Ontology-Based Information Search in the Real World Using Web Services

Hyun-Suk Hwang[1], Kyoo-Seok Park[2], and Chang-Soo Kim[3,*,**]

[1] PuKyong National University, Institute of Engineering Research, Korea
hhs@mail1.pknu.ac.kr
[2] KyungNam University, Div. of Computer Engineering, Korea
kspark@kyungnam.ac.kr
[3] PuKyong National University, Dept. of Computer Science, Korea
cskim@pknu.ac.kr

Abstract. The ontology is an essential component to demonstrate the semantic Web, which is being described as the next Web generation. A semantic information search based on the ontology can provide the inferred and associated information between data. To develop an ontology application in the real world, we design the architecture of search systems based on the ontology using Web services. Our system consists of ontology modules, search procedure modules for searching, RDQL generator modules, and client modules for user interfaces. Also, we construct a hotel ontology integrated with the related terms and implement a search example with the defined ontology.

1 Introduction

The Semantic Web is a technology which adds well-defined documents on the Web for computers as well as people to understand the meaning of the documents more easily, and to automate the works such as information searches, interpretation, and integration.

The ontologies, which are an essential component of the semantic Web, define the common words and concepts used to describe and represent an area of knowledge [6]. The construction of ontologies in some domains such as travel [13], education [22], and medical data [12] has developed to integrate different data structure on the Web and to provide semantic information.

Most current Web sites have major limitations in finding search results and presenting them. The keyword-based search is not efficient because it often results in too many or too few hits. Also, the provided information has many redundant and unrelated results, so it takes a long time to find the information that users want. As a result, such searching is time consuming and often frustrates the Web search users [8].

Ontologies are linked to each other on the Web, and the linked ontologies provide the various applications with shared terminologies and understanding [12], [23]. Therefore, searching for information on the semantic Web will provide the search results with less redundancy, integrated terms, and inferred knowledge.

* Corresponding author.
** This research has been funded by Kyungnam University Masan, Korea.

M. Gavrilova et al. (Eds.): ICCSA 2006, LNCS 3982, pp. 125 – 133, 2006.

The most researches related to the Semantic Web have been focused on the standards of Web Ontology Language(OWL), the ontology constructions, its infrastructure based on the crawlers [4], [7], and agents [8], [9], [10], [11]. However, searches based on the ontology are actively not supported for users in the real world because of there being a shortage of standards of the ontologies between the same domains, content based on the ontologies, and connection information to the related sites. Therefore, in this paper we will present the search architecture based on the ontology using Web services. We will design the ontology of a realistic hotel search domain with a search scenario and implement the system. Also, we will present the potential benefits in searching based on the ontologies in the real world.

This paper is organized in the following manner. In the next section, we describe the semantic search examples with ontology and explain how Web services work. In section 3, we present the architecture for searching based on the ontology. We define a hotel search ontology and implement a search example in the defined ontology in section 4. Finally, we summarize this research and describe future work.

2 Related Research

2.1 The Semantic Search with Ontology

The ontology defines the common words and concepts used to describe and represent an area of knowledge [20]. Until recently, the ontology has been researched by the AI ontology community. This has begun to move for applying ontologies on the Web, especially in the area of search and retrieval of information repositories.

The Semantic Web [1] is a technology to add information on the Web, to enable computers as well as people to understand the meaning of the Web documents more easily, and to automate the works such as information search, interpretation, and integration. The Semantic Web is an extension of the current Web for the next Web generation started at the W3C in 1998, when it was working on the OWL, a standardized Ontology-Specification-Language for the semantic Web.

Especially, the semantic search is an application of the Semantic Web to search and is designed to improve traditional Web searching. The search method using the ontology is gathering strength as another new way of Web searching [10], [15], [21].

Passin [16] presented the advantages of the ontology compared to conventional databases in data structure points of view. The search based on the ontologies can also provide the relationships between resources and can exchange data with other applications.

Sugumaran and Storey [21] presented the efficiency and the approach method of semantic-based search compared to keyword search method. The method of keyword-based search is to find the occurrence of string patterns specified by the users in component attributes and descriptions. On the other hand, the semantic-based approach is to use a natural language interface for generating initial queries and to augment the searching with domain information. To support the semantic search, the ontology of the domain model is constructed to integrate different sets of terms.

