Global Journal of Computational Intelligence Research.
ISSN 2249-0000 Volume 2, Number 1 (2012), pp. 39-43

http://www.ripublication.com/gjcir.htm

Ontology Based Ranking Approach for Web Documents

N. Praveena and S.Nagaraju

Dept.Of CSE
KITS,Warangal
e-mail:nagishettipraveena@gmail.com

Associate Professor
Dept.Of CSE
e-mail: nag_sangepu@yahoo.com,

Abstract

With the tremendous growth of information which is available through the web for end users, computing ontology comes to play an important role. The Governments now a days can be more effective with the use of information and communication technologies. The main challenge for the government is to interact and utilize internet effectively. Several techniques have proposed for processing of data, which allows efficient information retrieval but, do not take concept of relationship into consideration. In this proposed system, we develop ontology with respective concepts belonging to a particular domain. This enables the users to understand information more effectively. The proposed system also provides ranking based approach in conjunction with semantic ontology's which relies on information that can be extracted from web or annotated resources.

Keywords— Ontology, Semantic Web,

I. Introduction

In the E-government (E-gov) context, governments use the most innovative information and communication technologies to improve the quality of the services and to provide greater opportunities to participate in democratic institutions and processes. Delivering efficient and cost effective services are big challenges for many governments worldwide. The extensive use of the Internet required new software engineering methodologies. Simpler and faster methodologies have been introduced to organizations and governments in order to satisfy their demand. Examples of these methodologies are: rapid prototyping, agile development, extreme programming, and others. Software Engineering faces many key challenges such as heterogeneity, delivery, trust, cost, timelines, and quality challenges. Requirements engineering, an

essential phase in the software life cycle, has been a challenging problem. Software developers need to be able to capture the semantics of the required intelligent processing based on the application domain. Requirements engineering is concerned with domain specifications of software behavior and with their evolution over time. Getting high quality and well-defined domain for software requirements process for E-gov application is difficult and critical.

II. Previous Work

Domain Engineering is a process for creating reusable software abstractions for the development of a family of software applications in a domain. The process consists of the phases of analysis, design and implementation. Domain analysis identifies reuse opportunities and specifies the common requirements of a family of applications. The product of this phase is a domain model. Domain design looks for a documented solution to the problem specified in a domain model. The product of this phase is composed of one or more frameworks and a collection of design patterns, documenting good solutions in that domain. Reusable components integrating the framework are constructed during the phase of domain implementation.

Our research group is working on a process for Multi-Agent Domain Engineering. We are constructing a software development environment composed of a set of development tools and libraries of high-level reusable software abstractions (domain models, domain specific languages, user models, agent-based architectural and detailed design patterns, multi-agent frameworks, and reusable software agents) for both compositional and generative agent-based application development. Experiments are being conducted on the legal, tourism, and pecuary domains and for problem-solving tasks of information access and user modeling.

Ontologies are knowledge representation structures especially useful for the specification of high-level reusable software, like domain, user models and frameworks. They provide an unambiguous terminology that can be shared by all involved in the development process. Also, ontology can be as generic as needed allowing its reuse and easy extension. Ontologies are being used to represent both the knowledge of techniques for Domain Engineering and generated products.

III. Proposed System

Data Retrieval

In this module we crawl to retrieve multiple documents from the data set path specified. This module will run as a multithreaded program continuously, which will take the documents and the page, it will extract text from the document pages. The downloaded pages are saved in the temp directory located in the system. In this module we also strip any of the HTML tags that the document may contain, thus the save files will in plain text format. This module will retrieve the downloaded pages and for each page it will remove HTML tags as well as any special characters. The pages will now contain only the data. The data is saved in the Knowledge database.

Elimination of Stopwords
In this module we retrieve the plain text documents already saved, we eliminated the stopping words extremely common words like use, can, the, of, etc which are part of preposition and adverbs from the documents. The resultant documents will provide the important words which are considered as concept words.

Concept and Ontology Computation
In this module we retrieve the concepts from each document and store in a temporary data structure. The concepts from each document are sorted in the dictionary order such that the same concept words occur in the subsequent lines in the documents. The number of occurrence of the concepts is computed by counting the same and similar words in the subsequent lines of the documents. The frequencies of each concept words are stored in the data structure. The final stage of the module computes the occurrence of concept words from the entire set of documents and the results are displayed.

Ontology Ranking Logic
This module will retrieve the documents from the dataset and will check for the keyword that is present in the documents. It will only retrieve the pages that matched the keyword. Then program will consider the retrieved pages and construct the ontology sub graph. This module will consider all the documents that are retrieve as well as the sub graphs and will compute page spanning forests. Using this it will compute the scores based on the relations. And merge it with the original documents saved in the dataset. The ranking is calculated based on the scores of the concept words. The result is out to the user.

