would yield high quality ontology.

The **Practical Quality** of ontology is how easy the ontology is to understand, which is essentially related to the complexity of conceptualization. A more complex model would be difficult to comprehend by user/agent, thus decreasing the usability of the ontology. **Modifiability**: In Semantic Web, where the Knowledge chunks are correlated and interdependent, it is critical when the changes made at some particular *knowledge node*. Current Semantic Web representation models (Martin 2002) do not have explicit recommendations with regards to such ad hoc modifications. We have formalized and grouped the constructs that characterize the notion of modifiability, which yield some additional information that is useful in process of modifying Ontologies. **Complexity** can be measured by the Depth of ontology , Width of the ontology (the number of children per concept), Total number of relations, Total number of attributes and total number of instances.

Availability measures the proportion of time the knowledge is available for usage. This parameter is related to the probability of a particular referenced Ontology being not available or not accessible over Web (caused by server down or other network problems at ontology publisher). **Reusability** is one of the other feature that needs to be ta ken into consideration as one of goals of ontology development. It renders Ontologies to be open ended, thereby accommodating the needs for future applications not envisioned at time of Ontology creation. Given the fact that the Ontologies are largely designed by domain experts, we can assert that an Ontology designed by extending other Ontologies (designed by some other domain expert) results in cumulatively richer and more expressive Ontology. Moreover reusing knowledge is also in harmony with spirit of Semantic Web propounding extensive reusability.

One important feature of an ontology is **Specificity** that determines how close it is to the universe of discourse. If we view Ontologies as taxonomy, specificity is governed by the width and depth. Width is average number of subclasses in a class while depth is ratio of total number of concepts to the width. For example, Ontology A and Ontology B represent a domain of interest that has 500 concepts. Average number of subclasses in Ontology A and Ontology B are 25 and 20 respectively. Ontology A [width = 25, depth = 20] is shallower than Ontology B [width = 20, depth = 25]. Intuitively a more specific ontology is indicative of high quality of knowledge.

Quantifying the Features

A knowledge-based system (KBS) generally uses only one model to store and exploit knowledge, e.g. a semantic network model such as Conceptual Graphs (CGs) or the Resource Description Format (RDF), but may import/export (i.e. accept/present) representations in various notations, e.g. KIF (Knowledge Interchange Format) or RDF/XML (the XML linearization of RDF), it turns out that some of notations are better than others for Knowledge Representation and exchange (Martin 2002). In this section we analyze the language specific features in RDF and the ontology languages DAML+OIL and OWL (Web Ontology Language) that is under final phases of standardization by W3C. We examined the detailed language specifications and grouped the constructs according to our proposed model for quality determination. Table 1 shows the categorization of the constructs as per the characterizing feature.

We quantify the characterizing features mentioned in Table 1 by keeping track of number of occurrences of a particular construct for a given Ontology. The OWL and related Semantic Web technologies are relatively new and only a handful of OWL based Ontologies are available currently on Web, hence our experimental Ontologies do not provide explicit characterization for each of the construct. Although, we do believe that since the Semantic Web standards are on verge of finalization, there will be lots of Ontology based applications over Web in near future.

Characterizing Feature	Semantic Web Language Specific Constructs
Expressibility	owl:transitiveProperty owl:functionalProperty owl:symmetricProperty owl:inverseOf
Reusability	rdfs:subClassOf rdfs:subPropertyOf
Citations	daml:sameClassAs rdfs:seeAlso owl:imports
Axioms and Logical Constraints	owl:equivalentClass owl:complementOf owl:disJointWith owl:AllDifferent owl:differentFrom
Maintainability, Modifiability	owl:versionInfo owl:backwardCompatibleWith owl:inCompatibleWith owl:deprecartedClass owl:deprecatedProperty owl:dataTypeProperty owl:objectProperty

Table 1: Categorizing Language Specific Characteristics

Table 2 gives a formal treatment for some of the quantifiable features as discussed earlier. We employed the depicted formulae to measure and quantify the various Ontological features. We were able to capture most of the obvious features, but some of the features like Completeness were difficult to capture, because the currently available Ontologies are largely designed in a centralized fashion rather than using a distributed approach where Ontologies that reference other ontologies over web (similar argument holds for the Availability feature).

Feature based Ontology Analysis

Quality model proposed in this paper provides the ontology of features that characterize knowledge on Semantic Web. As we previously described, our aim is to determine correctness of an ontology. Based on the Quality model, we determine the correctness of domain ontologies by ranking them according to their degree of correctness. The approach is straight forward one; we determine quality metrics for each of the domain ontology under consideration. Problem of ranking and characterizing Quality of domain ontologies is analogous to Multiple Attribute Decision Making (MADM) problem.

