Selection of Ontologies for the Semantic Web*

Adolfo Lozano-Tello[1], Asunción Gómez-Pérez[2], and Encarna Sosa[1]

[1] Escuela Politécnica de Cáceres. Universidad de Extremadura
Avda de la universidad, s/n. 10071 Cáceres. Spain
{alozano, esosa}@unex.es
[2] Facultad de Informática . Universidad Politécnica de Madrid
Campus de Montegancedo, s/n. 28660 Boadilla del Monte. Madrid. Spain
asun@fi.upm.es

Abstract. The development of the Semantic Web has encouraged the creation of ontologies in a great variety of domains. Web users currently looking for ontologies in order to incorporate them into their systems, just use their experience and intuition. This makes it difficult for them to justify their choices. Mainly, this is due to the lack of methods that help the user to measure that the most appropriate ontologies for the new system are. To solve this deficiency, this work proposes a method which allows the users to measure the suitability of the existent ontologies, regarding the requirements of their systems.

1 Ontologies for the Semantic Web

Nowadays, there is a growing interest to transform the huge amount of World Wide Web information in documents with "semantic markup" that can be processable by machines and humans. This new vision was created by Tim Berners-Lee [Ber01] to enable computers and people to work in better cooperation. Ontologies are the key pieces to provide semantic to the Web. In spite of the great increase that the use of ontologies has acquired (in Semantic Web, information search, knowledge management and electronic commerce [Fen01]), nowadays, the knowledge engineers need to look for ontologies disperse in quite a few of web servers. When they find several that can be adapted, they should examine their characteristics attentively and to decide which the best are to incorporate them to their system. Moreover, the ontologies are implemented in a great variety of languages [Cor00], applying several methodologies [Fer99], and using numerous technological platforms related with ontologies [Gol02]. Keeping in mind these dimensions, election procedures usually depend on the experience and the engineer's intuition. If the system is being developed with commercial goals, it will be very difficult for them to justify the taken election.

Although most of the methodologies for building ontologies [Fer99] propose a phase on reusing existent ontologies, there are not works that indicate the users how to choose ontologies for a new project, and there are not methodologies that quantify

* This work has been developed with the support of CICYT under contract TIC2002-04309-C02-01 and Junta de Extremadura under project 2PR01A023.

J.M. Cueva Lovelle et al. (Eds.): ICWE 2003, LNCS 2722, pp. 413–416, 2003.

the suitability of these ontologies for the system. This election problem would be palliated if it existed a metric that quantified, for each one of the ontologies candidates, how of appropriate they are for a new system. The method that is described in this work (called *ONTOMETRIC*) presents the set of processes that the user should carry out to obtain these measures.

2 *ONTOMETRIC*: A Method to Choose Ontologies

In this section we describe *ONTOMETRIC*, the adaptation of AHP for selection of ontologies. The Analytic Hierarchy Process (AHP) was devised by Thomas L. Saaty in the early seventies [Saa77]. It is a powerful and flexible tool for decision-making in complex multicriteria problems. This method allows people to gather knowledge about a particular problem, to quantify subjective opinions and to force the comparison of alternatives in relation to established criteria.

The *ONTOMETRIC* method is based on a taxonomy of 160 characteristics of ontologies [Loz03], called **multilevel framework of characteristics**, that provides the outline to choose and to compare existent ontologies. The multilevel framework of characteristics has, in the superior level of the taxonomy, five basic aspects on the ontologies that are denominated **dimensions**. These are: the **content** of the ontology and the organization of their contents, the **language** in which is implemented, the **methodology** that has been followed to develop it, the software **tools** used to build and edit the ontology, and the **costs** that the ontology will be necessary in a project. The multilevel framework of characteristics is the base to build an ontology in the ontology domain, call ***Reference Ontology*** (RO). The conceptual model of this ontology will gather the characteristics of the framework exposed. A direct relationship exists among all the descriptive characteristics identified in the dimensions, and the instance attributes specified in the RO. The development methodology of ontologies METHONTOLOGY [Fer97] and the development environment of ontologies (that follows METHONTOLOGY) WebODE [Cor02] were used to build the RO. A first version of the RO was used in $(ONTO)^2$Agent [Arp98] to search ontologies.

Taking into account the general steps of AHP, we have adapted the method to be used in the reuse of ontologies:

STEP 1: specify the project objectives. The knowledge engineer should know the exact guidelines of their company and available resources in relation to the new business.

STEP 2: build the decision tree from the MTC, so that the objective, "select the most appropriate ontology for a new project", is placed at the rood node; the dimensions (content, language, methodology, tool and costs) are placed at the first level; the factors of each dimension at the second level; and underneath these factors, the sub-trees of specific characteristics of the particular evaluation project. The general characteristics of ontologies should be specialised according to: the particular ontology, the specific target project and the organization that will develop the project.