IV. Results

The concept of this paper is implemented and different results are shown below, The proposed paper is implemented in Java technology on a Pentium-IV PC with 20 GB hard-disk and 256 MB RAM with apache web server. The propose paper's concepts shows efficient results and has been efficiently tested on different Datasets. The Fig 1, Fig 2, and Fig 3 shows the real time results compared.

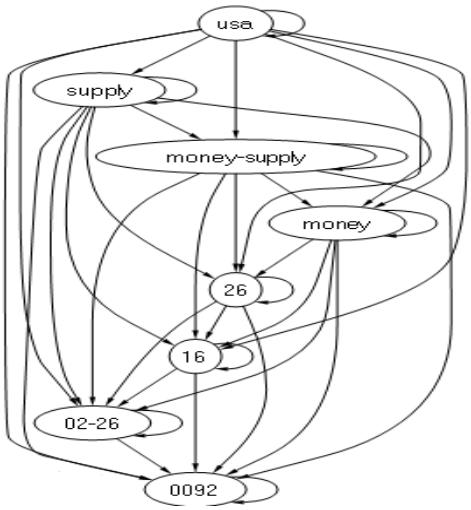

Fig. 1. Proposed system Ontology Graph.

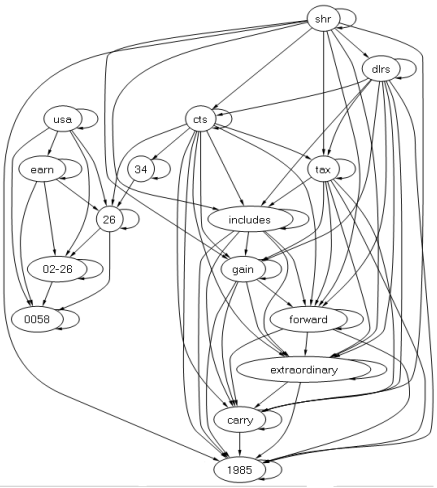

Fig. 1. Proposed system Ontology Graph

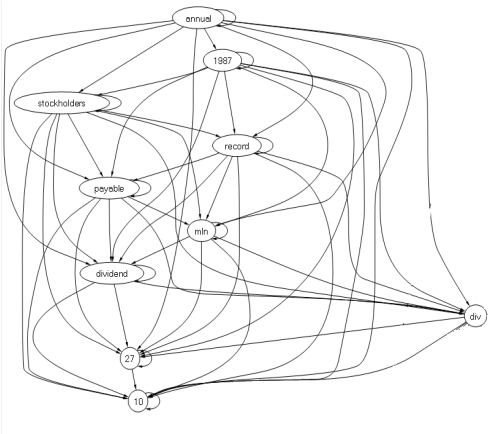

Fig. 3. Proposed system Ontology Graph

V. Conclusions

This article introduced a technique for the construction of domain models to be reused in the development of multi-agent applications. This system is an ontology-based domain model representing the knowledge of techniques for the specification of the requirements of a family of multi-agent systems in an application domain.Some of the advantages of using ontologies for the representation of reusable products have been shown in this article. Although it has been designed for its integration in a software development environment for Multi-agent Domain Engineering, the approach can be generalized to other development paradigms. For that, the ontology should be re-designed according to the particular knowledge of the techniques of those development paradigms.

References

[1] Samhan A., “Master Thesis: Towards An Ontology for Software Product Quality Attributes”, Middle East University for Graduate Studies, Amman Jordan, May, 2008.

[2] Liao L., Qu Y., and Leung H., “A software process ontology and its application”. In ISWC2005 Workshop on Semantic Web Enabled Software Engineering, 2005.

[3] Scacchi W., Jensen C., Noll J., and Elliott M.,"Multi-Modal Modeling, Analysis and Validation of Open Source Software Requirements Processes", Intern. J. Information Technology and Web Engineering, Vol. 1, Issue 3, 2006.

[4] Zhao G., Gao Y., and Meersman R., ”An Ontology-based Approach to Business Modeling”, In Proceedings of the International Conference of Knowledge Engineering and Decision Support, 2004.

[5] Girardi R. and Faria C., ”An ontology-based technique for the specification of domain and user models in multi-agent domain”. CLEI Electron. J. 7(1), 2004.

[6] Floridi L. (ed.), "Blackwell Guide to the Philosophy of Computing and Information", Preprint version of chapter “Ontology”, Oxford: Blackwell, 155–166, 2003.

[7] Gruber T., "Ontology, The Encyclopedia of Database Systems", Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2008.