ICML 2010
Program Chair Report

Johannes Fürnkranz
Thorsten Joachims
Program Chairs
Overview of Reviewing Process

Sep  Generate list of subject areas
Oct 1  Recruit area chairs (42+4)
Oct   Recruit reviewers (517, 80% accept)
Feb 1  Submissions (594)
Feb 3  Phase I: Bidding by AC and reviewers
       → 1 AC and 2 reviewers per submission
Mar 5  Author rebuttal
Mar 9  Phase II: ACs decide whether to
       a) reject clear cases (13%)
           b) manually assign at least 1 additional reviewer
       Discussion
Apr 5  AC recommendation and meta-review
       PC chairs clarify and make final decisions
Survey

• Many of upcoming slides will refer to the result of a survey
  – invited all area chairs, reviewers, authors
  – more than 550 responses!
• Will publish full survey.
Submissions Flowing in...

<table>
<thead>
<tr>
<th>Date</th>
<th>#Papers</th>
<th>#PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>06.1., 00.00</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>25.1., 08.30</td>
<td>39</td>
<td>21</td>
</tr>
<tr>
<td>26.1., 11.00</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>27.1., 11.00</td>
<td>62</td>
<td>38</td>
</tr>
<tr>
<td>28.1., 08.15</td>
<td>71</td>
<td>44</td>
</tr>
<tr>
<td>29.1., 08.15</td>
<td>87</td>
<td>58</td>
</tr>
<tr>
<td>30.1., 15.00</td>
<td>115</td>
<td>83</td>
</tr>
<tr>
<td>31.1., 08.05</td>
<td>129</td>
<td>97</td>
</tr>
<tr>
<td>31.1., 15.45</td>
<td>143</td>
<td>112</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>#Papers</th>
<th>#PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.2., 08.45</td>
<td>200</td>
<td>169</td>
</tr>
<tr>
<td>01.2., 16.40</td>
<td>250</td>
<td>214</td>
</tr>
<tr>
<td>01.2., 22.00</td>
<td>328</td>
<td>295</td>
</tr>
<tr>
<td>02.2., 00.30</td>
<td>366</td>
<td>334</td>
</tr>
<tr>
<td>02.2., 09.00</td>
<td>496</td>
<td>474</td>
</tr>
<tr>
<td>02.2., 11.25</td>
<td>568</td>
<td>558</td>
</tr>
<tr>
<td>02.2., 11.45</td>
<td>579</td>
<td>570</td>
</tr>
<tr>
<td>02.2., 12.00</td>
<td>599</td>
<td>592</td>
</tr>
<tr>
<td>02.2., 14.00</td>
<td>605</td>
<td>598</td>
</tr>
</tbody>
</table>
Submissions

- more or less the same as last years

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submissions</td>
<td>583</td>
<td>595</td>
<td>594</td>
</tr>
<tr>
<td>Accepted</td>
<td>158</td>
<td>160</td>
<td>152</td>
</tr>
<tr>
<td>Acceptance Rate</td>
<td>27.10%</td>
<td>26.89%</td>
<td>25.92%</td>
</tr>
</tbody>
</table>
Submissions

• more or less the same as last year
• ... not only in terms of numbers:

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accept</strong></td>
<td>7.854</td>
<td>7.480</td>
</tr>
<tr>
<td><strong>Reject</strong></td>
<td>8.082</td>
<td>8.204</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>8.021</td>
<td>8.012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accept</strong></td>
<td>141.92</td>
<td>139.40</td>
</tr>
<tr>
<td><strong>Reject</strong></td>
<td>140.09</td>
<td>142.15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>140.59</td>
<td>141.42</td>
</tr>
</tbody>
</table>
Supplementary Materials

• Submission of supplementary materials were allowed this year (as suggested last year)
  – ca. 20% of submissions had supplementary materials (121)
  – reviewers could choose whether to look at it or not

• did reviewers look at it?

