

The 2012 International Workshop on Web-scale Knowledge Representation, Retrieval, and Reasoning

Spyros Kotoulas
Smarter Cities Technology
Center
IBM Research
Dublin, Ireland
spyros.kotoulas@ie.ibm.com

Yi Zeng
Institute of Automation
Chinese Academy of Sciences
Beijing, P.R. China
yi.zeng@ia.ac.cn

Zhisheng Huang
Department of Computer
Science
VU University Amsterdam
Amsterdam, The Netherlands
huang@cs.vu.nl

ABSTRACT

The rapid and perpetual growth of knowledge on the Web has given rise to many grand challenges (such as scalability, inconsistency, uncertainty, distribution and dynamics) for traditional knowledge processing methods and systems. Knowledge representation, retrieval and reasoning methods need to evolve and adapt to the Web to face these challenges and make this vast, heterogenous knowledge useful and accessible. In this light, the International Workshop on Web-scale Knowledge Representation, Retrieval, and Reasoning (Web-KR) is initiated. This workshop serves as the third one in this workshop series. This summary discusses the scope of Web-KR and introduces the advances in this field through the accepted papers in the Web-KR 2012 workshop, co-located with CIKM 2012.

Categories and Subject Descriptors

I.2.4 [ARTIFICIAL INTELLIGENCE]: Knowledge Representation Formalisms and Methods; I.2.11 [ARTIFICIAL INTELLIGENCE]: Distributed Artificial Intelligence; H.3.3 [INFORMATION STORAGE AND RETRIEVAL]: Information Search and Retrieval

General Terms

Algorithms, Human Factors, Experimentation

Keywords

Knowledge Representation, Knowledge Retrieval, Semantic Search, Web reasoning, Scalability

1. INTRODUCTION

The World Wide Web has become the carrier for the largest human knowledge repository in history. As its knowledge bases are growing towards a practically 'infinite' volume, Web-scale Knowledge Representation, Retrieval, and Reasoning (Web-KR) is becoming a real issue and an urgent task. Although the Web community has developed a number of knowledge representation languages and reasoning methods, when the volume, goes Web-scale, existing approaches meet many challenging problems, such as scal-

ability, inconsistency, uncertainty and dynamics. Hence, a unified approach to Web-KR needs to be developed.

Following this motivation, a workshop series titled "Web-scale Knowledge Representation, Retrieval, and Reasoning (Web-KR)" was initiated. The workshop under this title has been held for the past two years, in conjunction with the IEEE/WIC/ACM International Conference on Web Intelligence (2010 in Toronto, Canada and 2011 in Lyon, France).

Building on previous successful iterations, this workshop is the third one in this workshop series, aiming to bring together researchers from Web research, Artificial Intelligence (AI), high performance computing, cognitive science, knowledge management, and machine learning to discuss all issues of Web-KR in a synergistic setting. The workshop is co-located with the 2012 ACM International Conference on Information and Knowledge Management (CIKM 2012) in Maui Hawaii, the United States.

2. TOPICS OF INTERESTS

Through the development of this workshop series in the recent three years, the following topics have been considered as important research trends for knowledge representation, retrieval and reasoning at Web scale:

- Expressivity of Web-scale common sense knowledge.
- Parallelization and distribution techniques for Web-scale knowledge retrieval and reasoning.
- Inconsistency, uncertainty, and relevancy of Web-scale reasoning.
- Granularity of knowledge representation, retrieval and reasoning.
- Web enabled real-time retrieval and reasoning.
- New forms of inductive and deductive reasoning (e.g. cognitive and nature inspired methods).
- Ontology and rule interchange and integration.(e.g. via OWL 2, RIF, and RuleML)
- Spatial and Temporal reasoning methods for the Web.
- Dynamics of Web knowledge updating and forgetting.
- Distributed query answering over Web-scale data.
- Cross-fertilizing Web-KR with Distributed Artificial Intelligence, Distributed Description Logic, and Distributed Logic Programming.

Many of the accepted papers in previous two years have investigated in these topics in depth from various perspectives. The following three papers deal with Web-scale reasoning and query processing: In the 2010 workshop in Toronto, “RDFS Backward Chained Reasoning Support in 4store” was introduced as early attempt for parallel backward chained reasoning for semantic Web data [9]. In the 2011 workshop in Lyon, TripleCloud was introduced as a cloud based Web-scale query engine which benefits from existing cloud based key-value stores (e.g. BigTable, HBase) and scales to billion triples for complex queries [3]. While in [7], the authors introduced an ant behavior inspired algorithm that deals with Web-scale range queries (e.g. queries on data annotated with location and time) [7].