Applications related to e-commerce, information retrieval, portals and Web communities based on the semantic Web and the ontologies have been actively researched in a few years. Especially, the noteworthy projects with respect to this

work are researches such as the OntoSeek [9] regarding the technical structure and environment in the USA, the OntoWeb [19] regarding the semantic portal in the E.U., the OntoBroker [5] regarding the general structure in the Germany, and the OntoKnowledge [20] regarding the frame work in the Germany.

In addition, there has been research to develop ontologies in applications on the Web. Clark et al. [3] insist on the importance of semantic Web in higher education. They said the effect of education on the Web depends on how the newly emerging semantic Web is explored, and the effects will be profound if the semantic Web becomes as ubiquitous as the Web today. Domingue et al. [8] developed an Alice, which is an ontology-based e-commerce project. This aims to support the dynamic query interface of online users by using five ontologies describing customers, products, typical shopping tasks, external context, and 'Alice' media. The combined ontology-based queries and dynamic queries will provide end users with the benefit of looking for relationships in large volumes of data.

The recent researchers [12], [15], [17] have used a Protege Tool to construct data structures and contents for supporting the semantic Web. The OWL [18] is widely accepted as the standard language for sharing semantic Web contents. The OWL plug-in [14], [15], [17] is a complex Protege plug-in with functions to load and save OWL files in various formats, to edit OWL ontologies with custom-tailored graphical widgets, and to provide access to reasoning based on description logic.

2.2 Semantic Web Services

The existing distribution systems have had disadvantages that they could not communicate with different protocol to one another. Web services can integrate the distributed computing environment using SOAP protocol with XML documents, not Resource Description Framework (RDF) documents. The Web services need the interface among Web service providers, brokers, and consumers. The providers publish the developed Web services with Universal Description Discovery and Integration (UDDI), and the consumers bind with Web Service Description Language (WSDL) and Simple Object Access Protocol (SOAP).

Passin et. al., [16] insist that the new version of SOAP makes it more practical to encode RDF data in a SOAP message if the current Web is oriented toward the semantic Web with semantic RDF contents and varied agents. Dameron et. al., [4] propose an architecture allowing the manipulation of ontologies using Web services. This enables users to implement such services like ontology Web services and their interfaces on the semantic Web. However, the functions rely on existing Web services technologies like SOAP and WSDL.

3 Architecture of an Ontology Based Search

3.1 System Architecture

Fig. 1 illustrates the architecture of semantic search using Web services. The systems consist of the content provider, semantic Web services, and search client.

The semantic Web services system includes the ontology server, Web services with remote search procedures, and Web server. The ontology server includes different

ontologies in varied domains and the ontologies can be exploited by different semantic Web applications. The remote search procedures support varied search functions for applications on client sides. The procedures need to connect the Jena API for querying the RDF contents and RDF Query Language (RDQL) Generator and RDQL Generator. The Web server needs to process the values of properties inputted by the client on the Web.

The content providers download the defined ontology from the ontology server and create RDF instances. The ontologies can be exploited by different semantic Web applications.

The search clients can attain required results by inputting the values for searching on the Web through Web server. Also, the client system can develop the applications by calling remote search procedures by Web services system on the mobile and personal system.

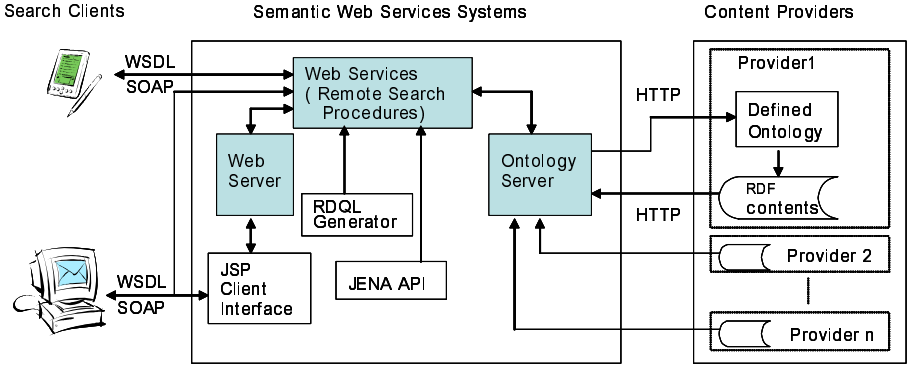

Fig. 1. Architecture of Ontology-Based Search

3.2 Implementation Module

Ontology Module. The ontologies could be defined by related industry. For example, Hotel Ontology could be provided by the hotel industry or hotel portal sites, and Geography Ontology could be defined by a government agency. The ontologies allow providers to get the defined ontology OWL files. The content providers submit the individual results as OWL or RDF files on their Web sites. The ontology server includes the interface form for providers to download the ontologies and to submit the URL with instances of the ontologies.

