GEOSPATIAL SEMANTIC QUERY BASED ON CASE-BASED REASONING SYSTEM

Kay Khaing Win
University of Computer Studies, Yangon,Myanmar

Keywords: Geospatial data, geospatial semantic, semantic query, spatial reasoning, geospatial semantic web.

Abstract: In today's fast-growing information age, currently available methods for finding and using information on the Web are often insufficient. Today's retrieval methods are typically limited to keywords searches or sub-string matches, therefore, users may often miss critical information when searching the web. After reviewing the real world Semantic Web, additional research is needed on the Geospatial Semantic Web. We are rich in geospatial data but poor in up-to-date geospatial information and knowledge that are ready to be used by anyone who wants to use. In this paper, we implement a framework of geospatial semantic query based on case based reasoning system that contributes to the development of geospatial semantic web. It is important to establish a geospatial semantics that support for effective spatial reasoning for performing geospatial semantic query. Compared to earlier keyword-based and information retrieval techniques that rely on syntax, we use semantic approaches in our spatial queries.

1 INTRODUCTION

Today's retrieval methods are offering no support for deeper structures that might lie hidden in the data: therefore, users may often miss critical information when searching the Web. At the same time, the structure of the posted data is flat, which increases the difficulty of interpreting the data consistently. There would exit a much higher potential for exploiting the Web if tools were available that better match human reasoning. In this vein, the research community has begun an effort to investigate foundations for the next stage of the Web, called Semantic Web.

A rich domain that requires special attention is the Geospatial Semantic Web. In the future, the Geospatial Semantic Web will allow the returning of both spatial and non-spatial resources to simple queries, using a browser. However, in the same way as with the Semantic Web, in order to approach the Geospatial Semantic Web it is necessary to solve several problems.

In this paper, we implement a framework of geospatial semantic query based on case-based reasoning system that contributes to the development of geospatial semantic web. We intend to develop a simple and powerful framework for people to interpret the semantics of geospatial entity classes. The remainder of this paper is organized as follows. Section 2 describes about the geospatial semantic web and geospatial semantic. Section 3 describes the role of semantic knowledge in case searches. We describe measuring semantic similarity in section 4. Section 5 describes case-based reasoning system. We describe conceptual framework of geospatial semantic query in section 6. Finally, we describe conclusion in section 7.

2 GEOSPATIAL SEMANTIC WEB AND GEOSPATIAL SEMANTIC

Geospatial Semantic Web is a natural extension of the current geospatial systems and applications that enable users to query more precisely the data they need. To accomplish the Geospatial Semantic Web, two research issues are apparent: 1^{st} is geospatial data query and 2^{nd} is method to assess the semantics of available data sources. The retrieval methods of semantic web are developed by incorporating the data's semantics and search process. Such a development needs the development of multiple spatial and terminological ontologies. The needs of all the above mentions enforce to the development of Geospatial Semantic Web.

The Geospatial Semantic Web will be a significant advancement in the meaningful use of

Khaing Win K. (2006).
GEOSPATIAL SEMANTIC QUERY BASED ON CASE-BASED REASONING SYSTEM.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - AIDSS*, pages 356-359
DOI: 10.5220/0002469203560359

spatial information. With the flexible incorporation of geospatial information retrieval; will become precise to the level that the results of user queries will be immediately useful, without weeding out irrelevant hits. In order to address geospatial semantic, one needs computational methods that go beyond syntax comparisons. In the case of the Geospatial Semantic Web, three types of geospatial semantics are distinguished, each requiring different computational methods (Egenhafer 2002):

- semantics of geospatial entity classes
- semantics of spatial predicates
- semantics of geospatial names

The classification of geographic entities is geospatial, even when no geometry is involved. Non-geometric concepts, such as building, road, and place are geospatial concepts that are used for describing the semantics of geospatial objects. Semantic relations are a typical way to describe knowledge about concepts. We refer to geospatial entity classes by words or sets of synonym interrelated by hyponymy and metonymy relations.

3 ROLE OF SEMANTIC KNOWLEDGE IN CASE SEARCHES

In order to understand and appreciate, the role and importance of semantic knowledge and sentence structure in case-retrieval process, we need to understand the working of systems that do not use this information and rely only on the word knowledge. Based on the inputs given by the user, a relevant case is retrieved and output to the user. A set of questions are posed to the user, the answers to which are compared to the ones listed out under the relevant cases that are retrieved. A question answer pair typically behaves as an attribute value pair. The answers provided by the user to the questions posed are compared to these values and a match is found.

In our method, similar concepts are playing similar roles in the sentences. A sentence is represented using an Interlingua called, Universal Networking Language (UNL). Information in every sentence is captured at three levels: the concepts that are involved, the role they play in the sentence, and attributes that describe their properties. Universal Networking Language (UNL), proposed by United Nations University, represents natural language in the form the links among them see in figure 1. UNL represents information sentence by sentence. This hypergraph is also represented as a set of directed binary relations, each between two concepts present in the sentence. Concepts are represented as character strings called Universal Words (UW).

