

# CONFERENCE PROGRAM



19th International Conference on Inductive Logic Programming  
7th International Workshop on Mining and Learning with Graphs  
6th International Workshop on Statistical Relational Learning

July 2-4, 2009, Leuven, Belgium





# Contents

<b>Conference Program</b>	<b>1</b>
<b>Sponsors</b>	<b>10</b>
<b>Organization</b>	<b>11</b>
<b>Programme Committees</b>	<b>12</b>
<b>About Leuven</b>	<b>13</b>
<b>Invited Speakers</b>	<b>14</b>
<b>Abstracts ILP</b>	<b>17</b>
<b>Abstracts SRL</b>	<b>33</b>
<b>Abstracts MLG</b>	<b>40</b>
<b>Notes</b>	<b>47</b>

## Dear Participant,

Welcome to ILP-MLG-SRL 2009, which combines the 19th International Conference on Inductive Logic Programming, the 7th International Workshop on Mining and Learning with Graphs, and the 6th International Workshop on Statistical Relational Learning.

While the above three events each have their own specific focus, it has been observed many times that there are clear connections between them, that they in a sense study the same problem, and that there is a significant overlap between the research areas they cover. This overlap has become more important over the years, thanks to a growing general interest in logical and relational learning, learning from graphs, statistical relational learning, and probabilistic logics. This has also caused the different research areas to drift towards each other. The corresponding research communities, however, have remained more separate.

It is this observation that provided the main motivation for co-locating the three events. The combined event has been set up to cater maximally for interaction between researchers from the different communities. Sessions have been organized mostly thematically, rather than adhering to a strict separation between the different workshops. Six invited speakers, a panel and a careful selection of orally presented papers (with an acceptance rate of about 30%) guarantee an exciting program, while an extensive poster program will foster interaction between a maximal number of researchers. The *Stuk* cultural center where the event takes place provides ample opportunities for smaller groups of researchers to separate themselves from the crowd in order to discuss new ideas and perhaps start new collaborations. Finally, the small but lively town of Leuven provides many opportunities for continuing scientific discussions into the evening in a somewhat more relaxed atmosphere, or simply enjoying the food, drinks, music, and other diversions that Leuven has to offer.

We wish you an enjoyable and intellectually inspiring stay in Leuven.

The ILP, MLG and SRL program chairs,

*Hendrik Blockeel, Karsten Borgwardt, Luc De Raedt,  
Pedro Domingos, Kristian Kersting and Xifeng Yan*

# CONFERENCE PROGRAM

Time	Wednesday (1/7) Evening
17:00	<b>Registration</b> <i>STUK</i>
19:30	Welcome Reception in Town Hall

**STUK:** The conference takes place in the Arenberg institute of the Katholieke Universiteit Leuven (for talks and posters) and the art centre STUK (for breaks, lunches, ...). Arenberg institute and STUK are part of the same building, at a convenient walking distance from the centre of Leuven with its many restaurants and pubs.

**City hall:** The welcome reception is given in Leuven's magnificent city hall. Built in the the medieval heydays of Leuven (15th century), its various architects gave the building its flamboyant Gothic look. In the 19th century it was further embellished with neo-gothic elements, including the 236 statues of artists, scientists, historic persons and other people that played important roles in the history of the city.



**Figure 1:** *Outside the STUK*



**Figure 2:** *Leuven Town Hall*

Conference Program

---

Time	Thursday (2/7) Morning	
08:20	<b>Registration</b> <i>STUK</i>	
08:50	<b>Opening</b> <i>Auditorium</i>	
09:00	<b>Invited Talk: Philip Yu</b> <i>Auditorium</i>	
10:00	<b>Coffee</b> <i>STUK Café</i>	
	<b>ILP</b> <i>Auditorium</i>	<b>MLG/SRL</b> <i>Verbeeck</i>
10:30	<i>Grant Anderson and Bernhard Pfahringer</i> Relational Random Forests based on Random Relational Rules	<i>(M) Leander Schietgat, Fabrizio Costa, Jan Ramon and Luc De Raedt</i> Maximum Common Subgraph Mining: A Fast and Effective Approach towards Feature Generation
10:56	<i>Katsumi Inoue, Koichi Furukawa and Ikuo Kobayashi</i> Abducing Rules with Predicate Invention	<i>(M) Christian Desrosiers, Alain Hertz, Philippe Galinier and Hansen Pierre</i> Mining Graphs to Discover New Theorems in Mathematics
11:22	<i>Louis Oliphant, Elizabeth Burnside and Jude Shavlik</i> Boosting First-Order Clauses for Large, Skewed Data sets	<i>(M) Shankar Vembu, Thomas Gärtner and Mario Boley</i> Abstract: MAP Structured Prediction by Sampling
11:48	<i>Ondřej Kuželka and Filip Zelezny</i> Block-Wise Construction of Acyclic Relational Features with Monotone Relevancy	<i>(S) Sebastian Riedel</i> Cutting Plane MAP Inference for Markov Logic
12:14	<i>Marco Lippi, Manfred Jaeger, Paolo Frasconi and Andrea Passerini</i> Relational Information Gain	<i>(S) Tuyen Huynh and Raymond Mooney</i> Max-Margin Weight Learning for Markov Logic Networks
12:40	<b>Lunch</b> <i>Stuk Café</i>	

Time	Thursday (2/7) Afternoon
14:00	<p><b>Invited Talk: Raymond Mooney</b>  <i>Auditorium</i></p>
15:00	<p><b>Joint Poster Spotlight Session</b>  <i>Auditorium</i></p> <p><i>Julien Lesbegueries, Agnès Braud and Nicolas Lachiche</i>  A Propositionalisation that Preserves More Continuous Attribute Domains</p> <p><i>Francesca Alessandra Lisi and Floriana Esposito</i>  Nonmonotonic Onto-Relational Learning</p> <p><i>Stephen Muggleton, Aline Paes, Vítor Santos Costa and Gerson Zaverucha</i>  Chess Revision: Acquiring the Rules of Chess Variants through FOL Theory Revision from Examples</p> <p><i>Sriraam Natarajan, Gautam Kunapuli, Ciaran O'Reilly, Richard Maclin, Trevor Walker, David Page and Jude Shavlik</i>  ILP for Bootstrapped Learning: A Layered Approach to Automating the IL P Set-Up Problem</p> <p><i>Niels Pahlavi and Stephen Muggleton</i>  Higher-order Logic Learning</p> <p><i>Anup Patel, Ganesh Ramakrishnan and Pushpak Bhattacharya</i>  Incorporating Linguistic Expertise Using ILP for Named Entity Recognition in Data Hungry Indian Languages</p> <p><i>Scott Proper and Prasad Tadepalli</i>  Transfer Learning via Relational Templates</p> <p><i>Oliver Ray, Ken Whelan and Ross King</i>  Automatic Revision of Metabolic Networks through Logical Analysis of Experimental Data</p> <p><i>Lothar Richter, Regina Augustin and Stefan Kramer</i>  Finding Relational Associations in HIV Resistance Mutation Data</p> <p><i>Hannes Schulz, Kristian Kersting and Andreas Karwath</i>  ILP, the Blind, and the Elephant: Euclidean Embedding of Co-Proven Queries</p> <p><i>Ashwin Srinivasan and Ganesh Ramakrishnan</i>  Parameter Screening and Optimisation for ILP using Designed Experiments</p> <p><i>Gabriel Synnaeve, Andrei Doncescu and Katsumi Inoue</i>  Kinetic Models for Logic-Based Hypothesis Finding in Metabolic Pathways</p> <p><i>Ingo Thon</i>  Don't Fear Optimality: Sampling for Probabilistic-Logic Sequence Models</p>

Time	Thursday (2/7) Afternoon
	<p><i>Daisuke Tokoro, Ai Fukunaga, Kiyoharu Hamaguchi, Toshinobu Kashiwabara and Shin-ichi Minato</i> Exploiting Global Structures in Bayesian Network Compilation by Zero-suppressed BDDs</p> <p><i>Hiroaki Watanabe and Stephen Muggleton</i> Can ILP be Applied to Large Dataset?</p> <p><i>Elma Akand, Michael Bain and Mark Temple</i> Learning Responsive Cellular Networks by Integrating Genomics and Proteomics Data</p> <p><i>Guillaume Damiand, Colin de la Higuera, Jean-Christophe Janodet, Emilie Samuel and Christine Solnon</i> A Polynomial Algorithm for Subisomorphism of Open Plane Graphs</p> <p><i>Frank Eichinger and Klemens Böhm</i> Towards Scalability of Graph-Mining Based Bug Localisation</p> <p><i>Risi Kondor and Nino Shervashidze</i> The Graphlet Spectrum</p> <p><i>Olana Missura and Thomas Gärtner</i> Classification on Graphs for Dynamic Difficulty Adjustment: More Questions Than Answers</p> <p><i>Tsuyoshi Murata</i> New Modularity for Evaluating Communities in Bipartite Networks</p> <p><i>Alberto Pinto</i> A Statistical Generalization of Music Graph Spectral Mining</p> <p><i>Albert Bifet and Ricard Gavaldà</i> Adaptive XML Tree Mining on Evolving Data Streams</p> <p><i>Angelika Kimmig, Bernd Gutmann and Vítor Santos Costa</i> Trading Memory for Answers: Towards Tabling ProbLog</p> <p><i>Stanley Kok and Pedro Domingos</i> Hypergraph Lifting for Structure Learning in Markov Logic Networks</p> <p><i>Matt Michelson and Sofus Macskassy</i> Layered, Multivariate Anomaly Explanations: A First Look</p> <p><i>Sriraam Natarajan, Prasad Tadepalli, Gautam Kunapuli and Jude Shavlik</i> Knowledge Intensive Learning - Directed vs Undirected SRL Models</p> <p><i>Vítor Santos Costa and Aline Paes</i> On Connecting PRISM to CLP(BN)</p>



Time	Thursday (2/7) Afternoon
	<p style="text-align: center;"><i>Joost Vennekens, Angelika Kimmig, Theofrastos Mantadelis, Bernd Gutmann, Luc De Raedt and Maurice Bruynooghe</i> From ProbLog to First Order Logic: A First Exploration</p> <p style="text-align: center;"><i>Hiroaki Watanabe and Stephen Muggleton</i> Abstracting Probabilistic Relational Model by Specialisation</p> <p style="text-align: center;"><i>Maurice Bruynooghe, Daan Fierens, Bernd Gutmann, Angelika Kimmig, Niels Landwehr, Wannes Meert, Ingo Thon and Luc De Raedt</i> An Exercise with Statistical Relational Learning Systems</p>
16:15	<p><b>Coffee followed by poster reception</b> <i>STUK Café and Ensemble room</i></p>
18:45	Guided tour
20:00	Banquet in Faculty Club

**Guided tour:** After the last session we meet for a guided tour through the historic centre of Leuven. We jump back into the early middle ages, when Leuven blossomed into a wealthy and important city and became the capital of Brabant. Various sites, like the church, the town hall and the various markets still pay tribute to these heydays. In the late middle ages, the capital was moved to Brussels and trade started to dwindle. That’s when Leuven assumed its role as an attraction pole for both religious and worldly scholars. The university (the first in the low countries and the world’s oldest still existing Catholic university) was founded. We enter the days of Erasmus, Vesalius and Mercator. Our walk ends in the beautifully conserved Groot Begijnhof (Beguinage).

**Conference Dinner:** The conference dinner takes place in the Faculty Club restaurant in the Groot Begijnhof. Founded in the middle ages, Beguinages were communities of independent lay women supporting themselves by manual labour, or by teaching the children of burghers. They were a movement unique to the low countries. Leuven’s Beguinage, founded in 1234, is among the most well-preserved and is recognized as a UNESCO World Heritage site.