The ontologies can be constructed by using Protégé/OWL tool [14] which provides access to reasoning based on description logic and creates individuals with custom-tailored graphical widgets.

RDQL Generator Module. RDQL is one of the query languages for querying RDF contents. RDQL Generator generates the query string with the parameters of RDF models, properties needed by users for searching, and search conditions including subjects, properties, and objects which are elements of RDF statements.

Information Search Module. The ontology based information search can be divided into general keyword searching and ontology browsing searching. The general keyword method means the search based on data properties in total ontology structure, and the ontology browsing method means searches connected by object properties with relationships between classes. The browsing search allows users to search the information across the path connected between classes of the ontology.

The search procedures need the input information such as the URL of instances from providers, resulting properties from users, and RDQL query string. Also, the search procedures include data structures of properties of classes. Fig. 2 shows input values needed to attain search results.

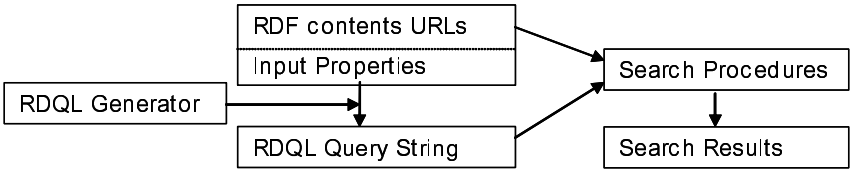

Fig. 2. Input Parameter of Search Procedures

Client Module. Clients can search for information based on the defined ontologies through their Web server and Web services. The clients need to submit input values for searches and then attain the search results. Also, clients can construct the applications on their client system by referencing WSDL files of Web services with defined ontologies. The applications can be generated on the mobile systems as well as personal systems.

4 Realizing a Hotel Search Ontology

We made up a search example based on the Hotel Ontology. We constructed the Web services with Java Web Services Development Pack 1.1 (JWSDP) including Tomcat server, and create the Hotel Ontology with the Protégé/OWL tool. We used the Jena API [2] to search requested information of users from the RDF-based contents generated by the Protégé/OWL. Also, we semantically searched RDF contents through RDQL, a Query Language for RDF. We used Java Server Page (JSP) to provide the user interface of information searches.

4.1 The Hotel Ontology

In this section, we show how we designed the ontology of a hotel domain from a search scenario which adjusts users' requirements. We imagined the search behavior of Web users who want to find a hotel with regards to some conditions like areas, room type, prices, and other facilities. The Web users try to find some candidate hotels of family suites with facilities like fitness centers and swimming pools, and they want to know them on for free. Therefore, we constructed the hotel ontology with more categorized classes including contract, service, room type, facility, and

rating. The ontology can provide more specific information by extracting information associated between data. For example, the users can obtain the information about what services are provided as the room type and whether the services are free or not. Fig. 3 shows the hierarchy of ontology for the hotel search and the relation connected by object properties between the classes in the defined ontology. The 'Hotel' Class is connected to the 'Room' class by the object property 'hasRoom', The 'Room' class is connected to the 'Service' class by the Property 'hasService'.

