IRMLeS 2009

1st ESWC Workshop on Inductive Reasoning and Machine Learning on the Semantic Web

Foreword

The goal of the *Semantic Web* research is to extend the current World Wide Web into one where information is given semantically precise meaning. The Semantic Web is supposed to provide a medium to share such information and is often perceived to be a global, distributed database of meaningfully represented data. From the speed of the development of the initiatives, like Linked Open Data or community efforts to integrate various ontological resources like bio-ontologies, we may expect to have, in the near future, access to enormous data resource.

Hence, the problem of mining from the Semantic Web currently becomes as important as it used to be that of data mining from local databases. Exploiting this global resource of data requires new kinds of approaches for data mining and data analysis, which would be able to deal with the complexity and expressiveness of the representation languages and with novel assumptions that underlie reasoning services within the Semantic Web. Open, distributed and inherently incomplete nature of the Semantic Web environment poses problems for deductive approaches, traditionally employed to reason with logic-based ontological data. Hence, one may witness a recent trend in the Semantic Web community to propose complementary forms of reasoning, preferably more efficient and noise-tolerant.

The IRMLeS workshop was meant to be a forum for scientific exchange amongst researchers interested in an interdisciplinary research on the intersec-

tion of the Semantic Web with Knowledge Discovery and Machine Learning fields. It was meant to act as a meeting point for the sub-communities from these fields that are interested in research on the challenging problems of such intersection.

The central goal was to analyze how machine learning techniques, such as statistical learning methods and inductive forms of reasoning, can be directly applied on richly structured Semantic Web data, and to discuss the value added of machine learning methods for the Semantic Web context. The focus of the workshop covered the following main topics from the general scope of the conference: machine learning methods for Semantic Web Data and Ontologies (especially ontology learning, creation, evolution, querying and mapping); Applications on the Semantic Web (data mining). Moreover, it addressed the need for new forms of reasoning on the Semantic Web that would complement the deductive ones, traditionally studied in this context.

Specifically, the workshop received a total of 18 submissions. The review phase produced a selection of 10 full papers and 3 short papers for the program. IRMLeS'2009 program was further enriched by two invited talks from prominent researchers. Professor Ross D. King presented in his talk how ontologies may help in automating science, by providing means to formalize and expose the experimental data, and how such formalization opens up new opportunities to apply data mining to discover new knowledge. Dr Rayid Ghani gave a talk from an industry perspective, in which he discussed how the structured content from internal databases of companies can be augmented by applying machine learning techniques to the unstructured content for a variety of business problems. The workshop was also successful in terms of registrations and attendance.

The application of knowledge discovery techniques to the Semantic Web, which is traditionally the most prominent form of interaction between the two fields, regarded topics related to data management, semantic retrieval and text mining, ontology matching and rule design. However a number of the presented works covered the opposite direction (applying Semantic Web techniques in the knowledge discovery process) especially as concerns mining and evaluation.

In the final wrap-up discussion, a number of open problems and promising directions were elicited. The integration of induction and deduction has been recognized as essential for the Semantic Web, since deductive reasoning has been found insufficient for dealing with real, noisy data. Relational learning was reckoned as very appropriate for the web of data, which still calls for standardization.

On the other hand, a bit of semantics often makes a difference for the applications of knowledge discovery. However the complexity of the problems posed by the standard knowledge representation calls for new ideas and techniques.

In order to establish a research community for this field, common benchmarking datasets must be shared, e.g. a collection of ontologies with large A-boxes. Even more so, it is the Semantic Web community which will pose new challenges for the machine learning community. It was also proposed the possibility of forming a W3C interest group (or at least an incubator) where this community may gather to discuss the mentioned problems.

Scalability and distribution of computing resources were pointed out as problems which were not sufficiently covered during the workshop presentations, while they are reckoned as worthwhile for directing the practitioners' future research. Given such open issues and the success of this first edition, we sense that this event is going to have a sequel in the near future.

Acknowledgements The workshop chairs are grateful to all the people who contributed to the event from the program committee members to the additional reviewers, the presenters and the participants. A special thank is due to the invited speakers who shared their vision of the possible applications of the semantic technologies to classic inductive methods. Finally, we are grateful also to the ESWC 2009 Chairs for the positive response to our proposal.

Heraklion, June 1st, 2009

Claudia d'Amato
Nicola Fanizzi
Marko Grobelnik
Agnieszka Ławrynowicz
Vojtěch Svátek

Program Committee Members

- Sarabjot S. Anand - University of Warwick
- Bettina Berendt - Katholieke Universiteit Leuven
- Sonia Bergamaschi - University of Modena-Reggio Emilia
- Sebastian Blohm - University of Karlsruhe
- Floriana Esposito - University of Bari
- Mohand-Said Hacid - University of Lyon 1
- Pascal Hitzler - University of Karlsruhe
- Andreas Hotho - University of Kassel
- Jos Iria - University of Sheffield
- Maciej Janik - University of Koblenz-Landau
- Matthias Klusch - DFKI - Saarbrücken
- Jens Lehmann - University of Leipzig
- Francesca A. Lisi - University of Bari
- Thomas Lukasiewicz - Oxford University
- Matthias Nickles - University of Bath
- Sebastian Rudolph - University of Karlsruhe
- Jetendr Shamdasani - University of the West of England
- Steffen Staab - University of Koblenz-Landau
- Umberto Straccia - ISTI-CNR, Pisa
- York Sure - University of Koblenz-Landau
- Valentina Tamma - University of Liverpool
- Johanna Völker - University of Karlsruhe

Additional Reviewers

Giovanna Castellano - University of Bari
Maurizio Vincini - University of Modena-Reggio Emilia

Workshop Homepage

http://irmles2009.di.uniba.it