

Chapter 78

Multidimensional Ontology– Based Information Retrieval for Academic Counseling

S. S. Lam

The Open University of Hong Kong, Hong Kong

Samuel P. M. Choi

The Open University of Hong Kong, Hong Kong

ABSTRACT

Conventional information retrieval can only locate documents containing user specified keywords. Integrating domain ontology with information retrieval extends the keyword-based search to semantic search and thus potentially improves the precision and recall of the document retrieval. In this paper, a set of new multidimensional ontology-based information retrieval algorithms is proposed for searching both specific and related terms. In particular, the relevant data properties of an instance, the relevant concepts, the relevant related concepts, and the related instances of a given user query can be identified from the domain ontology via the multidimensional search. Using the proposed algorithms, an intelligent counselling system which provides 24x7 online academic counselling services is developed. Through an interactive user-interface and domain ontology, the system facilitates students to find desired information by reviewing and refining their query. The article also outlines how to enable ontology-based searching for a conventional website.

INTRODUCTION

Higher education is moving towards flexible learning in order to meet the needs of more diverse student body. The Higher Education Funding Council of England has piloted the funding of flexible education since 2005. Being an important player in the field of open and distance learning, the Open University of Hong Kong (OUHK) offers highly flexible credit system supported by multiple learning modes including distance learning, face-to-face teaching and online learning that allows students studying on their pace. However, flexible learning creates new challenges to students on choosing the programmes or courses

DOI: 10.4018/978-1-5225-5191-1.ch078

that are suitable to their career development. Although the university has published the related information on the website with the support of Google search engine, prospective and current students still feel frustrated when searching for relevant information scattered across different webpages. Information presented on the website is written in webpages aimed at human consumption. They do not, however, contain metadata to support efficient searching and ranking of results to match students' query. Students have to review searching results page-by-page to find the right answer to their queries. Furthermore, conventional keyword-based search engines cannot interpret the sense of user's search and the ambiguity of the query leads to the retrieval of irrelevant information. New approach is needed for organizing the information and providing specific and relevant search results.

LITERATURE REVIEW AND RELATED WORK

Recent development of semantic Web has fuelled researches on developing semantic search engine and methods to exploit the potential of semantic Web, particularly the ontology, in information retrieval. Vallet et al. (2005) combine semantic search with keyword-based search. Ontology-based scheme is used to annotate documents. Weighted average is used to combine the semantic similarity measure with the similarity measure derived from the classic vector-space model. Dong et al. (2008) presents a preliminary survey on semantic search technologies and discusses common issues in the current semantic search engines and methods. Semantic search approaches are classified into various categories including semantic search engines and semantic search methods that completely adopt the semantic Web technology, hybrid semantic search engines that integrate semantic web technology into key-word-based search engines to improve the precision of traditional text search, search engines that can query objects in XML documents, and search engines that are designed for querying ontological files. Major common issues of semantic search methodologies include differentiation between designers and users' perceptions of relevancy of context, lack of adaptability in the knowledge structure, and lack of experimental tests to verify the model.

Mangold (2007) also studied 22 semantic document retrieval systems and proposed a categorization scheme based on seven features: architecture, coupling, transparency, user context, query modification, ontology structure and ontology technology.

Architecture concerns whether the underlying design of system stores an index of documents. A system is considered as a stand-alone search engine if it maintains indexes of the document data. Examples include SHOE (Heflin & Hendler, 2000) and SCORE (Sheth et al. 2002). A system is called a Meta search engine if it sends queries to subordinate search-engines. For instance, Inquirus2 (Glover et al., 2001) and TAP (Guha et al., 2003) employs the idea of Meta search engine.

Coupling is about how close the relationship between the documents and ontologies. In tight coupling, the ontology concepts are closely related to the documents. For instance, both hybrid spreading activation (Rocha et al., 2004) and librarian agent (Stojanovic, 2003) adopts tight coupling approach. On the contrary, the loose coupling refers to the case that documents are not committed to any ontology. Both ISRA (Burton-Jones et al., 2003) and TAP (Guha et al., 2003) make use of loose coupling.

Transparency refers to the degree of interaction between users and systems. In the one end of the spectrum, some systems, such as audio data retrieval (Khan et al., 2004) and Inquirus2 (Glover et al., 2001), remain silence to users' requests and their semantic features are totally transparent to the users. In the other end, some systems, such as Ontogator (Hyvönen et al., 2003) and SCORE (Sheth et al. 2002),

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/multidimensional-ontology-based-information-retrieval-for-academic-counseling/198622

Related Content

Finding Answers to Questions, in Text Collections or Web, in Open Domain or Specialty Domains

Brigitte Grau (2012). *Next Generation Search Engines: Advanced Models for Information Retrieval* (pp. 344-370).

www.irma-international.org/chapter/finding-answers-questions-text-collections/64432

Complex Terminology Extraction Model from Unstructured Web Text Based Linguistic and Statistical Knowledge

Fethi Fkihhand Mohamed Nazih Omri (2012). *International Journal of Information Retrieval Research* (pp. 1-18).

www.irma-international.org/article/complex-terminology-extraction-model-unstructured/78311

Efficacious Hyperlink Based Similarity Measure Using Heterogeneous Propagation of PageRank Scores

Vasanth Thangasamy (2019). *International Journal of Information Retrieval Research* (pp. 36-49).

www.irma-international.org/article/efficacious-hyperlink-based-similarity-measure-using-heterogeneous-propagation-of-pagerank-scores/236655

Big Data Analytics in Bioinformatics

Dharmendra Trikam Lal Patel (2018). *Handbook of Research on Biomimicry in Information Retrieval and Knowledge Management* (pp. 321-338).

www.irma-international.org/chapter/big-data-analytics-in-bioinformatics/197708

Applications of AI in Computer-Aided Drug Discovery

Reet Kaur Kohli, Seneha Santoshi, Sunishtha S. Yadav and Vandana Chauhan (2023). *Applying AI-Based IoT Systems to Simulation-Based Information Retrieval* (pp. 77-89).

www.irma-international.org/chapter/applications-of-ai-in-computer-aided-drug-discovery/322851