<table>
<thead>
<tr>
<th></th>
<th>Always</th>
<th>Sometimes</th>
<th>Never</th>
<th>My papers didn't have any</th>
<th>I don't know whether my papers had any</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Subject Area</td>
<td>acc</td>
<td>sub</td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reinforcement Learning</td>
<td>16</td>
<td>49</td>
<td>32,65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature Selection and Dimensionality Reduction</td>
<td>9</td>
<td>31</td>
<td>29,03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probabilistic Models</td>
<td>6</td>
<td>28</td>
<td>21,43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kernel Methods</td>
<td>6</td>
<td>26</td>
<td>23,08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering</td>
<td>8</td>
<td>26</td>
<td>30,77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimization Algorithms</td>
<td>9</td>
<td>25</td>
<td>36,00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervised Learning</td>
<td>3</td>
<td>23</td>
<td>13,04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-Supervised Learning</td>
<td>6</td>
<td>22</td>
<td>27,27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Learning</td>
<td>6</td>
<td>20</td>
<td>30,00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer and Multi-Task Learning</td>
<td>3</td>
<td>20</td>
<td>15,00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary Subject Area</th>
<th>acc</th>
<th>sub</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-Series Analysis</td>
<td>5</td>
<td>8</td>
<td>62,50</td>
</tr>
<tr>
<td>Deep Architectures</td>
<td>6</td>
<td>10</td>
<td>60,00</td>
</tr>
<tr>
<td>Multi-Agent and Co-Operative Learning</td>
<td>4</td>
<td>7</td>
<td>57,14</td>
</tr>
<tr>
<td>Cognitive Models of Learning</td>
<td>2</td>
<td>4</td>
<td>50,00</td>
</tr>
<tr>
<td>Planning and Control</td>
<td>1</td>
<td>2</td>
<td>50,00</td>
</tr>
<tr>
<td>Graph Mining</td>
<td>1</td>
<td>2</td>
<td>50,00</td>
</tr>
<tr>
<td>Bayesian Inference</td>
<td>7</td>
<td>16</td>
<td>43,75</td>
</tr>
<tr>
<td>Learning Theory</td>
<td>6</td>
<td>15</td>
<td>40,00</td>
</tr>
<tr>
<td>Cost-Sensitive Learning</td>
<td>2</td>
<td>5</td>
<td>40,00</td>
</tr>
<tr>
<td>Optimization Algorithms</td>
<td>9</td>
<td>25</td>
<td>36,00</td>
</tr>
</tbody>
</table>
Diversity of Submissions

• We tried to attract papers from areas that have not seen many submissions lately
  – by nominating area chair in these areas
  – by announcing the keywords + AC in the Cfp
  – by instructing ACs to solicit papers in their area
  – by posting the cfp to lists in these subareas
## Diversity of Submissions

- It did not work...

<table>
<thead>
<tr>
<th>Area</th>
<th>Primary Subject Area</th>
<th>Primary or Secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>acc</td>
<td>sub</td>
</tr>
<tr>
<td>Constructive Induction and Theory Revision</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Case-Based Reasoning</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fuzzy Learning Systems</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Knowledge-Intensive Learning</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Evolutionary Computation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pattern Mining and Inductive Querying</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Meta-Learning</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Discovery</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>None of the provided KW</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
Diversity of Submissions

• It did not work...

• Possible reasons:
  – same time slot as KDD?
    • on the other hand: we also overlapped with AI-STATS and have many statistical papers
  – nobody is working in these areas?
    • no, many still have separate conferences
  – they think we are not interested in them?
    • we tried to convince them otherwise
  – they are not interesting in going to ICML?
AC and Reviewer Recruiting

• Area Chairs (46, 4 after submissions)
  – Generated list of keywords to cover all areas
  – Invite area chairs to each cover 2-4 areas
  – Many Areas were covered by more than 1 AC

• Reviewers (517)
  • ACs +PC Chairs + Local Chairs nominated reviewers
  • 643 invitations (80% acceptance rate)
The ICML Reviewer Pageant

# Nominations

# Reviewers

- Luc De Raedt
- Thomas Gärtner
- Bernhard Pfahringer
- Ruslan Salakhutdinov
- Matthias Seeger

- Karsten Borgwardt
- Jerry Zhu
Reviewing Process

1. Blindness
2. Assigning area chairs to papers
3. Assigning reviewers to papers → 2 Phase
4. Author rebuttal
5. Reviewing quality
6. Reviewing load
7. Making decisions
What reviewing model do you prefer?