**Figure 3:** Faculty Club

Time	Friday (3/7) Morning	
09:00	<b>Invited Talk: Jure Leskovec</b> <i>Auditorium</i>	
10:00	<b>Coffee</b> <i>STUK Café</i>	
	<b>MLG</b> <i>Auditorium</i>	<b>ILP/SRL</b> <i>Verbeeck</i>
10:30	<i>Christoph Lippert, Nino Shervashidze and Oliver Stegle</i> <b>Relational Models for Generating Labeled Real-world Graphs</b>	<i>(S) Rodrigo de Salvo Braz, Sri- raam Natarajan, Hung Bui, Jude Shavlik and Stuart Russell</i> <b>Anytime Lifted Belief Propagation</b>
10:56	<i>Andreas Maunz, Christoph Helma and Stefan Kramer</i> <b>Large-Scale Graph Mining Using Backbone Refinement Classes</b>	<i>(S) Jacek Kisynski and David Poole</i> <b>Lifted Aggregation in Directed First-order Probabilistic Models</b>
11:22	<i>Galileo Mark Namata and Lise Getoor</i> <b>A Pipeline Approach to Graph Identification</b>	<i>(I) Wannes Meert, Jan Struyf and Hendrik Blockeel</i> <b>CP-Logic Theory Inference with Contextual Variable Elimination and Comparison to BDD Based Inference Methods</b>
11:48	<i>Juuso Parkkinen, Janne Sinkko- nen, Adam Gyenge and Samuel Kaski</i> <b>A Block Model Suitable for Sparse Graphs</b>	<i>(I) Lilyana Mihalkova and Matthew Richardson</i> <b>Speeding up Inference in Sta- tistical Relational Learning by Clustering Similar Query Literals</b>
12:14	<i>Christian Desrosiers and George Karypis</i> <b>Within-network Classification Using Local Structure Similarity</b>	<i>(S) Nima Taghipour, Wannes Meert, Jan Struyf and Hendrik Blockeel</i> <b>First-Order Bayes-Ball for CP- Logic</b>
12:40	<b>Lunch</b> <i>Stuk Café</i>	

Time	Friday (3/7) Afternoon
14:00	<p><b>Invited Talk: James Cussens</b> <i>Auditorium</i></p>
15:00	<p><b>Joint Poster Spotlight Session</b> <i>Auditorium</i></p> <p><i>Kamal Ali, Kevin Leung, Tolga Konik, Dongkyu Choi and Dan Shapiro</i> <b>Knowledge-Directed Theory Revision</b></p> <p><i>Laura-Andreea Antanas, Ingo Thon, Martijn van Otterlo, Niels Landwehr and Luc De Raedt</i> <b>Probabilistic Logical Sequence Models for Video</b> <i>Jianzhong Chen and Stephen Muggleton</i> <b>Decision-Theoretic Logic Programs</b> <i>Anton Dries and Luc De Raedt</i> <b>Towards Clausal Discovery for Stream Mining</b> <i>Daan Fierens</i> <b>On the Relationship between Logical Bayesian Networks and Probabilistic Logic Programming based on the Distribution Semantics</b> <i>Beatriz García Jiménez, Agapito Ledezma Espino and Araceli Sanchis de Miguel</i> <b>Modular Multi-Relational Framework for Gene Group Function Prediction</b> <i>Frank Hoonakker, Nicolas Lachiche, Alexandre Varnek and Alain Wagner</i> <b>Condensed Graph of Reaction: Considering a Chemical Reaction As one Single Pseudo Molecule</b> <i>Srihari Kalgi, Chirag Gosar, Ganesh Ramakrishnan and Ashwin Srinivasan</i> <b>Application of Theory of Optimal Search to ILP</b> <i>Hassan Khosravi, Oliver Schulte and Bahareh Bina</i> <b>Virtual Joins With Nonexistent Links</b> <i>Tolga Konik, Negin Nejati and Ugur Kuter</i> <b>Inductive Generalization of Analytically Learned Goal Hierarchies</b> <i>Mario Boley, Tamas Horvath, Axel Poigne and Stefan Wrobel</i> <b>New Results on Listing Closed Sets of Strongly Accessible Set Systems</b> <i>Charles Bouveyron, Hugh Chipman and Etienne Côme</i> <b>Supervised Classification and Visualization of Social Networks Based on a Probabilistic Latent Space Model</b></p>

## Conference Program

---

Time	Friday (3/7) Afternoon
	<p><i>Mickey Brautbar</i>  <b>Online Learning a Binary Labeling of a Graph</b>  <i>Thomas Cason, Pierre-Antoine Absil and Paul Van Dooren</i></p> <p><b>Iterative Methods for Low Rank Approximation of Graph Similarity Matrices</b>  <i>Christophe Costa Florencio, Jan Ramon, Jonny Daenen, Jan Van den Bussche and Dries Van Dyck</i></p> <p><b>Context-free Graph Grammars as a Language-bias Mechanism for Graph Pattern Mining</b>  <i>Christophe Costa Florencio</i></p> <p><b>Identification of Hyperedge-Replacement Graph Grammars</b>  <i>Ronen Brafman and Yagil Engel</i></p> <p><b>Lifted Optimization for Relational Preference Rules</b>  <i>Michael Chiang and David Poole</i></p> <p><b>Incremental EM for Latent-class Relational Models</b>  <i>James Cussens</i></p> <p><b>On Generative Parameterisations of Markov Logic Networks</b>  <i>Jan De Belder, Wim De Smet, Raquel Mochales Palau and Marie-Francine Moens</i></p> <p><b>Does Google own Youtube? Entity Relationship Extraction with Minimal Supervision</b>  <i>Arjen Hommersom, Nivea Ferreira and Peter Lucas</i></p> <p><b>Learning Parameters of Chain Logic Theories</b>  <i>Kristian Kersting and Zhao Xu</i></p> <p><b>Relational Gaussian Processes for Learning Preference Relations</b></p>
16:00	<p><b>Coffee followed by poster reception</b>  <i>STUK Café and Ensemble room</i></p>
19:30	<p><b>Brewery visit</b></p>
21:30	<p><b>Dinner on your own / Beleuvenissen (Jazz)</b></p>

**Stella Artois Brewery:** The walk to the Stella brewery takes approximately 20 minutes (see the map in the middle of this booklet). Alternatively, you can take bus 1 or 2 till the railway station, from where the walk is 5 minutes. During the visit to the brewery, we will find out all the secrets of the beer brewing process. We will visit the brewing hall, the filtration and the filling plant, where we can witness how bottles, cans and kegs are filled at an enormous speed. The visit will end with some beer tasting, where we can try out different brands produced by ABInBev, the world's largest beer brewer.

Time	Saturday (4/7) Morning	
09:00	<b>Invited Talk: Scott Sanner</b> <i>Auditorium</i>	
10:00	<b>Coffee</b> <i>STUK Café</i>	
	<b>ILP/SRL</b> <i>Auditorium</i>	<b>ILP</b> <i>Verbeeck</i>
10:30	<i>(I)</i> Lisa Torrey and Jude Shavlik Policy Transfer via Markov Logic Networks	Joris Gillis and Jan Van den Bussche Induction of Relational Algebra Expressions
10:56	<i>(I)</i> Taisuke Sato, Masakazu Ishihata and Katsumi Inoue Constraint-based Probabilistic Modeling for Statistical Abduction	Jens Lehmann and Christophe Haase Ideal Downward Refinement in the EL Description Logic
11:22	<i>(S)</i> Matthias Broecheler and Lise Getoor Probabilistic Similarity Logic	Stephen Muggleton, Jose Santos and Alireza Tamaddon-Nezhad Towards a System Based on Asymmetric Relative Minimal Generalisation
11:48	<i>(S)</i> Aniruddh Nath and Pedro Domingos A Language for Relational Decision Theory	Houssam Nassif, Hassan Al-Ali, Sawsan Khuri, Walid Keirouz and David Page An ILP Approach to Model and Classify Hexose Binding Sites
12:14	<i>(S)</i> Volker Tresp, Yi Huang, Markus Bundschuh and Achim Rettinger Materializing and Querying Learned Knowledge	Tamas Horvath, Gerhard Paass, Frank Reichartz and Stefan Wrobel A Logic-Based Approach to Relation Extraction from Texts
12:40	<b>Lunch</b> <i>Stuk Café</i>	

## Sponsors

---

Time	Saturday (4/7) Afternoon
14:00	<b>Invited Talk: Jason Eisner</b> <i>Auditorium</i>
15:00	<b>Panel</b> <i>Auditorium</i>
16:00	<b>Business Meeting</b> <i>Auditorium</i>
17:00	<b>Closing</b> <i>Auditorium</i>

## SPONSORS



Belgium-Dutch Association for  
Artificial Intelligence



Science Foundation Flanders



Network of Excellence "Pascal"

# ORGANIZATION

<i>General Chair</i>	Luc De Raedt
<i>ILP Program Chair</i>	Luc De Raedt
<i>MLG Program Chairs</i>	Hendrik Blockeel Karsten Borgwardt Xifeng Yan
<i>SRL Program Chairs</i>	Pedro Domingos Kristian Kersting

## **Organizing team:**

<i>Submissions</i>	Jan Struyf
<i>Registration, Hotels, and Conference Materials</i>	Siegfried Nijssen Martijn van Otterlo Joaquin Vanschoren
<i>Webmaster</i>	Fabrizio Costa Christophe Costa Florencio
<i>Publicity and Proceedings</i>	Kurt Driessens
<i>Logistics</i>	Daan Fierens
<i>Catering and Social Events</i>	Celine Vens
<i>Administration</i>	Karin Michiels
<i>Sponsoring</i>	Karsten Borgwardt Jan Ramon
<i>Volunteers</i>	Laura Antanas, Eduardo Costa, Tom Croonenborghs, Anton Dries, Robby Goetschalckx, Fabián Güiza, Angelika Kimmig, Parisa KordJamshidi, Theofrastos Mantadelis, Wannes Meert, Dhany Saputra, Leander Schietgat, Nima Taghipour, Ingo Thon, Vinicius Tragante do O., Bernd Gutmann

**ILP**

Akihiro Yamamoto  
 Alan Fern  
 Annalisa Appice  
 Bernhard Pfahringer  
 Celine Rouveinol  
 Christel Vrain  
 David Page  
 David Poole  
 Donato Malerba  
 Filip Zelezny  
 Francesca Lisi  
 Gerson Zaverucha  
 James Cussens  
 John Lloyd  
 Jude Shavlik  
 Katsumi Inoue  
 Kristian Kersting  
 Kurt Driessens  
 Lise Getoor  
 Nada Lavrac  
 Niels Landwehr  
 Paolo Frasconi  
 Peter Flach  
 Prasad Tadepalli  
 Ramon Otero  
 Raymond Mooney  
 Ross King  
 Saso Dzeroski  
 Stan Matwin  
 Stefan Kramer  
 Stefan Wrobel  
 Stephen Muggleton  
 Taisuke Sato  
 Tamas Horvath  
 Vitor Costa Santos

**MLG**

Alan Fern  
 Andrea Passerini  
 Andrew Goldberg  
 Christian Borgelt  
 Craig Saunders  
 Edo Airoidi  
 Fabrizio Costa  
 Francesco Bonchi  
 Gemma Garriga  
 George Karypis  
 Ingrid Fischer  
 Janne Sinkkonen  
 Jan Ramon  
 Jean-Philippe Vert  
 Joe Blitzstein  
 Jose Balcazar  
 Jure Leskovec  
 Koji Tsuda  
 Kristiaan Pelckmans  
 Luc De Raedt  
 Manfred Jaeger  
 Marco Gori  
 Mehryar Mohri  
 Nicos Angelopoulos  
 Pierre Dupont  
 Samuel Kaski  
 Siegfried Nijssen  
 Spiros Papadimitriou  
 Stefan Kramer  
 Stefan Wrobel  
 Susanne Hoche  
 Takashi Washio  
 Takeaki Uno  
 Tamas Horvath  
 Tanya Berger-Wolf  
 Tapani Raiko  
 Thomas Gaertner  
 Thorsten Meinl  
 Tsuyoshi Murata  
 Ulf Brefeld  
 Volker Tresp  
 William Stafford Noble  
 Yasemin Altun  
 Yun Chi

**SRL**

Alan Fern  
 Angelika Kimmig  
 Bhaskara Marthi  
 Brian Milch  
 Christian Plagemann  
 David Page  
 David Poole  
 Edo Airoidi  
 Foster Provost  
 Hendrik Blockeel  
 James Cussens  
 Jennifer Neville  
 John Lloyd  
 Kai Yu  
 Kee Siong Ng  
 Lise Getoor  
 Luc De Raedt  
 Luke Zettlemoyer  
 Lyle Ungar  
 Manfred Jaeger  
 Niels Landwehr  
 Noah Goodman  
 Paolo Frasconi  
 Prasad Tadepalli  
 Ricardo Silva  
 Rodrigo de Salvo Braz  
 Scott Sanner  
 Sofus Attila Macskassy  
 Sriraam Natarajan  
 Stephen Muggleton  
 Taisuke Sato  
 Vitor Santos Costa  
 Zhao Xu



# ABOUT LEUVEN

Although the first references to the town can be traced back as far as the 9th century and in spite of its strategic location on the river Dyle, it was not until around the 11th-12th century that Leuven began to develop as an important trading centre within the Duchy of Brabant. It was at this time that its first town wall, churches, monasteries and abbeys were built.

The town's once flourishing cloth trade had fallen into decline by the 14th century, but a new golden age dawned with the 15th century. The university was founded, various industrial concerns flourished and the main market square with its fine Gothic town hall was laid out.

Its finely sculpted Town Hall and Saint Peter's church with its magnificent interior which is hung with the famous 'Last Supper' by Thierry Bouts, are fine examples of flamboyant Gothic architecture at its best. Dating from the same period are the Cloth Hall, 't Sestich House and many churches, including St Gertrude's with its magnificent choir stalls.

Other architectural styles are not quite so much in evidence, but they can still be found. The Romanesque Gate, the Saint Lambert's church and the remains of the first town walls are still clearly in evidence.

The 13th century beguinage complex has cobbled streets and fascinating 17th century houses.

The Baroque style abounds in Leuven, from St Michael's church and the interior of the St John the Baptist's church to the unrivalled stucco of the Park Abbey. Both the Park Abbey and the Arenberg Castle are exciting mixtures of architectural styles.

The town still owes much of its character to the numerous university college buildings dating from the 16th and 17th centuries, many of which were renovated in the 18th century. The Van Dale College is pure Renaissance style and the College of the Falcon, Pope's College and the Arras College are neo-classical style. Vlierbeek Abbey is a fine example of this trend.