From the perspective of Web-scale knowledge acquisition and representation, and in particular large scale categorized Wikipedia contents, [5] introduced an approach for mining fuzzy domain ontologies based on concept vectors. From the application perspective, efficient representation methods for dynamic Web pages have also been investigated [2]. In addition, effective visualization techniques have been applied to support browsing online Multimedia Digital Library [6]. Nevertheless, there are still several topics related to Web-KR remained to be explored.

3. ACCEPTED PAPERS

We are very pleased to find that the scope and interesting topics for Web-KR has been extended by the submissions in the 2012 version of this workshop.

For more general purpose knowledge processing on the Web, OmpiJava is proposed and implemented as a reaction to grand challenges for data-centric supercomputing on the Web [1]. The author introduces the Message Passing Interface (MPI) bindings to Java, seamlessly integrated in Open MPI. Based on the author’s contribution, the implementation “enables Java based Semantic Web applications to be successfully ported to the most of modern HPC systems” [1]. A parallel version of random indexing is implemented as an illustrative example to validate OmpiJava.

In order to provide a human centered approach for Web-scale knowledge organization and retrieval, in [8], the authors propose to apply semiotic (i.e., sub-syntactical) and inductive (i.e., probabilistic) methods for inferring concept associations. Fuzzy semantic networks are used to represent knowledge on the Web. In addition, the authors provide interactive visualization tools to visualize these cognitive concept maps, and an exploratory search paradigm is taken for browsing and searching the organized semantic network [8].

In contrast with the deductive approaches for Web-scale reasoning introduced in this workshop series during these three years, in [4], the authors take a data mining approach for finding useful knowledge. More specifically, this paper formally proposes and defines the task of mining frequent correlated sequential patterns from a sequential database. The PSBspan algorithm is proposed based on pattern-growth methodology. The validation through finding popular tourism trails based on Flickr data shows the effectiveness of the proposed algorithm on processing Web entities with sequential characteristics. We see approaches that combine large scale deductive reasoning and inductive reasoning on the Web as a very promising direction for Web-KR.

Acknowledgments

The organizers of the Web-KR 2012 workshop would like to thank all the PC members for their continuous support on the workshop. We would also like to thank the CIKM 2012 conference organizers for providing an interdisciplinary venue (with researchers and practitioners from various backgrounds, including Information Retrieval, Database, and Knowledge Management, etc.) to disseminate and discuss research topics and results related to Web-scale knowledge representation, retrieval and reasoning.

4. REFERENCES

- [1] A. Cheptsov. Ompijava - a tool for development of high-performance reasoning applications for the semantic web. In *Proceedings of the 2012 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, 2012.
- [2] M. Farouk and M. Ishizuka. Semantic structure content for dynamic web pages. In *Proceedings of the 2010 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 253–256, 2010.
- [3] C. Gueret, S. Kotoulas, and P. Groth. Triplecloud: An infrastructure for exploratory querying over web-scale rdf data. In *Proceedings of the 2011 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 245–248, 2011.
- [4] C. X. Lin, M. Ji, M. Danilevsky, and J. Han. Efficient mining of correlated sequential patterns based on null hypothesis. In *Proceedings of the 2012 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, 2012.
- [5] C.-Y. Lu, S.-W. Ho, J.-M. Chung, F.-Y. Hsu, H.-M. Lee, and J.-M. Ho. Mining fuzzy domain ontology based on concept vector from wikipedia category network. In *Proceedings of the 2011 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 249–252, 2011.
- [6] S. Mishra, A. Gorai, T. Oberoi, and H. Ghosh. Efficient visualization of content and contextual information of an online multimedia digital library for effective browsing. In *Proceedings of the 2010 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 257–260, 2010.
- [7] H. Muhleisen, T. Walther, and R. Tolksdorf. Web-scale range queries in a self-organized storage service. In *Proceedings of the 2011 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 241–244, 2011.
- [8] E. Portmann, M. A. Kaufmann, and C. Graf. A distributed, semiotic-inductive, and human-oriented approach to web-scale knowledge retrieval. In *Proceedings of the 2012 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, 2012.
- [9] M. Salvadores, G. Correndo, T. Omitola, N. Gibbins, S. Harris, and N. Shadbolt. 4s-reasoner: Rdfs backward chained reasoning support in 4store. In *Proceedings of the 2010 International Workshop on Web-Scale Knowledge Representation, Retrieval, and Reasoning*, pages 261–264, 2010.