- double-blind (reviewers do not know the authors, and vice versa) (this year)
- single-blind (reviewers know the authors, but authors don't know the reviewers)
- open (reviewers and authors know each other's identities)
What model do you prefer for assigning area chairs to papers?

- AC bidding with known areas (this year)
- Automatic match based on areas
- Authors bid on area chairs. (last year)
- Alternative proposal
2-Phase Reviewer Assignments

Goal: Get the right reviewers to the papers.
Approach: Use everybody’s input!

• Reviewers:
  • Bid on papers
  • Suggest other reviewers in their Phase I reviews

• Authors:
  • Focus AC and reviewer bidding via keywords
  • (Confidential) comments in rebuttal of Phase I reviews

• Area Chairs:
  • May manually assign any reviewer from global pool in Phase II
  • May recruit additional outside reviewers
2-Phase Reviewer Assignments

Goal: Get the right reviewers to the papers.
Approach: Use everybody’s input!

- **Reviewers**:
  - Bid on papers
  - Suggest other reviewers in their Phase I reviews

- **Authors**:
  - Focus AC and reviewer bidding via keywords
  - (Confidential) comments in rebuttal of Phase I reviews

- **Area Chairs**:
  - May manually assign any reviewer from global pool in Phase II
  - May recruit additional outside reviewers

For Phase II Assignments, AC knows:
- Phase I reviews
- Reviewer suggestions by Phase I reviewers
- Reviewer bids
- Author rebuttal (confidential AC comments)
What model do you prefer for assigning reviewers to papers?

- Two-phase, first bidding, then manual. (this year)
- Two-phase, both manual.
- Single-phase with bidding.
- Alternative proposal

AreaChairs | Reviewers | AuthorsAccepted | AuthorsRejected
--- | --- | --- | ---
Two-phase, first bidding, then manual. (this year) | 60.0% | 50.0% | 40.0% | 30.0%
Two-phase, both manual. | 40.0% | 30.0% | 20.0% | 10.0%
Single-phase with bidding. | 20.0% | 10.0% | 0.0% | 0.0%
Single-phase, manual. | 10.0% | 5.0% | 0.0% | 0.0%
Alternative proposal | 0.0% | 0.0% | 0.0% | 0.0%
Do you think the Phase I reviews were different in quality from the Phase II reviews?

- Phase I substantially better: [Data visualization showing percentage]
- Phase I slightly better: [Data visualization showing percentage]
- Equal: [Data visualization showing percentage]
- Phase II slightly better: [Data visualization showing percentage]
- Phase II substantially better: [Data visualization showing percentage]
Ratings PI vs. PII

<table>
<thead>
<tr>
<th>Accept</th>
<th>2.5</th>
<th>2.5</th>
<th>2.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak accept</td>
<td>36%</td>
<td>33%</td>
<td>35%</td>
</tr>
<tr>
<td>Reject</td>
<td>0%</td>
<td>5%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Avg Rating

Strong Accept/Reject Rating (% of Reviewers)
Confidence PI vs. PII

**Expert**
- Phase I (prelim): 2.1
- Phase I: 2.1
- Phase II: 2.3

**Knowable**
- Phase I (prelim): 2.1
- Phase I: 2.1
- Phase II: 2.3

**Outsider**
- Phase I (prelim): 2.1
- Phase I: 2.1
- Phase II: 2.3

**"Expert" Reviews (%)**
- Phase I (prelim): 26%
- Phase I: 25%
- Phase II: 36%
Predictiveness PI vs. PII

Agree on Accept/Reject

Agree on Accept/Reject ("Strong" Rating)

Agree on Accept/Reject ("Expert" Confidence)

Phase I (prelim)  |  Phase I  |  Phase II

<table>
<thead>
<tr>
<th>Agree on Accept/Reject</th>
<th>69%</th>
<th>73%</th>
<th>78%</th>
<th>82%</th>
<th>87%</th>
<th>89%</th>
<th>80%</th>
<th>81%</th>
<th>84%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree on Accept/Reject (&quot;Strong&quot; Rating)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree on Accept/Reject (&quot;Expert&quot; Confidence)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How did you like your reviewing assignments?