The building of paved roads, the canal in the 18th century and the expansion of the brewery created a great impetus for trade and population growth. Unfortunately, this prosperity came to a sudden end under the reign of Joseph II when the French invaded.

The two World Wars in the 20th century destroyed much of Leuven and a great deal of the reconstruction work lasted well into the nineteen-sixties.

# INVITED SPEAKERS

## **A Tutorial on Logic-Based Approaches to SRL**

*by James Cussens (University of York)*

The relations in Statistical Relational Learning are often expressed using first-order logic, leading to formalisms which combine both logical and probabilistic representations. In this talk I intend to explain the most important consequences of adopting a logical approach to SRL. Defining distributions over 'possible worlds' is a common theme to many such approaches. Two prominent logic-based formalisms - Markov logic networks and PRISM programs - will be used as exemplars. Although the talk is tutorial in nature, I hope to make it interesting to those already familiar with this area!

## **Weighted Deduction as an Abstraction Level for AI**

*by Jason Eisner (Johns Hopkins University)*

The field of AI has become implementation-bound. We have plenty of ideas, but it is increasingly laborious to try them out, as our models become more ambitious and our datasets become larger, noisier, and more heterogeneous. The software engineering burden makes it hard to start new work; hard to reuse and combine existing ideas; and hard to educate our students.

In this talk, I'll propose to hide many common implementation details behind a new level of abstraction that we are developing. Dyna is a declarative programming language that combines logic programming with functional programming. It also supports modularity. It may be regarded as a kind of deductive database, theorem prover, truth maintenance system, or equation solver.

I will illustrate how Dyna makes it easy to specify the combinatorial structure of typical computations needed in natural language processing, machine learning, and elsewhere in AI. Then I will sketch implementation strategies and program transformations that can help to make these computations fast and memory-efficient. Finally, I will suggest that machine learning should be used to search for the right strategies for a program on a particular workload.

**Networks, clusters and Kronecker products***by Jure Leskovec (Cornell University)*

Emergence of the web and online computing applications gave rise to rich large scale social activity data. One of the principal challenges then is to build models and understanding of the structure of such large social and information networks. In this talk I will present our work on the clustering/community structure in large networks, where clusters are thought of as sets of nodes that are better connected internally than to the rest of the network. We find that large networks have very different clustering structure from well studied small social networks and graphs that are well-embeddable in a low-dimensional structure. In networks of millions of nodes tight clusters exist at only very small size scales up to around 100 nodes, while at large size scales networks becomes expander like. A network model based on Kronecker products efficiently models such core-periphery network structures. The results suggest broader implications for data analysis in sparse and noisy high-dimensional social and information networks, where intuitive notions about cluster quality fail, and thus traditional clustering methods which rely on “nice” statistical assumptions are inappropriate.

**Bottom-up Search and Transfer Learning in SRL***by Raymond Mooney (University of Texas at Austin)*

This talk addresses two important issues motivated by of our recent research in SRL. First, is the value of data-driven, “bottom-up” search in learning the structure of SRL models. Bottom-up induction has a long history in traditional ILP; however, its use in SRL has been somewhat limited. We review recent results on several structure-learning methods for Markov Logic Networks (MLNs) that highlight the value of bottom-up search. Second, is the value of transfer learning in reducing the data and computational demands of SRL. By inducing a predicate mapping between seemingly disparate domains, effective SRL models can be efficiently learned from very small amounts of in-domain training data. For example, by transferring a model learned from data about a CS department, we have induced reasonably accurate models for IMDB movie data given training data about only a single person.

**First-order Models for Sequential Decision-making: Insights, Caveats, and Tricks-of-the-Trade***by Scott Sanner (NICTA)*

In this talk I will discuss first-order models and algorithms for sequential decision-making, specifically those approaches that admit exact lifted solutions. The first emphasis of the talk will be on the insights that underlie these models and algorithms along with potential caveats for their practical application. The second emphasis of the talk will be on a variety of extensions of the first-order Markov decision process (MDP) framework such as the fac-

tored first-order MDP and the first-order partially observable MDP. The third emphasis of the talk will be on the algorithmic tricks-of-the-trade that allow the practical application of these models; this includes (a) useful data structures, (b) efficient solution techniques for first-order linear programs, (c) new techniques for first-order variable elimination, and (d) practical methods for maintaining compact, consistent first-order representations without theorem proving.

### **Scalable Link Mining and Analysis on Information Networks**

*by Philip S. Yu (University of Illinois at Chicago)*

With the ubiquity of information networks and their broad applications, there have been numerous studies on the construction, online analytical processing, and mining of information networks in multiple disciplines, including social network analysis, World-Wide Web, database systems, data mining, machine learning, and networked communication and information systems. Algorithms like PageRank and HITS have been developed in late 1990s to explore links among Web pages to discover authoritative pages and hubs. Links have also been popularly used in citation analysis and social network analysis. However, there is a lack of systematic treatment on how to fully explore the power of links in scalable data analysis. In this talk, the power of links are examined in details to improve the effectiveness and efficiency of typical data analysis tasks, including information integration, on-line analytic processing, and other interesting data mining tasks, especially in the multi-relational databases and/or the World-Wid e Web environments.

# ABSTRACTS ILP

## Oral Presentation

### **Relational Random Forests based on Random Relational Rules**

*Grant Anderson and Bernhard Pfahringer*

Random Forests have been shown to perform very well in propositional learning. FORF is an upgrade of Random Forests for relational data. In this paper we investigate shortcomings of FORF and propose an alternative algorithm, R4F, for generating Random Forests over relational data. R4F employs randomly generated relational rules as fully self-contained Boolean tests inside each node in a tree and thus can be viewed as an instance of dynamic propositionalization. The implementation of R4F allows for the simultaneous or parallel growth of all the branches of all the trees in the ensemble in an efficient shared, but still single-threaded way. Experiments favorably compare R4F to both FORF and the combination of static propositionalization together with standard Random Forests. Various strategies for tree initialization and splitting of nodes, as well as resulting ensemble size, diversity, and computational complexity of R4F are also investigated. A longer version of this paper has been accepted by IJCAI2009.

### **Induction of Relational Algebra Expressions**

*Joris Gillis and Jan Van den Bussche*

We consider the induction of relational algebra expressions from examples consisting of a relational database and an output relation. This problem fits perfectly in the ILP context but has not been investigated in specific detail. We discuss the importance of negation (difference, complementation, universal quantification), propose a new heuristic to take complementation into account, and propose the use of cylindrical set algebra as a more flexible form for searching expressions. We present some modest experimental results which nevertheless show that our approach allows the induction of typical relational database queries involving universal quantification, such as Codd's relational division operator.

### **A Logic-Based Approach to Relation Extraction from Texts**

*Tamas Horvath, Gerhard Paass, Frank Reichartz and Stefan Wrobel*

In recent years, text mining has moved far beyond the classical problem of text classification with an increased interest in more sophisticated processing of large text corpora, such as, for example, evaluations of complex queries. This and several other tasks are based on the essential step of relation extraction. This problem becomes a typical application of learning logic programs by considering the dependency trees of sentences as relational structures and examples of the target relation as ground facts of a target

predicate. In this way, each example is represented by a definite first-order Horn-clause. We show that Plotkin's LGG operator can effectively be applied to such clauses and propose a simple and effective divide-and-conquer algorithm for listing a certain set of LGGs. We use these LGGs as binary features and obtain the final hypothesis by applying SVM to the corresponding feature vectors. Empirical results on the ACE-2003 benchmark dataset indicate that the performance of our approach is comparable with state-of-the-art kernel methods.

### **Abducing Rules with Predicate Invention**

*Katsumi Inoue, Koichi Furukawa and Ikuo Kobayashi*

This paper addresses discovery of unknown relations from incomplete network data using abduction. Given a network information such as causal relations and metabolic pathways, we want to infer missing links and nodes in the network to account for observations. This is implemented in SOLAR, an automated deduction system for consequence finding, using first-order representation of algebraic relations and full-clausal ground formulas of network information. Abduction by SOLAR is powerful enough to infer unknown rules and to realize predicate invention by inferring unknown causes. In particular, we point out the importance of existentially quantified formulas to express hypotheses including new variables representing missing nodes to be introduced. We also show an application of rule abduction to discover knack in performing skillful tasks.

### **Block-Wise Construction of Acyclic Relational Features with Monotone Relevancy**

*Ondřej Kuželka and Filip Zelezny*

We describe an algorithm for constructing a set of acyclic conjunctive relational features by combining smaller conjunctive blocks. Unlike traditional level-wise approaches which preserve the monotonicity of frequency, our block-wise approach preserves a form of monotonicity of relevancy, which is important in propositionalization employed in the context of classification learning. With pruning based on this property, our block-wise approach efficiently scales to features including tens of first-order literals, far beyond the reach of state-of-the-art propositionalization or inductive logic programming systems.

### **Ideal Downward Refinement in the EL Description Logic**

*Jens Lehmann and Christophe Haase*

With the proliferation of the Semantic Web, there has been a rapidly rising interest in description logics, which form the logical foundation of the W3C standard ontology language OWL. While the number of OWL knowledge bases grows, there is an increasing demand for tools assisting knowledge engineers in building up and maintaining their structure. For this purpose, concept learning algorithms based on refinement operators have been investigated. In this paper, we provide an ideal refinement operator

for the popular description logic EL and show that it is computationally feasible on large knowledge bases.

### **Relational Information Gain**

*Marco Lippi, Manfred Jaeger, Paolo Frasconi and Andrea Passerini*

We introduce relational information gain, a refinement scoring function measuring the informativeness of newly introduced variables. The gain can be interpreted as a conditional entropy in a well-defined sense and can be efficiently approximately computed. In conjunction with simple greedy general-to-specific search algorithms such as FOIL, it yields an efficient and competitive algorithm in terms of predictive accuracy, compactness of the learned theory, and ability to learn a good structure for probabilistic logical models.

### **CP-Logic Theory Inference with Contextual Variable Elimination and Comparison to BDD Based Inference Methods**

*Wannes Meert, Jan Struyf and Hendrik Blockeel*

There is a growing interest in languages that combine probabilistic models with logic to represent complex domains involving uncertainty. Causal Probabilistic Logic (CP-logic) is such a probabilistic logic language, which has been designed to model causal processes. This paper investigates inference algorithms for CP-logic; these are crucial for developing learning algorithms. It proposes a new CP-logic inference method based on contextual variable elimination and compares this method to variable elimination and to inference methods based on binary decision diagrams.

### **Speeding up Inference in Statistical Relational Learning by Clustering Similar Query Literals**

*Lilyana Mihalkova and Matthew Richardson*

Markov logic networks (MLNs) have been successfully applied to several challenging problems by taking a "programming language" approach where a set of formulas is hand-coded and weights are learned from data. Because inference plays an important role in this process, "programming" with an MLN would be significantly facilitated by speeding up inference. We present a new meta-inference algorithm that exploits the repeated structure frequently present in relational domains to speed up existing inference techniques. Our approach first clusters the query literals and then performs full inference for only one representative from each cluster. The clustering step incurs only a one-time up-front cost when weights are learned over a fixed structure.

### **Towards a System Based on Asymmetric Relative Minimal Generalisation**

*Stephen Muggleton, Jose Santos and Alireza Tamaddoni-Nezhad*

Over the last decade Inductive Logic Programming systems have been dominated by use of top-down refinement search techniques. In this paper we re-examine the use of bottom-up approaches to the construction of logic programs. In particular, we explore an asymmetric variant of Plotkin's Relative Least General Generalisation (RLGG) which

is based on subsumption relative to a bottom clause. With Plotkin's RLGG, clause length grows exponentially in the number of examples. By contrast, in the Golem system, the length of ij-determinate RLGG clauses were shown to be polynomially bounded for given values of i and j. However, the determinacy restrictions made Golem inapplicable in many key application areas, including the learning of chemical properties from atom and bond descriptions. In this paper we show that with Asymmetric Relative Minimal Generalisations (or ARMGs) relative to a bottom clause, clause length is bounded by the length of the initial bottom clause. ARMGs, therefore do not need the determinacy restrictions used in Golem but still have a polynomial-time construction. An algorithm is described for constructing ARMGs together with some initial results on the clause lengths for some well-known ILP applications. In a longer follow-up paper we aim to extend the analysis and implementation of a system called ProGolem which combines bottom-clause construction in Progol with a Golem control strategy for constructing multiple clause predicate definitions.

### **An ILP Approach to Model and Classify Hexose Binding Sites**

*Houssam Nassif, Hassan Al-Ali, Sawsan Khuri, Walid Keirouz and David Page*

Hexoses play a key role in many cellular pathways and in the regulation of development and disease mechanisms. As such, hexose binding proteins and their binding properties are of great interest to biomedical researchers. Current protein-sugar computational models are based, at least partially, on prior biochemical knowledge. We investigate the empirical support for biochemical findings by comparing ILP-induced rules to actual biochemical results. Our method matches lab findings, is able to identify hexose binding sites, and in addition reveals a Trp-Glu dependency. Our classifier achieves a similar accuracy as other black-box classifiers while providing insight into the discriminating process.