The knowledge within a document is represented in three dimensions: Universal Words (UW): describe concepts that are present in a document; UNL Relations: describe the relations between the concepts involved in the sentence and the roles (e.g. subject or object in case of nouns) that they play in conveying the meaning of a sentence; UW attributes: capture and represent properties of concepts like tense of a verb.

4 MEASURING SENTENCE SIMILARITY

Similarity between two sentences is measured on two counts: how similar are the concepts involved in the two sentences and how similar roles do the concepts play in the sentence? Since relations describe the roles that concepts play in the meaning of the sentence, similar structure sentences will have similar relations in their respective UNL representations. Taking UNL representation, we compare each of the concepts occurring in UNL representation with the concepts that appear in the UNL representation for the problem sentence. The similarity score is computed using the method proposed by (Resnik, 1999) where similarity of two concepts is determined by the information that they share indicated by the most specific concept that subsumes them both in a concept hierarchy. Resnik used WordNet for this hierarchy of concepts. For every concept, its likelihood of occurring in the document is calculated by counting the number of instances of itself and the concepts subsumed by it in the document. Therefore, the more general a concept, the more number of occurrences it will have. Probability (or likelihood) of occurrence of a concept is given as

$$P(c) = Nc/N$$

where Nc is the number of times a concept C occurs in the document and N is the total number of words in the document. Using the Information Content Theory, the Information Values associated with each concept C is negative log of the likelihood of occurrence of the concept.

$$IC(c) = -\ln (P(c))$$

We too used WordNet to arrange concepts in a hierarchy and assign them Information Content Values in the manner proposed by (Resnik, 1999). However, in Resnik's method, the sense of the concepts being matched is not known.

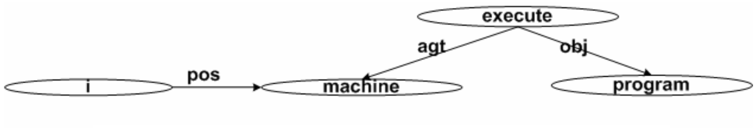

Figure 1: UNL graph for “My machine is executing programs”.

Therefore, a similarity score is measured for all senses of the two concepts and the maximum among them is chosen. We uses K-Nearest Neighbor (kNN) algorithm to determine the similarity between cases. Use of UNL Universal Words helps us restrict our attention to only one sense of a concept and therefore produces the most useful similarity score. If there are N1 and N2 nodes (or words) in the two sentences S1 and S2 respectively, then the concept similarity measure is calculated as

$$(\sum_{n1 \in s1} \sum_{n2 \in s2} \text{SimScore}(n_1, n_2)) / (N_1 * N_2)$$

The sum of all the similarity scores over all pairs of concepts that are matched for two sentences is taken and averaged over the number of comparisons made. We use this concept similarity in our geospatial semantic query system. In our proposed system (CBR system), the case base reasoner use this concept similarity measure to retrieve the most similar case (or cases) comparing the case to the library of past cases.

5 CASE BASED REASONING SYSTEMS

CBR is a view of knowledge acquisition method for problem-solving and interpretation, and a method for machine learning. CBR is to solve a problem by remembering a previous similar situation and by reusing information and knowledge of that situation. In Case-Based Reasoning (CBR) systems expertise is embodied in a library of past cases, rather than being encoded in classical rules. Each case typically contains a description of the problem, plus a solution and/or the outcome.

We implement a framework of geospatial semantic query based on case based reasoning system that contributes to the development of geospatial semantic web. In the framework of geospatial semantic query, spatial relations are stored in case based library of case based reasoning system. Spatial relationships between objects provide details of the retrieve location of objects e.g., “the train is to the right of the platform.” Such relationships are especially needed to locate objects in which case the location of such objects may be approximate during the spatial relationships between the objects it is near. It is impossible to determine the spatial relationship between x and y when x completely obscures y, as would be the case if x was inside of y. Since this information is not always determineable from object co-ordinates, without this aspect, semantic query cannot be supported. Spatial relationships may be represented explicitly via relations. We intend this reasoning system to be geospatial reasoning using the case based reasoning methodology. Geospatial reasoning is widely used by humans to understand, analyze, and draw conclusions about the spatial environment.

6 CONCEPTUAL FRAMEWORK OF GEOSPATIAL SEMANTIC QUERY

To illustrate how the principles of the conceptual framework of geospatial semantic query based on case-based reasoning system. Our proposed framework shown in figure 2 is intended to implement on the geospatial semantic query system that contributes to the development of geospatial semantic web. The sole purpose of this task is to make the system understand the terms appearing in the user’s query input. The reasoning capability of models of spatial relations is critical to complete the task of geospatial semantic query. To keep pace with future developments in geospatial reasoning, we argue that the system should be designed with an open architecture to allow for new models and extensions of existing models to be incorporated into the system easily.