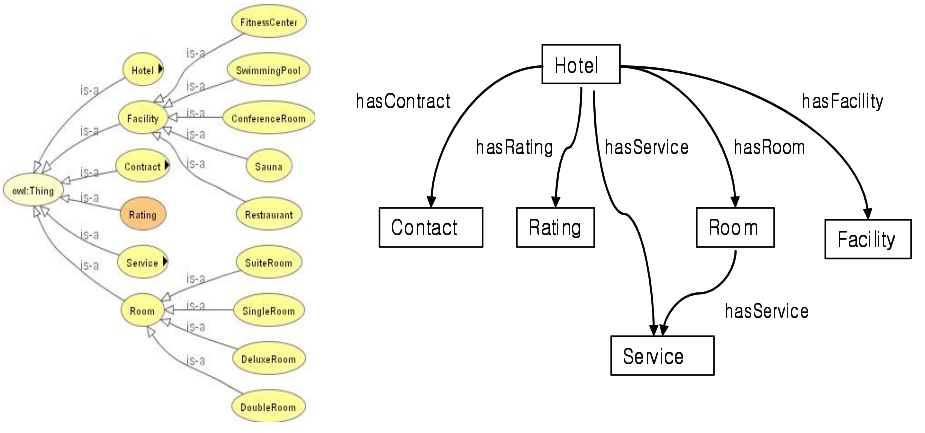

Fig. 3. The Fragment of the Hotel Ontology and the Associated Classes

4.2 Semantic Information Search

Our search systems support for clients to develop new search applications by accessing Web services. Fig. 4 shows the fragment of the WSDL needed to develop new applications and the part of calling search procedures of Web services on client systems.

Fig. 4. WSDL and Calling a Search Procedure

Fig. 5(a) shows an example of a keyword search by users' requirement in the defined Hotel Ontology. The search can provide the abstract information from classes associated by hasFacility and hasService with search keywords like address, rating, and price. Fig. 5(b) shows the search example by ontology browsing, and the search provides the hierarchy of the terms and relations between classes to help the decision in finding candidate hotels. We can search some lists of hotels with some services followed by a room type through searching between connected classes.

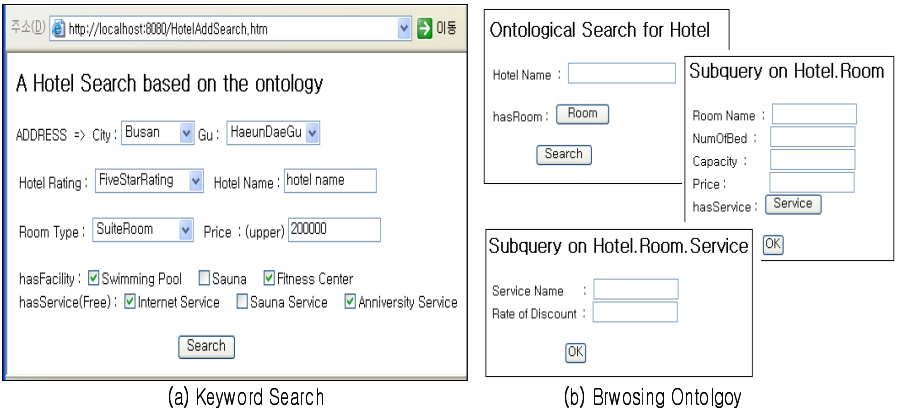

Fig. 5. Ontology-Based Search

4.3 Advantages of the Ontology-Based Search

The advantages of ontology-based search are follows. First, data integration on the Web can be accomplished by searching from ontology with integrated terms as domains, not extracting data from the different kind of database systems. Second, the ontology-based search provides more specific and hierarchical information by considering the relation between categorized classes, so finds the information from related data, not having been found from keyword search. In addition, it can do logic reasoning to discover unstated relationships in the data even though we do not implement the search with inference. Therefore, users can save the time and execute higher quality of Web search.

5 Conclusion

This paper presented the architecture of the information search based on the ontology using the Web services. The system consists of the ontology module, search procedures module, and client module. We constructed a standard ontology of hotel search domain with integration terms. Also, we implemented a hotel search example based on the defined ontology to improve the search of the current Web based on databases, which have problems such as the redundant and unrelated results and which are time consuming.

Our future works are as follows. The ontology of the hotel can be comprised in the variety viewpoint like travel theme categorized by mountain, beach, park, event and travel object like golf, business, and leisure with family and friends. Next, we can make up a portal site for the search based on the ontology to help the users to find some hotels adaptive to personal information.

Acknowledgement

This research was supported by the Program for the Training Graduate Students in Regional Innovation which was conducted by the Ministry of Commerce, Industry and Energy of the Korean Government.