### **Boosting First-Order Clauses for Large, Skewed Data sets**

*Louis Oliphant, Elizabeth Burnside and Jude Shavlik*

Creating an effective ensemble of clauses for large, skewed data sets requires finding a diverse, high-scoring set of clauses and then combining them in such a way as to maximize predictive accuracy. We have adapted the RankBoost algorithm in order to maximize area under the recall-precision curve, a much better metric when working with highly skewed data sets than ROC curves. We have also explored a range of possibilities for the weak hypotheses used by our modified RankBoost algorithm beyond using individual clauses. We provide results on four large, skewed data sets showing that our modified RankBoost algorithm outperforms the original on area under the recall-precision curves.

### **Constraint-based Probabilistic Modeling for Statistical Abduction**

*Taisuke Sato, Masakazu Ishihata and Katsumi Inoue*

We introduce a new framework for logic-based probabilistic modeling called constraint-based probabilistic modeling which defines conditional distributions  $P(\text{---} | \text{KB})$  over



independent propositional variables constrained by a first-order knowledge base KB. We prove generative models such as PCFGs and discriminative models such as CRFs have an equivalent CBPM (constraint-based probabilistic model) as long as they are discrete. We apply this generality to abductive reasoning with KB allowing cyclic dependencies and derive an EM algorithm for the parameter learning of CBPMs.

### **Policy Transfer via Markov Logic Networks**

*Lisa Torrey and Jude Shavlik*

We propose an improved method for transfer in reinforcement learning via Markov Logic Networks. Our goal is to extract relational knowledge from a source task and use it to speed up learning in a related target task. We do so by learning a Markov Logic Network that describes the source-task policy, and then using it for decision making in the early learning stages of the target task. Through experiments in the RoboCup simulated-soccer domain, we show that this approach performs better than a previous MLN transfer method that modeled a value function rather than a policy.

## **Poster Presentation**

### **Learning Responsive Cellular Networks by Integrating Genomics and Proteomics Data**

*Elma Akand, Michael Bain and Mark Temple*

Systems biology is an important application area for ILP. We investigate learning networks of activity in response to stress in yeast cells. Diverse heterogeneous empirical data is used rather than curated knowledge sources. We find that although this representation is restricted and predictive accuracy is low this can still lead to useful outputs.

### **Knowledge-Directed Theory Revision**

*Kamal Ali, Kevin Leung, Tolga Konik, Dongkyu Choi and Dan Shapiro*

Using domain knowledge to speed up learning is widely accepted but theory revision of such knowledge continues to use general syntactic operators. Using such operators for theory revision of teleoreactive logic programs is especially expensive since such programs are often used to control games in which proof of a top-level goal involves playing a game. In such contexts, one should have the option to complement general theory revision with domain-specific knowledge. We have found it easy to extract and express such knowledge. Using American football as an example, we use ICARUS's multiplayer teleoreactive logic programming ability to encode a coach agent which infers faults during a game and at its conclusion applies procedural attachments to fix the programs of other agents. Our results show effective learning using as few as twenty training examples. We also show that structural changes made by such revision

can produce performance gains that cannot be matched using algorithms that do not make structural changes.

### **Probabilistic Logical Sequence Models for Video**

*Laura-Andreea Antanas, Ingo Thon, Martijn van Otterlo, Niels Landwehr and Luc De Raedt*

Understanding complex dynamic scenes of real-world activities from low-level sensor data is of central importance for intelligent systems. The main difficulty lies in the fact that complex scenes are best described in high-level, logical formalisms, while sensor data usually consists of many low-level features. We first propose a method to obtain a logical representation of real-world dynamic scenes based on input video stream solely. We focus on representing the video data using probabilistic relational sequences as a natural way to incorporate sensor uncertainty. They allow us to work with structured terms, and in addition they capture the inherent uncertainty of object detection. Further on, we employ r-grams as the probabilistic logical learning model for this application. In a first step we use r-grams in a simple setting and we show their viability in card games. We then show how r-grams can be upgraded to deal with uncertain observations.

### **Decision-Theoretic Logic Programs**

*Jianzhong Chen and Stephen Muggleton*

We propose a new Probabilistic ILP (PILP) framework, Decision-theoretic Logic Programs (DTLPs), in the paper. DTLPs extend PILP models by integrating decision-making features developed in Statistical Decision Theory area. Both decision-theoretic knowledge (e.g. utilities) and probabilistic knowledge (e.g. probabilities) can be represented and dealt with in DTLPs. An implementation of DTLPs using Stochastic Logic Programs (SLPs) is introduced and a DTLP parameter learning algorithm is discussed accordingly. The representation and methods are tested by performing regression on the traditional mutagenesis dataset.

### **Towards Clausal Discovery for Stream Mining**

*Anton Dries and Luc De Raedt*

With the increasing popularity of data streams it has become time to adapt logical and relational learning techniques for dealing with streams. In this note, we present our preliminary results on upgrading the clausal discovery paradigm towards the mining of streams. In this setting, there is a stream of interpretations and the goal is to learn a clausal theory that is satisfied by these interpretations. Furthermore, in data streams the interpretations can be read (and processed) only once.

### **On the Relationship between Logical Bayesian Networks and Probabilistic Logic Programming based on the Distribution Semantics**

*Daan Fierens*

A significant part of current research on ILP deals with probabilistic logical models.

Over the last decade many logics or languages for representing such models have been introduced. There is currently a great need for insight into the relationships between all these languages. One class of languages are those that extend probabilistic models with elements of logic, such as in the language of Logical Bayesian Networks (LBNs). Some other languages follow the converse strategy of extending logic programs with a probabilistic semantics, often in a way similar to that of Sato's distribution semantics. In this paper we study the relationship between the language of LBNs and languages based on the distribution semantics. Concretely, we define a mapping from LBNs to theories in the Independent Choice Logic (ICL). We also show that this mapping provides us with a learning algorithm for ICL.

### **Modular Multi-Relational Framework for Gene Group Function Prediction**

*Beatriz García Jiménez, Agapito Ledezma Espino and Araceli Sanchis de Miguel*

Determining the functions of genes is essential for understanding how the metabolisms work, and for trying to solve their malfunctions. Genes usually work in groups rather than isolated, so functions should be assigned to gene groups and not to individual genes. Moreover, the genetic knowledge has many relations and is very frequently changeable. Thus, a propositional ad-hoc approach is not appropriate to deal with the gene group function prediction domain. We propose the Modular Multi-Relational Framework (MMRF), which faces the problem from a relational and flexible point of view. The MMRF consists of several modules covering all involved domain tasks (grouping, representing and learning using computational prediction techniques). A specific application is described, including a relational representation language, where each module of MMRF is individually instantiated and refined for obtaining a prediction under specific given conditions.

### **Condensed Graph of Reaction: Considering a Chemical Reaction As one Single Pseudo Molecule**

*Frank Hoonakker, Nicolas Lachiche, Alexandre Varnek and Alain Wagner*

Chemical reactions always involve several molecules of two types – reactants and products. On the other hand, Quantitative Structure Activity Relationship (QSAR) methods are used to predict physic-chemical or biological properties of individual molecules. In this article, we propose to use Condensed Graph of Reaction (CGR) approach merging all molecules involved in a reaction into one molecular graph. This allows one to consider reactions as pseudo-molecules and to develop QSAR models based on fragment descriptors. Here Substructure Molecular Fragment descriptors calculated from CGRs have been used to build quantitative models for the rate constant of SN2 reactions in water. Three common attribute-value regression algorithms (linear regression, support vector machine, and regression trees) have been evaluated.

### **Application of Theory of Optimal Search to ILP**

*Srihari Kalgi, Chirag Gosar, Ganesh Ramakrishnan and Ashwin Srinivasan*

Most ILP systems attempt to find a hypothesis by searching through a space of possible

alternatives, the principal problem is that the size of the search space makes it impossible to perform even a complete (but non-exhaustive) search. Hence, greedy approaches such as FOIL and restricted variants of branch and bound techniques (Aleph), etc. are used to restrict the search space. In this paper, we will concern ourselves principally with systems like Progol, in which search proceeds by examining subspaces constrained by a set of most-specific clause  $\perp_1, \perp_2, \dots, \perp_n$  (where  $n$  here will be the number of positive examples) using a greedy covering procedure. In each iteration of the greedy procedure a Progol-like system selects a positive example  $e_i$ , construct a bottom clause  $\perp_i$  using the example and the background knowledge  $B$  and then searches for the best clause  $h_i$  in a lattice  $\text{lat}(\perp_i)$  subsuming  $\perp_i$ . Here "best" is in terms of some objective function (like pos-neg). Once this best clause  $h_i$  in the lattice  $\text{lat}(\perp_i)$  is found, all positive examples made redundant by  $B \wedge h_i$  are removed from further consideration; and the search continues. A number of questions still remain, for example: How is the  $e_i$  to be selected? How much effort should we expend on any one search? Current ILP systems answer these in ad hoc ways. It is our aim to show how a theory of optimal search can be applied to answer such questions.

### **Virtual Joins With Nonexistent Links**

*Hassan Khosravi, Oliver Schulte and Bahareh Bina*

Many approaches to multi-relational learning require the computation of database frequencies in the presence of nonexistent links. The corresponding ILP problem is to compute the number of groundings that satisfy a given conjunction of literals in a relational database, where one or more of the literals is negated. We present a fast new dynamic programming algorithm for this problem. The database table joins performed by our algorithm are restricted to joins of tables already existing in the database. Evaluation on three data sets confirms the efficiency of our algorithm; computing frequencies for negated literals added about 15% to the cost of computing frequencies for positive literals only.

### **Inductive Generalization of Analytically Learned Goal Hierarchies**

*Tolga Konik, Negin Nejati and Ugur Kuter*

We describe a new approach for learning procedural knowledge represented as Teleoreactive Logic Programs (TLPs) using relational behavior traces as input. A TLP organizes task decomposition skills hierarchically and attaches them explicitly defined goals. Our approach integrates analytical learning with inductive generalization in order to learn these skills. The analytical component predicts the goal dependencies given a successful behavior trace and produces the structure of the skill hierarchy. The inductive component generates successful and failed traces using those skills and refines their applicability conditions with an ILP component. We claim that explicit goal representation of TLPs allows the learning system to determine positive and negative examples used in learning. We claim that our approach uses less expert effort compared to a purely inductive approach and may perform better compared to a purely analytical approach.

**A Propositionalisation that Preserves More Continuous Attribute Domains***Julien Lesbegueries, Agnès Braud and Nicolas Lachiche*

The aim of this paper is to extend local transformation functions for propositionalisation. We follow the works of (Kroegel et al., 2001) by adding a transformation function that reverses the thresholding problem. Indeed, we propose an innovative method that manages numeric attributes with the hypothesis that features can be transformed without use of discretisation nor aggregation.

**Nonmonotonic Onto-Relational Learning***Francesca Alessandra Lisi and Floriana Esposito*

In this paper we carry on the work on Onto-Relational Learning by investigating the impact of having disjunctive Datalog with default negation either in the language of hypotheses or in the language for the background theory. The inclusion of nonmonotonic features strengthens the ability of our ILP framework to deal with incomplete knowledge by performing some form of commonsense reasoning. One such ability can turn out to be useful in application domains, such as the Semantic Web, that require that kind of reasoning.

**Chess Revision: Acquiring the Rules of Chess Variants through FOL Theory Revision from Examples***Stephen Muggleton, Aline Paes, Vítor Santos Costa and Gerson Zaverucha*

The game of chess has been a major testbed for research in artificial intelligence since it requires focus on intelligent reasoning. Particularly, several challenges arise to machine learning systems when inducing a model describing the rules of chess, including the creation and definition of the examples, the learning of a model which correctly represents the official rules of the game, covering all the branches and restrictions of the correct moves, and the comprehensibility of such a model. Besides, the game of chess has inspired the creation of numerous variants, ranging from faster to more challenging or to regional versions of the game. The question arises if it is possible to take advantage of an initial classifier of chess as a starting point to obtain classifiers for the different variants. We approach this problem as an instance of theory revision from examples. The initial classifier of chess is inspired by a FOL theory approved by a chess expert and the examples are defined as sequences of moves within a game. Starting from a standard revision system, we argue that abduction and negation are also required to best address this problem. Experimental results show the effectiveness of our approach.

**ILP for Bootstrapped Learning: A Layered Approach to Automating the ILP Set-Up Problem***Sriram Natarajan, Gautam Kunapuli, Ciaran O'Reilly, Richard Maclin, Trevor Walker, David Page and Jude Shavlik*

This paper introduces a new type of application for ILP called Bootstrapped Learning

(BL). BL brings several challenges to ILP, including the need to automate the “ILP Set-Up” problem; small numbers of examples, in some cases no negative examples; and the need to “bootstrap” or to automatically base learning in part on the results of earlier learning sessions. The paper introduces BL, discusses the general challenges it raises for ILP, and presents initial results in ILP for BL.