The goal of the paper is to use existing MADM solutions (Hwang & Yoon 1981) for selecting among ontologies associated with multiple, usually conflicting attributes. We formulate the problem of determining the most appropriate ontology for the given requirements from a set of domain ontologies as a Multiple Attribute Decision Making Problem. Let O = $\{O_1, O_2, O_3, \dots O_M\}$ be a set of ontologies (alternatives) for a given domain, Q = $\{Q_1, Q_2, \dots Q_N\}$ be a set of features that characterize Quality of knowledge. Decision Matrix $D = d_{ij}$ is a matrix representation of the scores. Score for each of n attributes for each of m alternatives is calculated on the basis of our proposed Quality model. The weight vector W reflects the importance of each attribute $W = \{W_1, W_2, W_3, \dots, W_N\},\ where \sum W_i$ = 1.

We use a simple additive weight based solution (Hwang & Yoon 1981) that determines best source ranking in all the alternatives. The method is comprised of three steps, scale the scores in range [0, 1] - best score 1, apply weights and sum up the values for each of the alternative. Scaling of scores is based on the following formula.

$$r_{ij} = (d_{ij} - d_{jmin})/(d_{jmax} - d_{jmin}) \tag{1}$$

The final preference score for each alternative is

$$S_i = \sum W_j r_{ij} \tag{2}$$

We also employ other MADM technique where in the DM just provides a relative preference matrix rather than precise weight matrix. In practical real world scenarios, it is difficult for a human Knowledge Engineer to come up with precise values for weight matrix. For example, one cannot strongly state value for *Specificity* to be 0.67 or 0.64. Human thinking tends to be fuzzy and relative, so one can confidently give relative values among parameters (e.g. *Specificity* = 2 * *Completeness*). Hence, we allow a Knowledge Engineer (Decision Maker) to feed in relative preference matrix (Table 6). The goal is to determine the Weight vector w for the given preference matrix, D. We use the solution as proposed in (Ma, Fan, & Huang 1998). Summarizing the problem mathematically,

$$Minimize\ \ z_1 = w^T F w = \sum_{k=1}^{n} \sum_{j=1}^{n} (d_{kj} w_j - w_k)^2 \tag{3}$$

subject to

$$e^T w = 1 \tag{4}$$

where $w = (w_1, w_2, \dots, w_n)^T, e = (1, 1, \dots, 1)^T, F = [f_{ij}]_{n \times n}$.

The elements in matrix F are

$$f_{ii} = n - 2 + \sum_{k=1}^{n} d_{ki}^2, i = 1, \dots, n \tag{5}$$

$$f_{ij} = -(d_{ij} + d_{ji}),\ where\, i, j = 1, \dots, n,\ \ i \neq j \tag{6}$$

Finally, solving the equations, the weight vector is calculated using

$$w^* = F^{-1} e / e^T F^{-1} e \tag{7}$$

Implementation and Experimental Analysis

We intend to determine correctness of knowledge for a given domain. For experimental purposes Tourism domain was selected as universe of discourse.

In an independent environment such as Semantic Web, where different agents, human or machines have varying conceptualization of a domain, we are more likely to find more than one Ontology for a domain of discourse. Currently Semantic Web doesnt have any respository or central index of Ontologies, hence, we implemented OntoKhoj, an Ontology search and classification tool. OntoKhoj is a Java based prototype system developed on Linux platform providing four major functionalities 1) Crawling Ontologies over the Semantic Web 2) crawled ontologies 3) Classifying each of the stored ontology 4) Ontology Visualization. We give a brief overview of implementation aspects of the aforementioned OntoKhoj functionality and its use in our experimental setup, detail description of the former can be found in (K. Supekar, C. Patel, & Y. Lee 2003).