References

1. Berners-Lee, T., Hendler, J., lassila, O.: The Semantic Web. Scientific American Vol. 284, No. 5 (2001) 34-43
2. Carroll, J. Jeremy, Dickinson, Lan, Dollin Chris: Jena: Implementing the Semantic Web Recommendations. Proceedings of the 13th International World Wide Web (2004)
3. Clark, K., Parsia, B., Hendler, J.: Will the Semantic Web Change Education. Journal of Interactive Media in Education (2004)
4. Dameron, O., Natalya F., Knublauch, H., Musen, A.M.: Accessing and Manipulating Ontologies Using Web Services. Third International Semantic Web Conference(ISWC2004), Hiroshima, Japan (2004)
5. Decker, S., Erdmann, M., Fensel D., Studer R.: Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information. Database Semantics: Semantic Issues in Multimedia Systems, Kluwer Academic Publisher (1999)
6. Devedzic, Vladan: Understanding ontological engineering. Communications of the ACM, Vol. 45, No. 4 (2002) 136-144
7. Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R. S., Sachs, J., Doshi, V., Reddivari, P., Peng, Y.: Swoogle: A Search and Metadata Engine for the Semantic Web. Thirteenth ACM Conference on Information and Knowledge Management (CIKM'04), Washington DC, USA (2004)
8. Domingue, J., Stutt, A., Martins, M.: Supporting Online Shopping through a Combination of Ontologies and Interface Metaphors. International Journal of Human-Computer Studies 59 (2003)
9. Guarino, Nicola, Masolo, Claudio, Vetere, Guido: OntoSeek: Content-Based Access to the Web. IEEE Intelligent Systems, Vol. 14, No. 3 (1999) 70-80
10. Guha, R., McCool, Rob, Miller, Eric: Semantic Search. Proceedings of the 12th International Conference on World Wide Web (2003)
11. Knublauch, Holger, Musen, Mark A, Rector, Alan, L.: Editing Description Logic Ontologies with the Protégé OWL Plugin. International Workshop on Description Logics – DL (2004)
12. Knublauch, Holger, Dameron, Olivier, Musen, Mark A.: Weaving the Biomedical Semantic Web with the Protégé OWL Plugin. First International Workshop on Formal Biomedical Knowledge Representation, Whistler, BC, Canada (2004)

13. Knublauch, Holger: Ontology-Driven Software Development in the Context of the Semantic Web: An Example Scenario with Protege/OWL. International Workshop on the Model-Driven Semantic Web, Monterey, CA (2004)
14. Knublauch, Holger, Protege OWL Plugin tutorial. 7th International Protege Conference, Bethesda, MD, http://protege.stanford.edu/plugins/owl/documentation. html (2004)
15. Knublauch, Holger, Fergerson, Ray W., Noy, Natalya F., Musen, Mark A.: The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. Third International Semantic Web Conference - ISWC (2004)
16. Passin, Thomas B., Explorer's Guide to the Semantic Web. Manning (2004)
17. Rector, A., Drummond, N., Horridge, M., Rogers, J., Knublauch, H., Stevens, R., Wang, H., Wroe, C.: OWL Pizzas: Practical Experience of Teaching OWL-DL: Common Errors & Common Patterns. 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW) (2004)
18. Smith, Michael K., Welth Chris, McGuinness, Deborah L.: OWL Web Ontology Language Guide. Http://www.w3.org/TR/owl-guide/ (2003)
19. Spyns, Peter, Oberle, D., Volz, R., Zheng, J., Jarrar, M., Sure, Y., Studer R., Meersman, R.: OntoWeb: a Semantic Web Community Portal. In Proc. Fourth International Conference on Practical Aspects of Knowledge Management (PAKM), Vienna, Austria (2002)
20. Staab, Steffen, Studer, R., Schnurr, Hans-Peter, Sure, Y.: Knowledge Management : Knowledge Processes and Ontologies. IEEE Intelligent Systems Journal (2001)
21. Sugumaran, Vijayan, Story, V. C.: A Semantic-Based Approach to Component Retrieval. The DATABASE for Advances in Information Systems, Vol. 34, No. 3 (2003)
22. Tane, J., Schmitz, C., Stumme, G.: Semantic Resource Management for the Web: An E-Learning Application. Proceedings of the 13th international World Wide Web (2004)
23. Uschold, M. F., Jasper, R. j.: A Framework for Understanding and Classifying Ontology Applications. Proceedings of the IJCAI-99 workshops on Ontologies and Problem-Solving Mehtod (KRR5), Stockholm, Sweden (1999)