### **Higher-order Logic Learning**

*Niels Pahlavi and Stephen Muggleton*

This paper presents Higher-order Logic Learning (HOLL), which consists of generalizing logic-based Machine Learning, and more particularly Inductive Logic Programming (ILP), from the first-order to the higher-order context. The power of expressivity of Higher-order Logic (HOL) is used to improve significantly the learning capacity and efficiency of logic-based Machine Learning – by both allowing to learn new tasks but also to perform better than first-order Machine Learning systems on some already learnable problems. We describe a Higher-order ILP system, called  $\lambda$ Progol, adapting the ILP system Progol and based on the HOL formalism  $\lambda$ Prolog, along with its first implementation and promising results on motivating worked examples. We intend to extend the implementation, tests and evaluation of  $\lambda$ Progol further, and to develop a theory of HOLL.

### **Incorporating Linguistic Expertise Using ILP for Named Entity Recognition in Data Hungry Indian Languages**

*Anup Patel, Ganesh Ramakrishnan and Pushpak Bhattacharya*

Developing linguistically sound and data-compliant rules for named entity annotation is usually an intensive and time consuming process for any developer or linguist. In this work, we present the use of two Inductive Logic Programming (ILP) techniques to construct rules for extracting instances of various named entity classes thereby reducing the efforts of a linguist/developer. Using ILP for rule development not only reduces the amount of effort required but also provides an interactive framework wherein a linguist can incorporate his intuition about named entities such as in form of mode declarations for refinements (suitably exposed for ease of use by the linguist) and the background knowledge (in the form of linguistic resources). We have a small amount of tagged data - approximately 3884 sentences for Marathi and 22748 sentences in Hindi. The paucity of tagged data for Indian languages makes manual development of rules more challenging. However, the ability to fold in background knowledge and domain expertise in ILP techniques comes to our rescue and we have been able to develop rules that are mostly linguistically sound that yield results comparable to rules hand-crafted by linguists. The ILP approach has two advantages over the approach of hand-crafting all rules: (i) the development time reduces by a factor of 240 when ILP is used instead of involving a linguist for the entire rule development and (ii) the ILP technique has the computational edge that it has a complete and consistent view of all significant patterns in the data at the level of abstraction specified through the mode declarations. The point (ii) enables the discovery of rules that could be missed out by

the linguist and also makes it possible to scale the rule development to a larger training dataset. The rules thus developed could be optionally edited by linguistic experts and consolidated either (a) through default ordering (as in TILDE[1]) or (b) with an ordering induced using [2] or (c) by using the rules as features in a statistical graphical model such a conditional random field (CRF) [3]. We report results using WARMR [4] and TILDE to learn rules for named entities of Indian languages namely Hindi and Marathi.

### **Transfer Learning via Relational Templates**

*Scott Proper and Prasad Tadepalli*

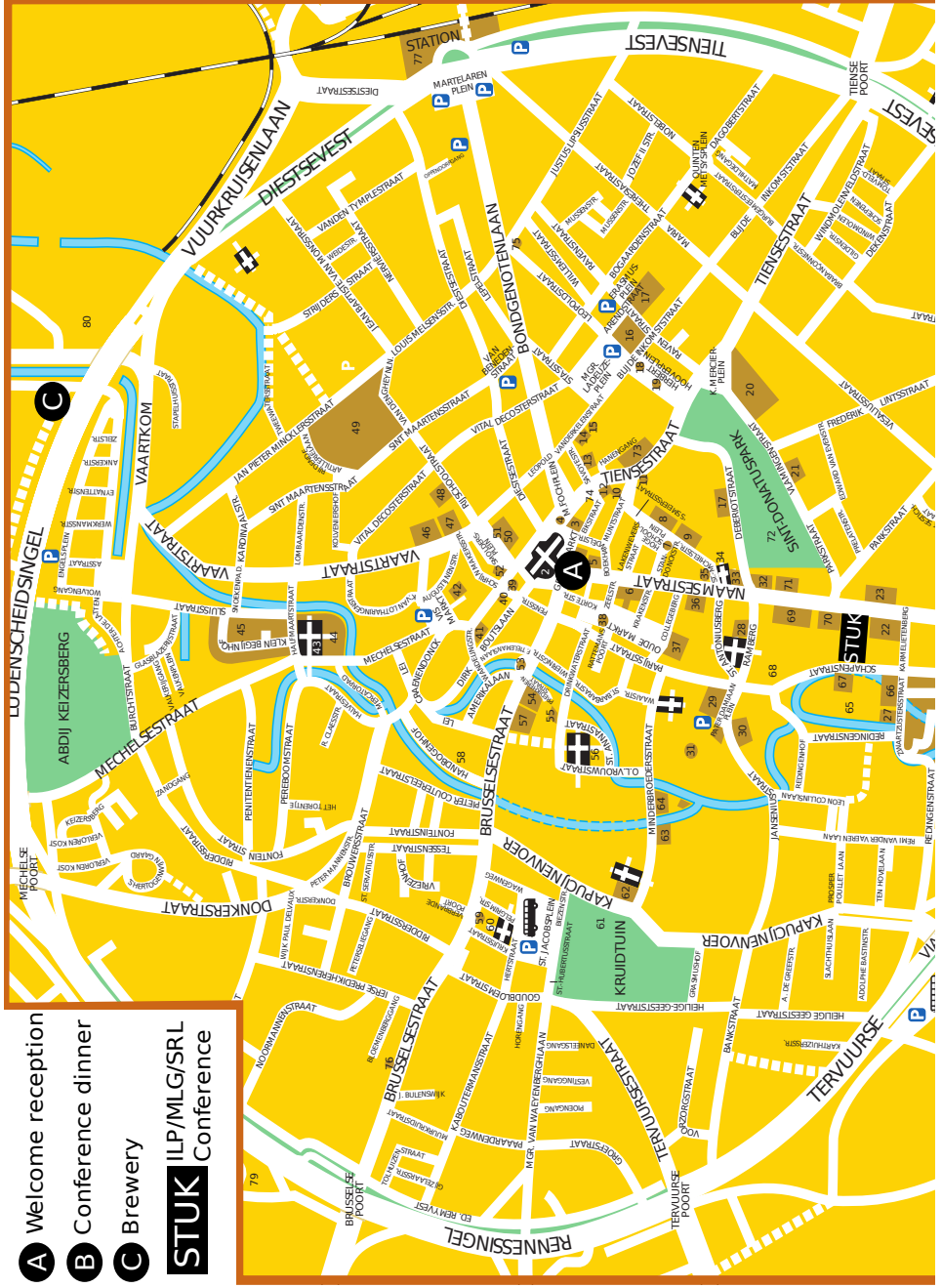
Transfer Learning refers to learning of knowledge in one domain that can be applied to a different domain. In this paper, we view transfer learning as generalization of knowledge in a richer representation language that includes multiple subdomains as parts of the same superdomain. We employ relational templates of different specificity to learn pieces of additive value functions. We show significant transfer of learned knowledge across different subdomains of a real-time strategy game by generalizing the value function using relational templates.

### **Automatic Revision of Metabolic Networks through Logical Analysis of Experimental Data**

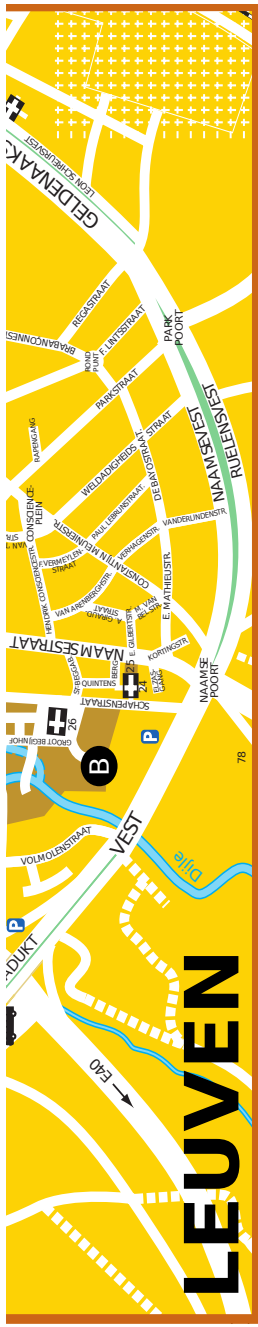
*Oliver Ray, Ken Whelan and Ross King*

This paper presents a nonmonotonic ILP approach for the automatic revision of metabolic networks through the logical analysis of experimental data. The method extends previous work in two respects: by suggesting revisions that involve both the addition and removal of information; and by suggesting revisions that involve a combination of gene function, enzyme inhibition and metabolic reactions. Our proposal is based on a new declarative theory of metabolism expressed in a nonmonotonic logic programming formalism. With respect to this theory, a mixture of abductive and inductive inference is used to compute a set of minimal revisions needed to make a given network consistent with some observed data. In this way, we describe how a hybrid reasoning system was able to usefully revise a state-of-the-art metabolic network in order to better account for real-world experimental data acquired by an autonomous laboratory platform known as the Robot Scientist.

- A** Welcome reception
  - B** Conference dinner
  - C** Brewery
- STUK** ILP/MLG/SRL  
Conference







Een schat van een stad

1. Stadhuis
2. Sint-Pieterskerk met Meester Jan en Museum voor Religieuze Kunst
3. Tafelrand
4. Fons Sapientiae (Fonske)
5. Dirk Bouts
6. Universiteitshal
7. André Dumont
8. Pauscollege
9. Maria-Theresiacollege
10. College van Luik
11. Fiere Margriet
12. Huis Den Bel
13. Stedelijk museum Vanderkelen-Mertens
14. College heilige Yvo + Vlaams filmmuseum en archief Kamerood 60
16. Centrale
17. Universiteitsbibliotheek Faculteit Letteren en Wijsbegeerte
18. Ballon van de Vriendenschap
19. Edouard Remy
20. Hoger instituut voor

Wijsbegeerte

21. Kapel van Onze Lieve Vrouw Ter Koornts
  22. Amerikaans college
  23. Vigliuscollege
  24. Sint-Kwintenskerk
  25. Kapel van Jezus in 't Steentje of Janneke de grijzer
  26. Groot Begijnhof
  27. Industriële maaldenij
  28. Dijiemolens
  29. Sint-Antoniuskapel
  30. Paridaensinstituut
  31. Iers college
  32. Janseniustoren
  33. Prenomstretcollege
  34. Sint-Michielskerk
  35. Renée
  36. Heilig Geestcollege + Spoelbergmuseum
  37. Heilig Drieuldighedscollege
  38. Kotmadam
- Een schat van een stad (uitbreiding)
39. Erasmus

40. Huis In de Ploeg

41. Huis De Kruiwagen
  42. Busleydencollege of Collegium Trilingue
  43. Sint-Geertruikerk
  44. Sint-Geertruabdij + scoutsmuseum
  45. Klein Begijnhof
  46. Luxemburgcollege
  47. Villerscollege
  48. Tweebronnen + Rijsschoolplein
  49. Sint-Maartensdaiwijk
  50. Gerechtshof
  51. Pieter Couterel
  52. Postkantoor
- Een ommetje waard
53. Paepe Toon
  54. Peltzcollege
  55. Huis Guy de M'orillon
  56. Kerk Onze-Lieve-Vrouw ten Predikheren
  57. Sint-Elisabethgasthuis (voormalig)
- met Romaanse Poort + Cultuurcentrum

58. Handbogenhof

59. Pater Darniaan
  60. Sint-Jacobskerk
  61. Stedelijke Kruituin
  62. Anatomisch Theater
  63. Justus Lipsiuscollege
  64. Herenhuis van de familie d'Udekem d'Accoz (voormalig)
  65. Dillepark
  66. Klooster van de Zwartzusters
  67. Huis Bethlehem
  68. Wolfenpoort
  69. Rambergpark
  70. Van Dalecollege
  71. Atrechtcollege
  72. Sint-Donatuspark
  73. College de Valk
  74. Huisbrouwerij Domus
- andere
75. Justus Lipsius
  76. Koeschietier
  77. Station
  78. Kasteel Arenbergen
  79. Universitaire Ziekenhuizen Gasthuisberg
  80. Brouwerij Interbrew

**Finding Relational Associations in HIV Resistance Mutation Data***Lothar Richter, Regina Augustin and Stefan Kramer*

HIV therapy optimization is a hard task due to rapidly evolving mutations leading to drug resistance. Over the past five years, several machine learning approaches have been developed for decision support, mostly to predict therapy failure from the genotypic sequence of viral proteins and additional factors. In this paper, we define a relational representation for an important part of the data, namely the sequences of a viral protein (reverse transcriptase), their mutations, and the drug resistance(s) associated with those mutations. The data were retrieved from the Los Alamos National Laboratories' (LANL) HIV databases. In contrast to existing work in this area, we do not aim directly for predictive modeling, but take one step back and apply descriptive mining methods to develop a better understanding of the correlations and associations between mutations and resistances. In our particular application, we use the Warmr algorithm to detect non-trivial patterns connecting mutations and resistances. Our findings suggest that well-known facts can be rediscovered, but also hint at the potential of discovering yet unknown associations.