- **Ontology Crawling** In OntoKhoj, an RDF crawler performs Ontology crawling on heterogenous Web sources including HTML, XML, RDF, DAML+OIL, OWL based resources. The ontology crawler provided additional features like aggregating

Feature	Formulae
Syntactic Warnings	SW = #Warnings Generated by Parser (Minor representation issues)
Syntactic Errors	SE = #Errors in the Representation Syntax (Violation of {Implicit Semantics (e.g. Cardinality), Explicit Semantics (e.g. loops)}
Completeness	CP = 1 - #Missing URI Links
Width	WD = Σ (#Child Nodes per Node)/ (Total# of Nodes)
Depth	DP = Total# of IS-A Links/WD
Specificity	$SP \propto DP \Longrightarrow SP = \Omega * DP$, where Ω is factor used to normalize Specificity
Usability	$USB \propto 1/SP \Longrightarrow USB = \Phi * 1/SP$, where Φ is usability normalization factor
Availability	Av = 1 - P(URI Unreachable) P(URI Unreachable) = Σ URI Unreachable / Σ Access Attempts

Table 2: Formal Quantification of characteristics

Number of Web pages visited	2018412
Number of Concepts crawled	19870
Number of Relationships Discovered	1321
Total Ontologies (after Aggregation)	418

Table 3: OntoKhoj statistics

crawled Ontologies that belong to same domain into a single RDF, distributed crawling and handling circular links of RDF URI's. We ran the OntoKhoj crawler for 48 hours (providing seed URLs as major Semantic Web websites and mailing lists). Table 3 shows some of the statistics obtained in our experiments.

- **Ontology Classification** We perform Ontology Classification to *fit* the Ontologies into a predefined directory of general categories. Our solution for Ontology classification is based on traditional classification algorithms and tool[6]. The classifier has been trained by the initial training data derived from plain categorized source [7] contains huge number of manually classified datasets. Each crawled and aggregated Ontology is handed over to the classifier which determines whether a new Ontology belongs to a particular topic with sufficient confidence. All the ontologies that were classified into Tourism category form our experimental dataset at http://ontobroker.semanticweb.org/ontos/compontos/.

We developed a Java based prototype that implements Simple Additive Weight (Hwang & Yoon 1981) to solve MADM problem. The system developed computes the decision matrix - scores of quality metrics associated with each Ontology. For experimental purposes we have chosen four quality attributes - Syntactic Correctness, Semantic Expressiveness, Specificity and Completeness. Table 4 shows related feature list for each of the eight ontologies. The scores in the decision matrix are then scaled based on the formalism (*Equation-1*), which is used to scale the scores between 0 and 1.

To have a better idea of the results obtained programmatically (table 5), we give a visualization of the RDF graph (using IsaViz [8] tool) for two Ontologies (O_1 and O_8) (Figure 2). The feature that can be viewed clearly is the Width (Average child per node), which is indicative of two important quality attributes - Specificity and Complexity. O_1 containing relatively fewer children per parent resulted in the graph to be narrower than the graph for Ontology O_8 that has higher value for Width parameter. From the visualization it is clear that ontology O_1 is more specific than O_8.

Let us assume W = {0.15, 0.15, 0.4, 0.3} be the weight vector provided by DM in our case the Knowledge Engineer.

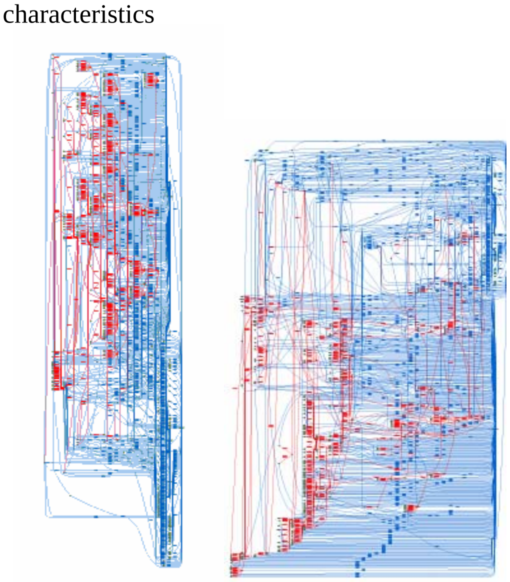

Figure 2: Tourism Ontologies, O_1 and O_8

Applying the Simple Additive Weight approach we calculate the score matrix (*Equation-2*) is used to calculate the overall scores for each alternative. Table 5 elicits overall scores for each of the alternative.

Table 5 provides the following order, $O_8 < O_7 < O_2 < O_4 < O_5 < O_6 < O_3 < O_1$, ontologies ordered according to degree of correctness. The results also show that O_1 has the maximum overall score (=0.89), thus identified as a *right* ontology for Tourism domain.

We now show the experimental results from the relative preference analysis. We incorporated the preference matrix, D, into above mentioned equations, and the final resultant Weight Vector w = {0.308, 0.245, 0.174, 0.271}. Correspondingly, we now see that the Ontology, O_5 (0.87) best suits the given preferences followed by O_1 (0.82), O_7(0.77) and so on. Table-6 provides overall scores for both the approaches, based on preferences of the knowledge engineer encoded in weight vector and preference matrix respectively.

