Lecture Notes in Artificial Intelligence 10326

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

James Cussens · Alessandra Russo (Eds.)

Inductive Logic Programming

26th International Conference, ILP 2016
London, UK, September 4–6, 2016
Revised Selected Papers

Springer

Editors
James Cussens
Department of Computer Science
University of York
York
UK

Alessandra Russo
Imperial College London
London
UK

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-63341-1 ISBN 978-3-319-63342-8 (eBook)
DOI 10.1007/978-3-319-63342-8

Library of Congress Control Number: 2017946684

LNCS Sublibrary: SL7 – Artificial Intelligence

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume constitutes the proceedings of the 26th International Conference on Inductive Logic Programming (ILP 2016) and includes a selection of the papers presented at the conference. ILP 2016 was held in London, during September 4–6, 2016, at the Warren House Conference Centre. Since its first edition in 1991, the annual ILP conference has served as the premier international forum for learning from structured relational data. Originally focusing on the induction of logic programs, over the years it has expanded its research horizon significantly and welcomed contributions on all aspects of learning in logic, multi-relational data mining, statistical relational learning, graph and tree mining, learning in other (non-propositional) logic-based knowledge representation frameworks, exploring intersections with statistical learning, and other probabilistic approaches. Theoretical advances in all these areas have also been accompanied by challenging applications of these techniques to important problems in fields like bioinformatics, medicine, and text mining.

Following the trend of past events, this edition of the conference solicited three types of submissions: (a) long papers describing original mature work containing appropriate experimental evaluation and/or representing a self-contained theoretical contribution; (b) short papers describing original work in progress, brief accounts of original ideas without conclusive evaluation, and other relevant work of potentially high scientific interest but not yet qualifying for the long paper category; and finally (c) papers relevant to the conference topics and recently published or accepted for publication by a first-class conference such as ECML/PKDD, ICML, KDD, ICDM, AAAI, IJCAI, or a journal such as MLJ, DMKD, JMLR etc.

The conference received 35 submissions: ten long papers, 19 short papers, and six published papers. Each of the long and short paper submissions was reviewed by three Program Committee (PC) members. Only four of the ten submitted long papers were accepted for presentation and publication. Short papers were initially evaluated on the basis of the submitted manuscript and the presentation, and authors of a subset of these papers were invited to submit an extended version. After a second review process, only six extended papers were finally accepted for publication. In summary, together with the four long papers, ten papers were accepted to be included in the present volume. The multiple-stage review process, although rather complex, has enabled the selection of high-quality papers for the proceedings. We thank the members of the PC for providing high-quality and timely reviews. Out of all the submitted papers, an additional 13 papers were accepted for publication in the CEUR workshop proceedings series.

The ILP 2016 program included five large technical sessions: Logic and Learning; Graphs and Databases; Probabilistic Logic and Learning; Algorithms, Optimisations and Implementations; and Applications. The papers in this volume represent well the current breadth of ILP research topics such as predicate invention, graph-based learning, spatial learning, logical foundations, statistical relational learning,

probabilistic ILP, implementation and scalability, and applications in robotics, cyber-security, and games, providing also an excellent balance across theoretical and practical research. ILP 2016 received generous sponsorship by the *Machine Learning* journal for best student paper awards. The two best student paper awards of ILP 2016 were given to Yi Huang for his paper entitled "Learning Disjunctive Logic Programs from Interpretation Transition," co-authored with Yisong Wang, Ying Zhang and Mingyi Zhang, and to Marcin Malec for his paper "Inductive Logic Programming Meets Relational Databases: An Application to Statistical Relational Learning," co-authored with Tushar Khot, James Nagy, Erik Blasch and Sriraam Natarajan. The conference also received sponsorship from Springer for a best paper award. This award was given to the paper "Generation of Near-Optimal Solutions Using ILP-Guided Sampling" by Ashwin Srinivasan, Gautam Shroff, Lovekesh Vig and Sarmimala Saikia.