**ILP, the Blind, and the Elephant: Euclidean Embedding of Co-Proven Queries***Hannes Schulz, Kristian Kersting and Andreas Karwath*

Relational data is complex. This complexity makes one of the basic steps of ILP difficult for the untrained user: understanding the data. And if the user cannot easily understand it, she gets frustrated soon or draws incomplete conclusions. The situation is very much as in the parable of the blind men and the elephant that appears in many cultures. In this tale the blind work independently and with quite different pieces of information, thereby drawing very different conclusions about the nature of the beast. In contrast, visual representations make it easy to shift from one perspective to another while exploring and analyzing data. This paper describes a method for embedding interpretations and queries into a single, common Euclidean space based on their co-proven statistics. We demonstrate our method on real-world datasets showing that ILP results can indeed be captured at a glance.

**Parameter Screening and Optimisation for ILP using Designed Experiments***Ashwin Srinivasan and Ganesh Ramakrishnan*

Reports of experiments conducted with an Inductive Logic Programming system rarely describe how specific values of hyperparameters of the system are arrived at when constructing models. Usually, no attempt is made to identify sensitive hyperparameters, and those that are used are often given "factory-supplied" default values, or values obtained from some non-systematic exploratory analysis. The immediate consequence of this is, of course, that it is not clear if better models could have been obtained if some form of hyperparameter selection and optimisation had been performed. Questions follow inevitably on the experiments themselves: specifically, are all algorithms being treated fairly, and is the exploratory phase sufficiently well-defined to allow the

experiments to be replicated? In this paper, we investigate the use of hyperparameter selection and optimisation techniques grouped under the study of experimental design. Screening and response surface methods determine, in turn, sensitive hyperparameters and good values for these hyperparameters. Screening is done here by constructing a stepwise regression model relating the utility of an ILP system's hypothesis to its input hyperparameters, using systematic combinations of values of input hyperparameters (technically speaking, we use a two-level fractional factorial design of the input hyperparameters). The hyperparameters used by the regression model are taken to be the sensitive hyperparameters for the system for that application. We then seek an assignment of values to these sensitive hyperparameters that maximise the utility of the ILP model. This is done using the technique of constructing a local "response surface". The hyperparameters are then changed following the path of steepest ascent until a locally optimal value is reached. This combined use of hyperparameter selection and response surface-driven optimisation has a long history of application in industrial engineering, and its role in ILP is investigated using two well-known benchmarks. The results suggest that computational overheads from this preliminary phase are not substantial, and that much can be gained, both on improving system performance and on enabling controlled experimentation, by adopting well-established procedures such as the ones proposed here.

### **Kinetic Models for Logic-Based Hypothesis Finding in Metabolic Pathways**

*Gabriel Synnaeve, Andrei Doncescu and Katsumi Inoue*

In this paper we propose a logical model of Glycolysis and Pentose phosphate pathways of E.Coli that allows the analysis of the dynamical response of a biological system perturbed by a pulse of glucose. Our goal is to give a better comprehension of the physiological state of the cell from a better interpretation of the interactions between metabolic and signaling networks. Starting with the discretization of concentrations of some metabolites considered in steady state, we fully explain our approach to build a symbolic model applied to kinetics. Finally, hypothesis finding produces logical formulas on the metabolites through abduction which cannot be measured during dynamical state.

### **Don't Fear Optimality: Sampling for Probabilistic-Logic Sequence Models**

*Ingo Thon*

One of the current challenges in artificial intelligence is modeling dynamic environments that change due to the actions or activities undertaken by people or other agents. The task of inferring hidden states, e.g. the activities or intentions of people, based on observations is called filtering. Standard probabilistic models such as Dynamic Bayesian Networks are able to solve this task efficiently using approximative methods such as particle filters. However, these models do not support logical or relational representations. The key contribution of this paper is the upgrade of a particle filter algorithm for use with a probabilistic logical representation through the definition of a proposal distribution. The performance of the algorithm depends largely on how well this distribution fits the target distribution. We adopt the idea of logical compilation

into Binary Decision Diagrams for sampling. This allows us to use the optimal proposal distribution which is normally prohibitively slow.

### **Exploiting Global Structures in Bayesian Network Compilation by Zero-suppressed BDDs**

*Daisuke Tokoro, Ai Fukunaga, Kiyoharu Hamaguchi, Toshinobu Kashiwabara and Shin-ichi Minato*

This paper addresses an algorithm for probabilistic inference for Bayesian Networks (BNs). Our algorithm uses Zero-suppressed BDDs for compiling BNs. We introduce "separation variable" to reflect global structures of BNs, which provides more compact ZBDDs. We show some experimental results to compare our method with the state-of-the-art tool.

### **Can ILP be Applied to Large Dataset?**

*Hiroaki Watanabe and Stephen Muggleton*

There exist large data in science and business. Existing ILP systems cannot be applied effectively for data sets with 10000 data points. In this paper, we consider a technique which can be used to apply for more than 10000 data by simplifying it. Our approach is called Approximative Generalisation and can compress several data points into one example. In case that the original examples are mixture of positive and negative examples, the resulting example is ascribed in probability values representing proportion of positiveness. Our longer term aim is to apply on large Chess endgame database to allow well controlled evaluations of the technique. In this paper we start by choosing a simple game of Noughts and Crosses and we apply mini-max backup algorithm to obtain database of examples. These outcomes are compacted using our approach and empirical results show this has advantage both in accuracy and speed. In further work we hope to apply the approach to large database of both natural and artificial domains.

# ABSTRACTS SRL

## Oral Presentation

### **Probabilistic Similarity Logic**

*Matthias Broecheler and Lise Getoor*

Many interesting research problems, such as ontology alignment and collective classification, require probabilistic and collective inference over imprecise evidence. Existing approaches are typically ad-hoc and problem-specific, requiring significant effort to devise and implement and provide poor generalizability. In this paper, we introduce probabilistic similarity logic (PSL), a simple, yet powerful language for describing problems which require probabilistic reasoning about similarity where, in addition to reasoning probabilistically, we want to capture both logical constraints and imprecision. We prove that PSL inference is polynomial and describe an efficient implementation which uses an off-the-shelf convex optimization package. Lastly, we outline a wide range of application areas for PSL.

### **Anytime Lifted Belief Propagation**

*Rodrigo de Salvo Braz, Sriraam Natarajan, Hung Bui, Jude Shavlik and Stuart Russell*

Lifted first-order probabilistic inference has been receiving increasing attention. To date, all lifted inference methods require a model to be shattered against itself and evidence, before inference start. In many situations this will lead to great fracturing of the model, producing a new one that is not far from propositionalized and therefore canceling the benefits of lifted inference. Even when shattering is not so severe, intuition suggests it should not involve an entire model, but only the most relevant parts to the evidence and query at hand. We present an algorithm that corresponds to this intuition by performing shattering during belief propagation inference, on a as-needed basis, starting on the most relevant parts of a model first. The trade-off is having an (exact) bound on the query's belief rather than an exact belief. In many problems one can find a very good bound for only a fraction of the cost for computing an exact solution. Moreover, the bound converges to the exact solution as inference and shattering converge to the entire model. Interestingly, this algorithm mirrors theorem-proving, helping to close the gap between probabilistic and logic inference.

### **Max-Margin Weight Learning for Markov Logic Networks**

*Tuyen Huynh and Raymond Mooney*

Markov logic networks (MLNs) are an expressive representation for statistical relational learning that generalizes both first-order logic and graphical models. Existing discriminative weight learning methods for MLNs all try to learn weights that optimize the

Conditional Log Likelihood (CLL) of the training examples. In this work, we present a new discriminative weight learning method for MLNs based on a max-margin framework. This results in a new model, Max-Margin Markov Logic Networks (M3LNs), that combines the expressiveness of MLNs with the predictive accuracy of structural Support Vector Machines (SVMs). To train the proposed model, we design a new approximation algorithm for loss-augmented inference in MLNs based on Linear Programming (LP). The experimental result shows that the proposed approach generally achieves higher F1 score than the current best discriminative weight learner for MLNs.

### **Lifted Aggregation in Directed First-order Probabilistic Models**

*Jacek Kiszyński and David Poole*

There is an ongoing effort to design efficient lifted inference algorithms for first-order probabilistic models. This paper focuses on directed first-order models that require an aggregation operator when a parent random variable is parameterized by logical variables that are not present in a child random variable. We describe our work on extending Milch et al.'s C-FOVE lifted inference algorithm with efficient lifted aggregation procedures.

### **A Language for Relational Decision Theory**

*Aniruddh Nath and Pedro Domingos*

In recent years, many representations have been proposed that combine graphical models with aspects of first-order logic, along with learning and inference algorithms for them. However, the problem of extending decision theory to these representations remains largely unaddressed. In this paper, we propose a framework for relational decision theory based on Markov logic, which treats weighted first-order clauses as templates for features of Markov networks. By allowing clauses to have utility weights as well as probability weights, very rich utility functions can be represented. In particular, both classical planning and Markov decision processes are special cases of this framework.

### **Cutting Plane MAP Inference for Markov Logic**

*Sebastian Riedel*

Abstract In this work we present Cutting Plane Inference (CPI) for MAP inference in Markov Logic. CPI incrementally solves partial Ground Markov Networks, adding formulae only if they are violated in the current solution. We show dramatic improvements in terms of efficiency, and discuss scenarios where CPI is likely to be fast.

### **First-Order Bayes-Ball for CP-Logic**

*Nima Taghipour, Wannes Meert, Jan Struyf and Hendrik Blockeel*

Efficient probabilistic inference is key to the success of statistical relational learning. One issue that affects inference cost is the presence of irrelevant random variables. The Bayes-ball algorithm can identify such irrelevant variables in a propositional Bayesian

network. This paper presents a lifted version of Bayes-ball, which works directly on the first-order level, and shows how this algorithm applies to CP-logic inference.

### **Materializing and Querying Learned Knowledge**

*Volker Tresp, Yi Huang, Markus Bundschuh and Achim Rettinger*

The Semantic Web (SW) presents new challenges to statistical relational learning. One of the main features of SW data is that it is notoriously incomplete. Consider friend-of-a-friend (FOAF) data. The purpose of the FOAF project is to create a web of machine-readable pages describing people, their relationships, and people's activities and interests, using SW technology. Obviously people vary in their motivation to communicate private information such that person-specific information varies from almost nothing to great detail. Another feature of SW data is the great sparsity of relational information such as friendship: friends in the knowledge base are a small subset of all potential friends. Finally, there is a variety of additional untested and potentially conflicting schema information specified by users. We are convinced that the data situation in the FOAF setting will be typical for many other SW domains. Naturally, missing information can be supplemented. First, deductive reasoning can be used to complement factual knowledge based on axioms. The advantage is that deduced facts can be queried with powerful query languages such as SQL for relational data bases and SPARQL for SW data. The disadvantage is that, generally, valid axioms are rare and many true facts in an application cannot be derived. Inductive reasoning is the second way for gaining additional knowledge by exploiting regularities in the data. The advantage of inductive reasoning is that one can perform inference on a large number of statements. Disadvantages are that only the probability of the truth values of statements can be derived and that machine learning needs to be performed by machine learning experts. The goal of the work presented here is to apply statistical relational learning in sparse and unreliable domains such as the SW. Several requirements need to be fulfilled. The statistical relational learning algorithms need to be able to deal with sparse data typical for SW domains, they must be scalable, they need to be easily configurable, finally it should be possible to integrate the learning results into querying. For the user there should be almost no difference between querying facts, deduced facts and induced information. We define an extension of SPARQL, which allows the integration of the learned probabilistic statements into querying. Statements that can be inferred via logical reasoning can readily be integrated into learning and querying. We study learning algorithms that are suitable for the resulting high-dimensional sparse data matrix. We present experimental results using a friend-of-a-friend data set.

## **Poster Presentation**

### **Lifted Optimization for Relational Preference Rules**

*Ronen Brafman and Yagil Engel*

In reasoning about preference we often have knowledge about the desirable behavior or state of a system of agents/objects, that applies to different instantiations of this system, while instantiations may differ in the number and properties of concrete objects. With this motivation we previously proposed relational preference rules, which is a formalism for modeling preference and value information about complex command and control systems. Preference rules allow to specify weights on first-order statements regarding properties (attributes) of object classes. The rules determine a value for any given set of grounded objects and an assignment to their properties. A major reasoning task with preference rules is to find an assignment to the attributes of the objects in the system, that maximizes this value. In the current work we focus on the computational problem of performing lifted optimization; that is finding an optimal assignment without explicit enumeration of the space of assignments to the grounded set of objects. As we show, lifted reasoning with preferences turn out to be significantly different from lifted probabilistic inference.

### **An Exercise with Statistical Relational Learning Systems**

*Maurice Bruynooghe, Daan Fierens, Bernd Gutmann, Angelika Kimmig, Niels Landwehr, Wannes Meert, Ingo Thon and Luc De Raedt*

We report on two exercises in modeling, inference and learning with seven statistical relational learning systems and use this as a basis for a simple and preliminary comparison between these systems.

### **Incremental EM for Latent-class Relational Models**

*Michael Chiang and David Poole*

Latent-class relational models have found much use in key areas such as recommender systems and social network analysis, where hidden group information about domain individuals can help improve predictive accuracy. Fitting such models, however, is generally infeasible due to the large number of latent quantities that must be inferred, and approximate inference techniques such as Gibbs sampling and variational Bayes are required. In this work, we show how an instance of Neal and Hinton's incremental EM algorithm can be derived for fitting latent-class relational models tractably. Empirical evaluations show good accuracy of models learned with the proposed algorithm in numerous social networks and movie rating domains.