Our proposed framework mainly consists of case based reasoning system. Spatial relations are stored in case based library of case based reasoning system. Case based reasoner matches the current problem on the query content with the cases in the case based library, and similar cases are retrieved. We examine the use of pairs of terms found in close proximity to each other to be used as the query terms. Additionally, we are conducting experiments using a more refined problem-specific sense of relevance: a case is considered relevant only if it is actually cited in the actual opinion of the problem case. The retrieved cases are used to suggest a solution which is reused and tested for success.

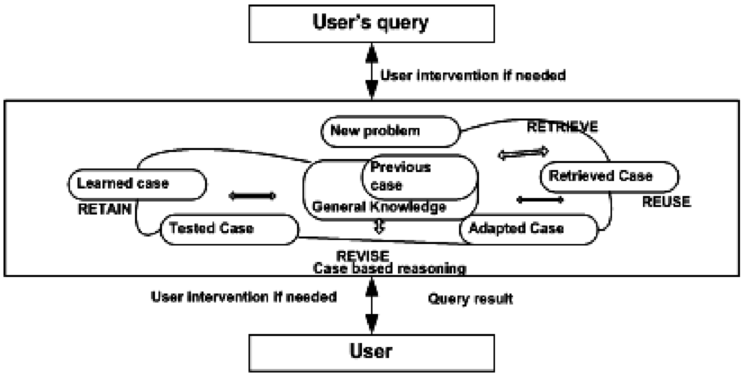

Figure 2: Conceptual framework of geospatial semantic query.

If necessary, the solution is then revised. Finally the current problem and the final solution are retained as part of a new case. The case based reasoner produces the query result (final solution or outcome) to the user.

7 CONCLUSIONS

This paper introduces the new concept of geospatial relations on case-based reasoning and process a solution to query of geospatial semantic. To handle geospatial semantic query, we propose a conceptual framework that takes advantage of case-based reasoning system. The emphasis of this paper has been reasoning on geospatial relations to handle geospatial semantic queries. In our system, we presented an approach for querying geospatial semantic based on case-based reasoning system that contributes to the development of Geospatial Semantic Web. In our system, we use geospatial relations to be efficient and effective reasoning system. This paper describes efforts in developing for conceptual framework of a geospatial semantic query system.

REFERENCES

Allen J. F., 1983 "Maintaing knowledge about temporal interval communications" in *ACM, 26 (11)*: 832-843.

Andrea M. R. and Egenhofer M.J, 2004 "Comparing Geospatial Entity Classes: An Asymmetric and context-Dependent Similarity Measure*" International Journal of Geographical Information Science* 18 (3).

Bhattacharyya P.,Choudhury S.R.,Gupta S.S., 1998 "Man power Planning with Case Based Reasoning: The Selector", *International Conference on Knowledge-based Computer Systems (KBCS).*

Chakraborti S.,Balaraman V.,Vattam S.S., 2000 "Using CBR inexact search for intelligent data retrieval", *International Conference on Knowledge-based Computer Systems (KBCS).*

Egenhofer, M. 2002, "Toward the Semantic Geospatial Web*" ACM-GIS Mclean, VI* (in press)

Jones C.B and Purves R. and Ruas A. 2001 "Spatial Information Retrieval and Geospatial Information Theory", *pages 322-335. Springer-verlag.*

Jones C.B and Purves R. and Ruas A. and Sanderson M. and Sester M and Van Kreveld M.J. and Weibel R. 2002 "Spatial Information Retrieval and Geographical Ontologies: an overview of the SPIRIT Project". *In Proceeding of the 25th Annual International ACM SIGIR Conference, Pages 387_388*

Jun-ichi A., Koioru H., Akira S., Yutaka Y. and Testsuji S. 2002 "Towards the Real World Semantic Web-Web Search based on Spatial and Temporal Metadata" *vol.1.No.3 June 2002.*

Mcauley K. 2001 "Geospatial Mapping and Navigation of the Web." *In proceedings of Tenth International World Wide Web Conference, page Session P7.ACM Press.*

Mclean, Virginia, 2002. "Toward the semantic geospatial web" *Proceedings of the 10th ACM international symposium on Advances in geographic information systems.*

Ohlbach, H. J and Lorenz B. 2005 "Geospatial Reasoning: Basic Concepts and Theory" *31 March 2005 REWERSE.*

Resnik P., 1999 "Semantic similarity in a taxonomy: An infomation-based measure and its application to problems of ambiguity in natural language", *Journal of Artificial Intelligence Research.*

Sistal P., Yu C. ,and Haddack R. 1994 "Reasoning about spatial relationships in picture retrieval systems*". In Proc. Of the 20th Int't Conf. on very large Data Bases.*

Sharma J. and Flewelling D. M. 1995 "Inferences from combined knowledge about topology and directions." *In Proc. Of 4th Int't Symposium on large Spatial Databases, Portland, Marine, USA*