Prior Knowledge for Text and Language Processing Workshop

Helsinki 9 July 2008

Organizers:
Guillaume Bouchard (Xerox Research Centre Europe)
Hal Daumé III (University of Utah)
Marc Dymetman (Xerox Research Centre Europe)
Yee Whye Teh (University College London)
Prior knowledge for text and language processing workshop: context

Pascal-2 Thematic Programme:
“Leveraging Complex Prior Knowledge for Learning”

Current Workshops:
- Bayesian Approaches to Prior Knowledge
  - Bayesian Research Kitchen, Grasmere, Lake District, U.K., 8th-10th June 2008
  - Nonparametric Bayes Workshop, Helsinki, 9th July 2008

Prior Knowledge encoded as Graphs
- Mining Learning and Graphs, Helsinki, July 4th-5th 2008

Prior Knowledge encoded as Differential Equations
- Approximate Inference in Stochastic Processes and Dynamical Systems,
  Cumberland Lodge, May 27th-29th 2008

Biological Data and Prior Knowledge
- Learning and Inference in Computational Systems Biology, Glasgow, March 26th-27th, 2008

Language and Prior Knowledge
- Prior Knowledge for Text and Language Processing, Helsinki, 9th July 2008
Prior knowledge for text and language processing workshop: morning

9:05-9:50  Invited Talk  Mark Johnson
9:50-10:00 Poster Preview
10:00-10:30 Coffee Break
10:30-10:55 Talk  Ming-Wei Chang, Lev Ratinov and Dan Roth
10:55-11:40 Invited Talk  Jason Weston
11:40-12:30 Posters
   Songfang Huang and Steve Renals
   Honglak Lee, Rajat Raina, Alex Teichman and Andrew Y. Ng
   Kinfe Tadesse Mengistu, Mirko Hanneman, Tobias Baum and
   Andreas Wendemuth
   Thomas J. Murray, Panayiotis G. Georgiou and Shrikanth S.
   Narayanan
   Andrea Vlachos, Zoubin Ghahramani and Anna Korhonen

12:30-14:30 Lunch
Prior knowledge for text and language processing workshop: **afternoon**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30-15:15</td>
<td><strong>Invited Talk</strong></td>
<td>Pedro Domingos</td>
</tr>
<tr>
<td>15:15-15:40</td>
<td><strong>Talk</strong></td>
<td>Mikaela Keller, John S. Brownstein and Clark C. Freifeld</td>
</tr>
<tr>
<td>15:40-16:05</td>
<td><strong>Talk</strong></td>
<td>Hanna M. Wallach, Charles Sutton and Andrew McCallum</td>
</tr>
<tr>
<td>16:05-16:30</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>16:30-17:30</td>
<td><strong>Panel</strong></td>
<td>David Blei, Fabrizio Costa, Pedro Domingos, Peter Grünwald, Mark Johnson, Jason Weston</td>
</tr>
<tr>
<td>17:30-17:55</td>
<td><strong>Talk</strong></td>
<td>Partha Pratim Talukdar, Ted Sandler, Mark Dredze, Koby Crammer, John Blitzer and Fernando Pereira</td>
</tr>
<tr>
<td>17:55-18:00</td>
<td>Wrap-up</td>
<td></td>
</tr>
</tbody>
</table>
Things to note

• Hardcopies of proceedings
• Video recordings
• Invited talks: 40 mns + 5 mns
• Contributed talks: 20 mns + 5 mns
• Poster preview: 9:50-10:00
• Posters co-located with NP-Bayes: 11:40-12:30
• Panel: 16:30-17:30
  – interactive, controversial
Panel: the right attitude

- Prior knowledge vs. raw data in text and language processing
- Paradigms for encoding prior knowledge
- Prior knowledge from humans vs. prior knowledge from other data (than the data for the current task)
- Different forms of human prior linguistic knowledge; approaches to encoding such knowledge
- Prior knowledge that is implicitly coded in features or algorithms vs. prior knowledge which is explicitly represented and manipulated
- Prior knowledge vs. regularization