### **On Generative Parameterisations of Markov Logic Networks**

*James Cussens*

A Markov logic network uses weighted clauses to define a probability distribution over Herbrand interpretations. PRISM programs also define distributions over Herbrand interpretations. In this paper we show how to represent any Markov logic network as a PRISM program, and then consider the pros and cons of so doing.

### **Does Google own Youtube? Entity Relationship Extraction with Minimal Supervision**



*Jan De Belder, Wim De Smet, Raquel Mochales Palau and Marie-Francine Moens*

This paper advances state of the art entity relationship extraction from text with minimal supervision in two ways. The contributions are a new weighting scheme, that solves different types of bias caused by having only a few training examples, and a new method to construct a support vector machine in the Multiple Instance Learning (MIL) setting, based on Weighted Least Squares (WLS).

### **Learning Parameters of Chain Logic Theories**

*Arjen Hommersom, Nivea Ferreira and Peter Lucas*

There has been a considerable amount of research on statistical relational learning (SRL) during the past decade, much of it focusing on generalised probabilistic logics. SRL formalisms are typically based on Bayesian or Markov networks. Chain graphs (CGs) generalise both acyclic directed graph (ADG) models, i.e., Bayesian networks, and undirected graph (UG) models, i.e., Markov networks, as they may contain both undirected and directed edges. As a consequence they constitute an attractive point of departure for SRL. In this paper, we start with a logical generalisation of chain graphs, called chain logic, to elaborate a method for learning the parameters of chain logic theories.

### **Relational Gaussian Processes for Learning Preference Relations**

*Kristian Kersting and Zhao Xu*

Gaussian processes have successfully been used to learn preferences among entities. They are attractive because they provide Bayesian approaches for model selection and probabilistic inference. For many entities encountered in real-world applications, however, there are complex relations between them. Considering a preference model taking these relations into account should improve learning performance as it allows for preference information between entities to be shared. In this paper, we show that this is indeed the case. Specifically, we propose a probabilistic kernel model to preference learning based on Silva et al.'s mixed graph Gaussian processes. A new prior distribution, enhanced with relational graph kernels, is proposed to capture the correlations between preferences in the Bayesian framework. We show competitive results against baseline methods when demonstrated on publicly available LETOR datasets.

### **Trading Memory for Answers: Towards Tabling ProbLog**

*Angelika Kimmig, Bernd Gutmann and Vítor Santos Costa*

ProbLog is a recent probabilistic extension of Prolog, where facts can be labeled with mutually independent probabilities that they belong to a randomly sampled program. The implementation of ProbLog on top of the YAP-Prolog system provides various inference algorithms that calculate the success probability of a query, i.e. the probability that the query is provable in a randomly sampled program. We discuss extensions of these algorithms with tabling that broaden the class of problems that can be handled. First, exploiting structure sharing can speed up inference in domains where different

proofs of a query share many subgoals. Second, we extend exact inference to deal with negated ground subgoals in clause bodies.

### **Hypergraph Lifting for Structure Learning in Markov Logic Networks**

*Stanley Kok and Pedro Domingos*

Markov logic networks (MLNs) combine logic and probability by attaching weights to first-order clauses, and viewing these as templates for features of Markov networks. Learning MLN structure from a relational database involves learning the clauses and weights. The state-of-the-art MLN structure learners all involve some element of greedily generating candidate clauses, and are susceptible to local optima. To address this problem, we present an approach that directly utilizes the data in constructing candidates. A relational database can be viewed as a hypergraph with constants as nodes and relations as hyperedges. We find paths of true ground atoms in the hypergraph that are connected via their arguments. To make this tractable (there are exponentially many paths in the hypergraph), we lift the hypergraph by jointly clustering the constants to form higher-level concepts, and find paths in it. We variabilize the ground atoms in each path, and use them to form clauses, which are evaluated using a pseudo-likelihood measure. In our experiments on three real-world datasets, we find that our algorithm outperforms the state-of-the-art approaches.

### **Layered, Multivariate Anomaly Explanations: A First Look**

*Matt Michelson and Sofus Macskassy*

An anomaly is a data point that deviates dramatically from some set of related data points, based on some common metric. While there is significant research on finding anomalous records within data sources, we are unaware of research on explaining why those anomalous data points are actually anomalous (beyond their deviance from the "standard" value for the metric). Generating these explanations is the focus of this paper.

### **Knowledge Intensive Learning - Directed vs Undirected SRL Models**

*Sriraam Natarajan, Prasad Tadepalli, Gautam Kunapuli and Jude Shavlik*

SRL models have become attractive for allowing the domain expert to succinctly encode the knowledge about the domain. The goal in this work is to present work in progress of comparing two kinds of models in their ability to use this knowledge: directed models with combining rules and undirected models based on weighted logics (MLNs). We compare the models in 2 real-world datasets and present the empirical results. From initial experiments, it appears that the directed models perform better when there are small number of rules while the undirected models improve their performance when there are a large number of weakly predictive rules.

### **On Connecting PRISM to CLP(BN)**

*Vitor Santos Costa and Aline Paes*

The last few years have seen significant progress in Statistical Relational Learning.

This progress requires the design and implementation of languages that can be used to express uncertainty over relational data. These languages differ on both the representation they use for structure, with logic-based formalisms being quite popular, and on the representations for uncertainty, with both directed and undirected graphical models being used. Recently, there has been interest on understanding the relationship between these different languages. Establishing a mapping between two different languages is useful to provide insight in understanding the language and sharing implementation technology and experience. In this work, we study the relation between CLP(BN) and PRISM. We will focus on the PRISM to CLP(BN) mapping first, following an example-based approach.

### **From ProbLog to First Order Logic: A First Exploration**

*Joost Vennekens, Angelika Kimmig, Theofrastos Mantadelis, Bernd Gutmann, Luc De Raedt and Maurice Bruynooghe*

ProbLog is a system that allows a user to compute (or approximate) the probability of a query in a theory consisting of Horn clauses guarded by probabilistic facts. This paper explores whether it is feasible to generalize the approach towards theories consisting of first order logic formulas guarded by probabilistic facts.

### **Abstracting Probabilistic Relational Model by Specialisation**

*Hiroaki Watanabe and Stephen Muggleton*

Knowledge abstraction is an essential activity of human intelligence. This paper introduces a logic-based specialisation algorithm for relational domains under uncertainty. The specialisation algorithm takes a pair of existentially quantified conjunctions of first-order literals and computes a most general specialisation of the pair by adopting Plotkin's least general generalisation algorithm. The representation and specialisation algorithm provide a natural way to upgrade non-relational probabilistic models to relational probabilistic models.

# ABSTRACTS MLG

## Oral Presentation

### **Within-network Classification Using Local Structure Similarity**

*Christian Desrosiers and George Karypis*

This paper presents a novel collective classification approach that classifies entities of a network using their local structure. Moreover, a new node similarity measure based on random walks, which takes into account label uncertainty and the degree of nodes, is introduced. Through experimentation on real-life datasets from different domains, we show our method to outperform several state-of-the-art approaches for this problem.

### **Mining Graphs to Discover New Theorems in Mathematics**

*Christian Desrosiers, Alain Hertz, Philippe Galinier and Hansen Pierre*

This paper introduces a new data mining problem of characterizing a class of graphs with a set of forbidden subgraphs. Efficient methods for this problem are presented, and the potential of these methods illustrated on the task of discovering novel and significant results in the field of graph theory.

### **Relational Models for Generating Labeled Real-world Graphs**

*Christoph Lippert, Nino Shervashidze and Oliver Stegle*

Analyzing and understanding the structure of social networks and other real-world graphs has become a major area of research in the field of data mining. An important problem setting is the creation of realistic synthetic graphs that resemble real-world social networks. While a range of efficient algorithms for this task have been proposed, current methods solely take the network topology into account ignoring any node labels. We propose a probabilistic approach to synthetic graph generation with node labels, building on concepts from relational learning.

### **Large-Scale Graph Mining Using Backbone Refinement Classes**

*Andreas Maunz, Christoph Helma and Stefan Kramer*

We present a new approach to large-scale graph mining based on so-called backbone refinement classes. The method efficiently mines tree-shaped subgraph descriptors under minimum frequency and significance constraints, using classes of fragments to reduce feature set size and running times, defined in terms of fragments sharing a common backbone. The method is able to optimize structural inter-feature entropy as opposed to occurrences, which is characteristic for open or closed fragment mining. In the experiments, the proposed method reduces feature set sizes by  $> 90\%$  and  $> 30\%$  compared to complete tree mining and open tree mining, respectively. Evaluation using

cross-validation runs shows that their classification accuracy is similar to the complete set of trees but significantly better than that of open trees. Compared to open or closed fragment mining, a large part of the search space can be pruned due to an improved statistical constraint (dynamic upper bound adjustment), which is confirmed in the experiments in lower running times compared to static upper bound pruning. Further analysis using large-scale datasets confirms that the novel descriptors render large training sets feasible which previously might have been intractable.

### **A Pipeline Approach to Graph Identification**

*Galileo Mark Namata and Lise Getoor*

There is a growing wealth of data describing networks of various types, including social networks, physical networks such as transportation or communication networks, and biological networks. At the same time, there is a growing interest in analyzing these networks, in order to uncover general laws that govern their structure and evolution, and patterns and predictive models to develop better policies and practices. However, a fundamental challenge in dealing with this newly available observational data describing networks is that the data is often of dubious quality – it is noisy and incomplete – and before any analysis method can be applied, the data must be cleaned, and missing information inferred. In this paper, we introduce the notion of graph identification, which explicitly models the inference of a “cleaned” output network from a noisy input graph. It is this output network that is appropriate for further analysis. We present an illustrative example and use the example to explore the types of inferences involved in graph identification, as well as the challenges and issues involved in combining those inferences. We then present a simple, general approach to combining the inferences in graph identification and experimentally show how the performance of graph identification is sensitive to the inter-dependencies among these inferences.

### **A Block Model Suitable for Sparse Graphs**

*Juuso Parkkinen, Janne Sinkkonen, Adam Gyenge and Samuel Kaski*

We introduce a new generative block model for graphs. Vertices (nodes) have mixed memberships in margin components, and edges arise from a multinomial defined over the cartesian product of the margin components. The model is able to represent block structures of “non-community” type, that is, it is able to model linkage between margin components. Compared to earlier mixed membership stochastic blockmodels which have a Bernoulli parameterization for the generation of links between each margin component pair, in the new model collapsed Gibbs samplers need to represent only those interactions with realized data in them, making possible large and sparse block models.

### **Maximum Common Subgraph Mining: A Fast and Effective Approach towards Feature Generation**

*Leander Schietgat, Fabrizio Costa, Jan Ramon and Luc De Raedt*

There exists a wide variety of local graph mining approaches that search for frequent,

correlated or closed patterns in graphs. All these methods return very large sets of patterns which can then be used as features to build classifiers. Here we take a different approach: rather than mining for all optimal local patterns, we randomly sample from the set of maximum common subgraphs. The advantages are that maximum common subgraphs are easier to compute than frequent or correlated patterns (for the class of outerplanar graphs this can even be accomplished in polynomial time), and that the resulting features lead to classification models that achieve significantly better predictive performance than models built on the patterns returned by traditional mining approaches.

### **MAP Structured Prediction by Sampling**

*Shankar Vembu, Thomas Gärtner and Mario Boley*

We consider maximum a posteriori parameter estimation for structured prediction with exponential family models. In this setting, the main difficulty lies in the computation of the partition function and the first-order moment of the sufficient statistics. We consider the case that efficient algorithms for exact uniform sampling from the output space exist. This assumption is orthogonal to the typical assumptions made in structured output learning. It holds, in particular, for the highly relevant problem of sampling potent drugs. Under our uniform sampling assumption, we show that exactly computing the partition function is intractable, but it can be approximated efficiently. Furthermore, we show that the first-order moment of the sufficient statistics can be approximated and that we can sample according to the estimated distribution. We also present some application settings.

## **Poster Presentation**

### **Adaptive XML Tree Mining on Evolving Data Streams**

*Albert Bifet and Ricard Gavaldà*

We propose a new method to classify trees, using closed and maximal frequent trees. Closed trees maintain the same information as frequent trees using less space and maximal trees maintain approximate information. We use them to reduce the number of classification features. We present a new framework for data stream tree classification. For the first component of our classification framework, using a methodology based in Galois Lattice Theory, we present three closed tree mining algorithms: an incremental one IncTreeMiner, a sliding-window based one, WinTreeMiner, and finally one that mines closed trees adaptively from data streams, AdaTreeMiner. To the best of our knowledge this is the first work on tree classification in streaming data varying with time. We give a first experimental evaluation of the proposed classification method.