Phase I

Phase II
How would reviewers like to receive their assignments?

- An area chair who knows me should assign them to me: 0.0%
- I want to bid for my papers: 10.0%
- We should try a mix of both models: 60.0%
- Don't care: 60.0%

Bar chart showing the distribution of preferences.
AC Selection of Phase II Reviewers

Phase I Reviews

- very important: 7
- important: 6
- some influence: 12
- no influence: 1

Author Rebuttal

- very important: 6
- important: 4
- some influence: 12
- no influence: 9
Move Rebuttal to after Phase II:
• Pro: less author frustration, less reviewer confusion
• Con: less info for AC, late w.r.t. discussion
ICML2010 substantially better
ICML2010 slightly better
About the same
Other conferences slightly better
Other conferences substantially better
Reviewing Load

- Submissions: 594
- Reviewers: 517
- Reviews: 1860
- Min # of reviews per reviewer: 2
- Avg # of reviews per reviewer: 3.6
- Max # of reviews per reviewer: 7
Decisions

• Recommendations by area chairs
• Final decisions by PC chairs
• Decisions entirely based on quality of papers
  – no comparison between papers
  – no satisfying a given capacity
• Regular accept/reject (after Phase II)
  – at least 3 reviews
• Early reject (after Phase I, 13% of papers)
  – 2 reviews
Should we have early rejects?

No, papers should get three reviews
Yes, in clear cases (this year)
ACs should use them more.

AreaChairs
Reviewers
AuthorsAccepted
AuthorsRejected
Should we have early accepts?

- No (this year)
- Yes, for exceptionally good papers.

Areas of focus:
- AreaChairs
- Reviewers
- AuthorsAccepted
- AuthorsRejected
Application Papers

• Application Papers do receive rather low scores in reviewing
  - „this is nice, but no real contribution to science“
• Invited Application Track
  - was suggested at last year's meeting for attracting application papers
  - committee looked for good application papers that were published elsewhere
  - speakers got free entry to conference (no travel)
• Other ideas: Separate submission track?
Omnipress offers packages for:

- order minimum number of printed proceedings (75 for $80 each, 50 for $106.50 each)
- order proceedings on USB ($13.40) or CD ($7.30)
- no order, available by print-on-demand

Last year: 75 books and 580 CDs

- < 10 copies of books were sold on site

This year we chose the last option:

- cheapest total price ($2910), but no immediate gain (gain is addition to ICML series in omnipress program)
Proceedings

• on-line only proceedings seem to be o.k.

<table>
<thead>
<tr>
<th>Response</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>89.1%</td>
</tr>
<tr>
<td>Yes, but I think it's a bad idea.</td>
<td>7.4%</td>
</tr>
<tr>
<td>No</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

• opinions on printed proceedings are diverse

<table>
<thead>
<tr>
<th>Response</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>35.0%</td>
</tr>
<tr>
<td>Yes, but try to find a cheaper way.</td>
<td>11.7%</td>
</tr>
<tr>
<td>No</td>
<td>53.3%</td>
</tr>
</tbody>
</table>
Videotaping

• We didn't hire videolectures.net this year
• Reason:
  – Cost:
    • only invited talks: €2560,-
    • invited talks + 1 track: €5800,-
  – Funding from PASCAL network was not as in previous years
  – Plenary talks will be filmed by local staff
Based on manual processing of 285 free-text replies
What else do we attend?
Discussion

- Phase I vs. phase II author rebuttal
- Benefit of Invited Application Track
- Usefulness of supplementary material
- Cost/benefit of videotaping
- Cost/benefit of printed proceedings
- ...