With the intent of stimulating collaborations and discussion between academia and industry, the program also featured three invited talks by academic and industrial distinguished researchers. In the talk "Inferring Causal Models of Complex Relational and Dynamic Systems," David Jensen, from the University of Massachusetts, presented key ideas, representations, and algorithms for causal inference, and highlighted new technical frontiers. Frank Wood, from the University of Oxford, gave a talk entitled "Revolutionising Decision Making, Democratising Data Science, and Automating Machine Learning via Probabilistic Programming." In his talk, he gave a broad overview of the emerging field of probabilistic programming, from the point of view of both programming (modelling) language and automated inference, and introduced the most important challenges facing this field. Finally, Vijay Saraswat, senior research scientist in the Cognitive Computing Research division at the IBM T.J. Watson Research Center, discussed in his talk "Machine Learning and Logic: The Beginnings of a New Computer Science?" the open challenges of building cognitive assistants in compliance, and the need to bring together researchers in natural language understanding, machine learning, and knowledge representation/reasoning to address them.

The conference featured, for the first time, an international competition, designed and managed by Mark Law, a member of our local Organizing Committee. The competition was aimed at testing the accuracy, scalability, and versatility of the learning systems that were entered. The competition had two main tracks for probabilistic and non-probabilistic approaches. The winners of the competition were Peter Schüller, from Marmara University, for his non-probabilistic approach and jointly Riccardo Zese, Elena Bellodi, and Fabrizio Riguzzi for their probabilistic approach. Results of the competition are publicly available on http://ilp16.doc.ic.ac.uk/competition.

The ILP 2016 conference was kindly sponsored by IBM Watson Research, the Association of Logic Programming, Springer's *Lecture Notes in Artificial Intelligence*, the *Artificial Intelligence* journal, and the *Machine Learning* journal. We would like to thank EasyChair for supporting the submission handling. We would like to thank the members of the local Organizing Committee of ILP 2016: Krysia Broda, Dalal Alrajeh, and Mark Law. Our thanks also go to Mark Law for running the competition and for setting up and maintaining the website. The conference would not have been possible without their hard work.

Finally, we would like to thank all those involved in making ILP 2016 such a success: our invited speakers, our sponsors, the PC and, of course, those who came to ILP 2016 to present and discuss their work.

May 2017

James Cussens
Alessandra Russo

Organization

Organizing Committee

Program Co-chairs

James Cussens	University of York, UK
Alessandra Russo	Imperial College, UK

Competition Chair

Mark Law	Imperial College, UK

Financial Chair

Dalal Alrajeh	Imperial College, UK

Publicity Chair

Krysia Broda	Imperial College, UK

Program Committee

Dalal Alrajeh	Imperial College London, UK
Alexander Artikis	Institute of Informatics and Telecommunications, Greece
Krysia Broda	Imperial College London, UK
Rui Camacho	LIACC/FEUP University of Porto, Portugal
Luc De Raedt	Katholieke Universiteit Leuven, Belgium
Sašo Džeroski	University of Ljubljana, Slovenia
Floriana Esposito	Università degli Studi di Bari, Italy
Nicola Fanizzi	Università degli Studi di Bari, Italy
Stefano Ferilli	Università degli Studi di Bari, Italy
Nuno Fonseca	European Bioinformatics Institute, Portugal
Katsumi Inoue	National Institute of Informatics, Japan
Kristian Kersting	TU Dortmund University, Germany
Ross King	University of Manchester, UK
Nicolas Lachiche	University of Strasbourg, France
Nada Lavrač	Jožef Stefan Institute, Slovenia
Francesca Lisi	Università degli Studi di Bari, Italy
Donato Malerba	Università degli Studi di Bari, Italy
Stephen Muggleton	Imperial College London, UK
Aline Paes	UFF, Federal Fluminense University, Brazil
Jan Ramon	Inria, France
Oliver Ray	University of Bristol, UK