Conclusion

In this paper, we have shown a rigorous knowledge quality model for supporting Software agents or Knowledge engineers to accurately judge the quality of Ontologies and determine reuse or extend them for their applications. We have

[6]shttp://www-2.cs.cmu.edu/ mccallum/bow/

[7]http://www.dmoz.org/

[8]http://www.w3.org/2001/11/IsaViz/

Feature	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8
Avg No of Child/Node	3.89	4.32	3.98	4.13	4.93	4.22	5.39	5.32
No of Concepts	343	311	310	302	306	304	302	304
Depth	88.1	72	77.9	72.1	62.1	72.2	56.3	56.1
No of Syntactic Warnings	4	14	0	95	1	5	1	76
No of rich DAML constructs	81	48	115	139	110	77	92	87
Syntactical Errors	No	No	No	No	No	No	No	No

Table 4: Ontology Feature Values

Score	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8
Overall	0.89	0.62	0.82	0.65	0.74	0.80	0.52	0.393
Relative Preferences	0.82	0.62	0.87	0.60	0.77	0.72	0.69	0.44

Table 5: Overall Scores

D = Preference Matrix	SC	SE	SP	CP
Syntactic Correctness (SC)	1	1	1	1
Semantic Expressiveness (SE)	0.5	1	2	1
Specificity (SP)	0.5	2	1	3
Completeness (CP)	2	1	1	1

Table 6: Preference Matrix

implemented the framework that searches, classifies and analyzes a set of domain Ontologies to quantify the features characterized by the proposed model. The experimental results showed the quality features of real world Ontologies on Semantic Web. Although determining the *right* ontology is still an open issue, we believe that the framework presented in this paper forms a suitable basis for selecting/reusing appropriate ontologies on Semantic Web.

References

Benjamin, R., and Fensel, D. 1998. The ontological engineering initiative-ka2. In *Proceedings of the 1st International Conference on FormalOntologies in Information Systems, FOIS'98*.

Berners-Lee, T. Hendler, J. and Lassila., O. 2001. The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*.

Chandrasekaran, B. Josephson, J. and Benjamins, R. 1998. The ontology of tasks and methods. In *In Proceedings of the 11th Knowledge Acquisition Modeling and Management Workshop, KAW'98*.

Foerster, V. 1969. Analysis and synthesis of cognitive processes and systems. Technical report, Biological Computer Laboratory, University of Illinois.

Gruber, T., and Olsen, G. 1994. . an ontology for engineering mathematics. In *In Proc. of the Fourth International Conference on Principles of Knowledge Representation and Reasoning*.

Gruber, T. 1993. *A Translation Approach to Portable Ontologies*. Knowledge Acquisition. chapter 5(2).

Guarino, N. 1998. Formal ontology and information systems. In *Proceedings of the 1st International Conference on FormalOntologies in Information Systems, FOIS'98*.

Hwang, C., and Yoon, K. 1981. *Multiple Attribute Decision Making*. Springer-Verlag.

Kalfoglou, Y., and Robertson, D. A. 1999. Case study in applying ontologies to augment and reason about the correctness of specifications. In *Proceedings of the 11th International Conference on Software Engineering and Knowledge Engineering, SEKE'99*.

K.Supekar C.Patel and Y.Lee. 2003. Ontorank: Ranking ontologies in semantic web. In *Workshop of Web Information and Data Management*.

Ma, J. Fan, Z. and Huang, L.-H. 1998. A subjective and objective integrated approach to determine attribute weights. In *European Journal of Operational Research*.

Maedhe, A., and Staab, S. 2001. Comparing ontologies-similarity measures and a comparison study. Technical Report 408, Institute AIFB, University of Karlsruhe.

Maganaraki, A. Alexaki, S. Christophides, V. and Plexousakis, D. 2002. Benchmarking rdf schemas for the semantic web. In *First International Semantic Web Conference (ISWC'02)*.

Martin, P. 2002. Knowledge representation in rdf/xml, kif, frame-cg and formalized-english, knowledge representation in cglf, cgif, kif, frame-cg and formalized-english. In *Tenth International Conference on Conceptual Structures ICCS 2002*.

Noy, N. F., and Hafner, C. 1997. The state of the art in ontology design: A survey and comparative review. *AI Magazine*.

Noy, N. F., and Musen, M. A. 2000. Prompt: Algorithm and tool for automated ontology merging and alignment. In *Seventeenth National Conference on Artificial Intelligence (AAAI2000)*.