### **New Results on Listing Closed Sets of Strongly Accessible Set Systems**

*Mario Boley, Tamas Horvath, Axel Poigne and Stefan Wrobel*

In a former work we defined different graph mining settings raising the problem of listing all closed sets of strongly accessible set systems and proposed a DFS-algorithm that solves this problem with polynomial delay and incremental polynomial space. In this extended abstract we summarize our new results on listing closed sets of strongly accessible set systems. In particular, (i) we show that while listing all closed sets is intractable for accessible set systems, it becomes tractable for the class of strongly accessible set systems, (ii) give a listing algorithm for this problem that has not only better complexity, but provides also an algorithmic characterization of strongly accessible set systems, and (iii) present new results on the relationship between closure operators and support-closedness of patterns as defined in data mining.

### **Supervised Classification and Visualization of Social Networks Based on a Probabilistic Latent Space Model**

*Charles Bouveyron, Hugh Chipman and Etienne Côme*

Graph-structured networks are widely used to represent relationships between persons in organizations or communities. Recently, the need of classifying and visualizing such data has suddenly grown due to the emergence of a large number of social network websites. We propose in this paper two supervised approaches for learning a latent space model of the network taking into account both the observed class labels and the graph structure. The first proposed approach introduces the class information through the conditional model of the link existence between two nodes whereas the second one considers the class labels as new observed variables. The learned models are then used to project and classify new nodes.

### **Online Learning a Binary Labeling of a Graph**

*Mickey Brautbar*

We investigate the problem of online learning a binary labeling of the vertices for a given graph. We design an algorithm, Majority, to solve the problem and show its optimality on clique graphs. For general graphs we derive a relevant mistake bound that relates the algorithm's performance to the cut size (the number of edges between vertices with opposite labeling) and the maximum independent set in the graph. We next introduce a novel complexity measure of the true labeling - the frontier and relate the number of mistakes incurred by Majority to this measure. This allows us to show, in contrast to previous known approaches, that our algorithm works well even when the cut size is bigger than the number of vertices. A detailed comparison with previous results is given.

### **Iterative Methods for Low Rank Approximation of Graph Similarity Matrices**

*Thomas Cason, Pierre-Antoine Absil and Paul Van Dooren*

In this paper, we analyze an algorithm to compute a low-rank approximation of the similarity matrix between nodes of two graphs. This problem can be reformulated as an optimization problem of a continuous function  $\Phi(S) = \text{tr}[S^T * M^2(S)]$  where  $S$

is constrained to have unit Frobenius norm, and  $M^2$  is a non-negative linear map. We restrict the feasible set to the smooth manifold of rank- $k$  matrices with unit Frobenius norm and  $k$  identical singular values. We analyze the convergence properties of our algorithm and prove that accumulation points are stationary points of  $\Phi(S)$ . We finally compare our method in terms of speed and accuracy to the full rank algorithm.

## **Context-free Graph Grammars as a Language-bias Mechanism for Graph Pattern Mining**

*Christophe Costa Florencio, Jan Ramon, Jonny Daenen, Jan Van den Bussche and Dries Van Dyck*

Many graph mining tasks involve search over languages of graphs. Generating all elements of a graph language is a non-trivial task, and there are often additional constraints, e.g. to allow for efficient pruning. There are several approaches to generate graphs. For some applications, however, it is desirable to constrain the search and only consider a subclass of all graphs. For this, the use of language bias has been considered, e.g. in the field of inductive logic programming. In the context of graph mining, one would expect of a similar general language bias mechanism that it allows one to specify many different classes of graphs in a declarative way. In the present paper, we propose the use of context-free graph grammars for this purpose. Graph grammars are well known, and a wide variety of classes of graphs can be specified using graph grammars. This makes them suitable for use as an expressive and flexible mechanism for graph pattern language bias. Examples of context-free graph languages include: all trees; all cycles; all outerplanar graphs; all graphs of treewidth at most  $k$ . A particularly elegant approach to graph grammars has been developed by Bauderon and Courcelle. As has been pointed out before, Courcelle's writings might be somewhat inaccessible for readers with little algebraic background. This paper has three purposes: (1) review Courcelle's graph grammars (more precisely: the HR-equation systems) in a simple and, hopefully, accessible manner, (2) define a canonical form for bounded treewidth graphs based on Arnborg's treewidth checking algorithm and (3) show how these can be used as a tractable formalism for the generation of graph patterns.

## **Identification of Hyperedge-Replacement Graph Grammars**

*Christophe Costa Florencio*

The last few years have seen an increasing interest in mining and learning from graphs. Most work in this area has aimed at finding approximate solutions to real-world problems, without much concern for more fundamental issues. This paper is intended as a step towards an approach that is theoretically more sound, by applying ideas from the field of grammar induction. We present results concerning the learning of hypergraph grammars. We consider identification in the limit from both derivation trees and graphs. We show that, by restricting the complexity of grammars, a learnable class of derivation languages is obtained. By exploiting a known invariance result this class is extended to a class that is learnable from hypergraphs.



**A Polynomial Algorithm for Subisomorphism of Open Plane Graphs**

*Guillaume Damiand, Colin de la Higuera, Jean-Christophe Janodet, Emilie Samuel and Christine Solnon*

We address the problem of searching for a pattern in a plane graph, that is, a planar drawing of a planar graph. We define plane subgraph isomorphism and give a polynomial algorithm for this problem. We show that this algorithm may be used even when the pattern graph has holes.

**Towards Scalability of Graph-Mining Based Bug Localisation**

*Frank Eichinger and Klemens Böhm*

(Semi-)automated bug localisation is an important issue in software engineering. Recent techniques based on call graphs and graph mining can locate bugs in relatively small programs, but do not scale for real-world applications. In this paper we describe a bug-localisation approach based on graph mining that has this property, at least according to preliminary experiments. Our main contribution is the definition and analysis of class-level call graphs, with encouraging results.

**The Graphlet Spectrum**

*Risi Kondor and Nino Shervashidze*

Current graph kernels suffer from two limitations: graph kernels based on counting particular types of subgraphs ignore the relative position of these subgraphs to each other, while graph kernels based on algebraic methods are limited to graphs without node labels. In this paper we present the graphlet spectrum, a system of graph invariants derived by means of group representation theory that captures information about the number as well as the position of labeled subgraphs in a given graph. In our experimental evaluation the graphlet spectrum outperforms state-of-the-art graph kernels.

**Classification on Graphs for Dynamic Difficulty Adjustment: More Questions Than Answers**

*Olana Missura and Thomas Gärtner*

Motivated by a problem of dynamic difficulty adjustment in computer games, we propose a new learning setting. We argue that video games require a mechanism for dynamic difficulty adjustment. Such mechanism can be formulated as a classification problem on a graph that combines the properties of online and active learning. The main contribution of this paper is the clarification of this novel learning setting and a summary of open questions.

**New Modularity for Evaluating Communities in Bipartite Networks**

*Tsuyoshi Murata*

Detecting dense subnetworks (communities) and evaluating them are practically important for finding similar vertices in a network. Although Newman's modularity is widely used for evaluating network division, it is for unipartite networks composed of

only one vertex type. As the attempts for evaluating divisions of bipartite networks, Guimera and Barber propose bipartite modularities. The author proposes another bipartite modularity that allows one-to-many correspondence of communities of different vertex types.

### **A Statistical Generalization of Music Graph Spectral Mining**

*Alberto Pinto*

Graph spectral representation of pitch class sequences provides an effective tool for mining structural music information. Here we show how it is possible to extend this approach by embedding rhythmical information into graph representation through a complex valued function defined on the graph edge set allowing for both pitch-class and rhythmical modelling through coefficients of variation. Experiments run on MIREX 05 dataset show that spectral distances calculated from this representation outperform existing methods.



















# Index

- Absil, P., 43  
Akand, E., 21  
Al-Ali, H., 20  
Ali, K., 21  
Anderson, G., 17  
Antanas, L., 22  
Augustin, R., 30
- Bain, M., 21  
Bhattacharya, P., 26  
Bifet, A., 42  
Bina, B., 24  
Blockeel, H., 19, 34  
Boley, M., 42  
Bouveyron, C., 43  
Brafman, R., 35  
Braud, A., 25  
Brautbar, M., 43  
Broecheler, M., 33  
Bruynooghe, M., 36, 39  
Bui, H., 33  
Bundschuh, M., 35  
Burnside, E., 20  
Böhm, K., 45
- Cason, T., 43  
Chen, J., 22  
Chiang, M., 36  
Chipman, H., 43  
Choi, D., 21  
Costa Florencio, C., 44  
Costa, F., 41  
Cussens, J., 36
- Côme, E., 43
- Daenen, J., 44  
Damiand, G., 45  
De Belder, J., 37  
de la Higuera, C., 45  
De Raedt, L., 22, 36, 39, 41  
de Salvo Braz, R., 33  
De Smet, W., 37  
Desrosiers, C., 40  
Domingos, P., 34, 38  
Doncescu, A., 31  
Dries, A., 22
- Eichinger, F., 45  
Engel, Y., 35  
Esposito, F., 25
- Ferreira, N., 37  
Fierens, D., 22, 36  
Frasconi, P., 19  
Fukunaga, A., 32  
Furukawa, K., 18
- Galinier, P., 40  
García Jiménez, B., 23  
Gavalda, R., 42  
Getoor, L., 33, 41  
Gillis, J., 17  
Gosar, C., 23  
Gutmann, B., 36, 37, 39  
Gyenge, A., 41  
Gärtner, T., 42, 45

- Haase, C., 18  
Hamaguchi, K., 32  
Helma, C., 40  
Hertz, A., 40  
Hommersom, A., 37  
Hoonakker, F., 23  
Horvath, T., 17, 42  
Huang, Y., 35  
Huynh, T., 33
- Inoue, K., 18, 20, 31  
Ishihata, M., 20
- Jaeger, M., 19  
Janodet, J., 45
- Kalgi, S., 23  
Karwath, A., 30  
Karypis, G., 40  
Kashiwabara, T., 32  
Kaski, S., 41  
Keirouz, W., 20  
Kersting, K., 30, 37  
Khosravi, H., 24  
Khuri, S., 20  
Kimmig, A., 36, 37, 39  
King, R., 27  
Kisyański, J., 34  
Kobayashi, I., 18  
Kok, S., 38  
Kondor, R., 45  
Konik, T., 21, 24  
Kramer, S., 30, 40  
Kunapuli, G., 25, 38  
Kuter, U., 24  
Kuželka, O., 18
- Lachiche, N., 23, 25  
Landwehr, N., 22, 36  
Ledezma Espino, A., 23  
Lehmann, J., 18  
Lesbegueries, J., 25  
Leung, K., 21
- Lippert, C., 40  
Lippi, M., 19  
Lisi, F.A., 25  
Lucas, P., 37
- Maclin, R., 25  
Macskassy, S., 38  
Mantadelis, T., 39  
Maunz, A., 40  
Meert, W., 19, 34, 36  
Michelson, M., 38  
Mihalkova, L., 19  
Minato, S., 32  
Missura, O., 45  
Mochales Palau, R., 37  
Moens, M., 37  
Mooney, R., 33  
Muggleton, S., 19, 22, 25, 26, 32, 39  
Murata, T., 45
- Namata, G.M., 41  
Nassif, H., 20  
Natarajan, S., 25, 33, 38  
Nath, A., 34  
Nejati, N., 24
- O'Reilly, C., 25  
Oliphant, L., 20
- Paass, G., 17  
Paes, A., 25, 38  
Page, D., 20, 25  
Pahlavi, N., 26  
Parkkinen, J., 41  
Passerini, A., 19  
Patel, A., 26  
Pfahringer, B., 17  
Pierre, H., 40  
Pinto, A., 46  
Poigne, A., 42  
Poole, D., 34, 36  
Proper, S., 27

- Ramakrishnan, G., 23, 26, 30  
Ramon, J., 41, 44  
Ray, O., 27  
Reichartz, F., 17  
Rettinger, A., 35  
Richardson, M., 19  
Richter, L., 30  
Riedel, S., 34  
Russell, S., 33
- Samuel, E., 45  
Sanchis de Miguel, A., 23  
Santos Costa, V., 25, 37, 38  
Santos, J., 19  
Sato, T., 20  
Schietgat, L., 41  
Schulte, O., 24  
Schulz, H., 30  
Shapiro, D., 21  
Shavlik, J., 20, 21, 25, 33, 38  
Shervashidze, N., 40, 45  
Sinkkonen, J., 41  
Solnon, C., 45  
Srinivasan, A., 23, 30  
Stegle, O., 40  
Struyf, J., 19, 34  
Synnaeve, G., 31
- Tadepalli, P., 27, 38  
Taghipour, N., 34  
Tamaddoni-Nezhad, A., 19  
Temple, M., 21  
Thon, I., 22, 31, 36  
Tokoro, D., 32  
Torrey, L., 21  
Tresp, V., 35
- Van den Bussche, J., 17, 44  
Van Dooren, P., 43  
Van Dyck, D., 44  
van Otterlo, M., 22  
Varnek, A., 23  
Vembu, S., 42
- Vennekens, J., 39
- Wagner, A., 23  
Walker, T., 25  
Watanabe, H., 32, 39  
Whelan, K., 27  
Wrobel, S., 17, 42
- Xu, Z., 37
- Zaverucha, G., 25  
Zelezny, F., 18