Fabrizio Riguzzi	University of Ferrara, Italy
Chiaki Sakama	Wakayama University, Japan
Vítor Santos Costa	Universidade do Porto, Portugal
Takayoshi Shoudai	Kyushu International University, Japan
Alireza Tamaddoni-Nezhad	Imperial College London, UK
Christel Vrain	University of Orleans, France
Stefan Wrobel	Fraunhofer IAIS and University of Bonn, Germany
Akihiro Yamamoto	Kyoto University, Japan
Gerson Zaverucha	PESC-COPPE, UFRJ, Brazil
Filip Železný	Czech Technical University, Czech Republic

Sponsoring Institutions

Association for Logic Programming
IBM Research
Machine Learning Journal (for best student paper awards)
Lecture Notes in Artificial Intelligence, Springer (for best paper award)
Artificial Intelligence Journal

Invited Speakers

Inferring Causal Models of Complex Relational and Dynamic Systems

David Jensen

Knowledge Discovery Laboratory, Computational Social Science Institute,
College of Information and Computer Sciences,
University of Massachusetts Amherst, Amherst, USA

Over the past 25 years, surprisingly effective techniques have been developed for inferring causal models from observational data. While traditional models reason about a given system by assuming that its behavior is stationary, causal models reason about how a system will behave under intervention. Unfortunately, nearly all existing methods for causal inference assume that data instances are independent and identically distributed, making them inappropriate for analyzing many social, economic, biological, and computational systems. In this talk, I will explain the key ideas, representations, and algorithms for causal inference, and I will describe very recent developments that extend those techniques to complicated systems with relational and dynamic behavior. I will describe practical methods for evaluating methods for causal inference and identify some of the most pressing research questions and new technical frontiers.

Machine Learning and Logic—The Beginnings of a New Computer Science?

Vijay A. Saraswat

IBM T.J. Watson Research Lab, New York, USA

Our long-term research goal in Cognitive Computing Research at IBM is to develop systems that know deeply, learn continuously, reason with purpose and interact naturally. To further this agenda, we are focusing on a few deep domains. This talk will address the challenges of building cognitive assistants in compliance assistants that deal with understanding and reasoning about the myriad (corporate, financial, privacy, ethical) laws and regulations within the context of which modern international businesses must operate. An interim goal for the compliance cognitive assistant is to clear the Uniform CPA exam, a professional certification attempted by master's level students. We will outline the tremendous technical challenges underlying this goal and our current approaches. We believe the key to achieving this goal is bringing together researchers in natural language understanding, machine learning, and knowledge representation/reasoning for a concerted attack on this problem.

Revolutionizing Decision Making, Democratizing Data Science, and Automating Machine Learning via Probabilistic Programming

Frank Wood

Department of Engineering Science, University of Oxford, Oxford, UK

Probabilistic programming aims to enable the next generation of data scientists to easily and efficiently create the kinds of probabilistic models needed to inform decisions and accelerate scientific discovery in the realm of big data and big models. Model creation and the learning of probabilistic models from data are key problems in data science. Probabilistic models are used for forecasting, filling in missing data, outlier detection, cleanup, classification, and scientific understanding of data in every academic field and every industrial sector. While much work in probabilistic modeling has been based on hand-built models and laboriously-derived inference methods, future advances in model-based data science will require the development of much more powerful automated tools than currently exist. In the absence of such automated tools, probabilistic models have traditionally co-evolved with methods for performing inference. In both academic and industrial practice, specific modeling assumptions are made not because they are appropriate to the application domain, but because they are required to leverage existing software packages or inference methods. This intertwined nature of modeling and computation leaves much of the promise of probabilistic modeling out of reach for even expert data scientists. The emerging field of probabilistic programming will reduce the technical and cognitive overhead associated with writing and designing novel probabilistic models by both introducing a programming (modeling) language abstraction barrier and automating inference. The automation of inference, in particular, will lead to massive productivity gains for data scientists, much akin to how high-level programming languages and advances in compiler technology have transformed software developer productivity. What is more, not only will traditional data science be accelerated, but the number and kind of people who can do data science also will be dramatically increased. My talk will touch on all of this, explain how to develop such probabilistic programming languages, highlight some exciting ways such languages are starting to be used, and introduce what I think are some of the most important challenges facing the field as we go forward.

Contents