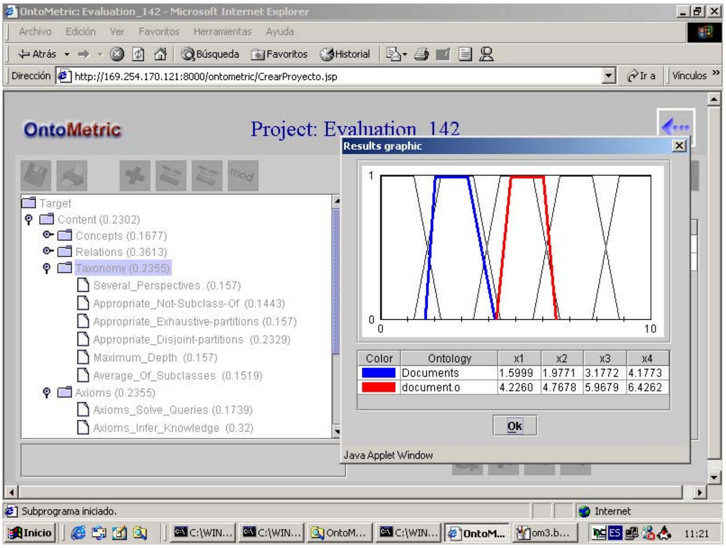

Fig. 1. Comparison of two ontologies (Documents and Document.o) in the factor "taxonomy" using *OntoMetric Tool*.

STEP 3: for each set of brother nodes, make the pairwise comparison matrixes [Saa77] with the criteria of the decision tree. The *eigenvectors* are calculated from these matrixes. These weights represent the relative importance between criteria.
STEP 4: for each alternative ontology, assess its characteristics. These values will (always multiplying by the weights calculated in step 3) ascend up to the superior nodes of the tree, until the node root is calculated. For each one of these characteristics, the user should establish a scale of appropriate ratings.
STEP 4.1: this method assigns linguistic values (non-numbers) to the alternatives because the human beings, in their daily activities, usually make this type of judgement. For example, if users evaluate the "essential relations for the system are defined in the ontology", they can assess this quality using the linguistic scale: (*very_low*, *low*, *medium*, *high* and *very_high*). It is more intuitive than a numeric scale between zero and ten. In this process, it is important that the groups of the linguistic values are precisely defined.

However it is not possible to perform calculations with linguistic values. One possible representation of these linguistic values is fuzzy intervals. Their angular points in a scale from 0 to 10, determine the fuzzy intervals (see graphic in figure 1).
STEP 4.2: with these established linguistic scales for each one of the criteria, the user will proceed to study each one of the ontologies that have been considered as alternatives, and to value them using these scales.
STEP 5: lastly, combine the vectors of weights *W* obtained in step 3 with the values of the alternatives *V* obtained in step 4, using (e.g.) the formula: $\Sigma_n w_i \; v_i$. Figure 1 shows an example with the comparison of two ontologies in the factor "taxonomy".

In large projects, which require a team of analysts, each person can provide their own values, and it will be necessary to reach an agreement. In this case, all the steps up to step 4.1 should reach a common consensus among the members of the evaluation team. Later, each analyst can value each one of the candidate ontologies in an individual way. Finally, the suitable ontology is chosen based on these results.

3 Conclusions

ONTOMETRIC is an adaptation of the AHP method to help knowledge engineers choose the appropriate ontology for a new project. This issue is being more important due to the enormous development of ontologies for the Semantic Web. In order to do this election, the engineer must compare the importance of the objectives, and study carefully the characteristics of ontologies. Although the specialisation of the characteristics and the assessment of the criteria of a particular ontology require a considerable effort, the above framework provides a useful schema to carry out complex multicriteria decision-making.

Feedback from knowledge engineers who have used the method, reveals that specifying the characteristics of a certain ontology is complicated and take time, and its assessment is quite subjective; however, they state that, once the framework has been defined, *ONTOMETRIC* helps to justify decisions taken, to "clarify ideas", and to weigh up the advantages and the risks involved in choosing one ontology from other options. The prototype ***OntoMetric Tool*** (figure 1) assists the user in applying the method. Shortly, this prototype will be integrated in WebODE platform [Cor02].

References

[Ar98] Arpírez J., Gómez-Pérez A., Lozano-Tello A., Pinto S.: (Onto)^{2}Agent: An Ontology-Based WWW Broker to Select Ontologies. European Conference of Artificial Intelligence (ECAI'98). Brighton. UK. (1998) 16–24.

[Ber01] Berners-Lee T., Hendler J., Lassila O.: The Semantic Web. Scientific American, Volume 284, N. 5. May. (2001) 34–43.

[Cor02] Corcho O, de Vicente O., Fernández-López M., Gómez-Pérez A.: WebODE: an Integrated Workbench for Ontology Representation, Reasoning and Exchange. Proceedings of EKAW'02. Sigüenza, Spain. (2002) 138–153.

[Cor00] Corcho O., Gómez-Pérez A.: A Roadmap to Ontology Specification Languages. Proceedings of EKAW'00. Jean-Les-Pins. (2000) 80–96.

[Fen01] Fensel D.: Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce. Springer-Verlag, Berlin (2001).

[Fer97] Fernández M., Gómez-Pérez A., Juristo N.: METHONTOLOGY: From Ontological Art Toward Ontological Engineering. Spring Symposium Series on Ontological Engineering. (AAAI'97). Stanford. (1997) 33–40.

[Fer99] Fernández M.: Overview of Methodologies for Building Ontologies. Proceedings of the IJCAI'99. Workshop on Ontologies and PSMs. Stockholm. (1999).

[Gol02] Golbeck J., Grove M., Parsia B., Kalyanpur A., Hendler J.: New Tools for the Semantic Web. Proceedings of EKAW'02. Sigüenza, Spain. (2002) 392–400.

[Loz03] Lozano-Tello A. and Gómez-Pérez A. ONTOMETRIC: A Method to Choose the Appropriate Ontology. Journal Of Database Management to be published in (2003).

[Saa77] Saaty, T.: A Scaling Method for Priorities in Hierarchical Structures. Journal of Mathematical Psychology, Vol.15, (1977) 234–281.