Micro 2012
Program Chair’s Remarks

Onur Mutlu
PC Chair
December 3, 2012
Vancouver, BC, Canada
Purpose of This Session

- Provide insight and transparency into the Micro-45 paper submission and selection process
- Provide statistics and some informal analyses
- Describe some new things we tried
- Get your feedback – involve the community
- Enable better future Micro conferences
  - Hopefully, enable the institution of best practices
Basic Process and Statistics

- 228 submissions; 40 papers accepted (17.5% Acc. Rate)

- Reviewed by
  - The 44-member Program Committee
  - The 106-member External Review Committee (fully involved)
  - At least 98 more external/aiding reviewers

- Review process
  - 1325 total reviews, 815 from PC members; 5.8 reviews/paper
  - Rebuttal period to allow for author response
  - Extensive online discussions (160 papers)
  - 1.5-day PC meeting on August 25-26, 2012
    - All PC members present the entire first day

- After decision
  - All papers shepherded
More Information on the Process …

- **Now – Program Chair’s Remarks**
  - Statistics on submitted and accepted papers
  - Paper selection process
  - Metrics used to rank papers for discussion
  - Best paper selection process
  - New things we tried this year
  - What worked and what did not? What can we do better?

- **Message from the Program Chair (in your proceedings)**

- **Feedback very much welcome**
  - I would appreciate all kinds of honest feedback
Our Goals in New Things Tried

- Improve quality, transparency; institute best practices

1. Quality of the reviews and paper selection process
2. Transparency of the selection process
3. Strong involvement of external reviewers and community
4. Authors’ ability to respond to initial reviewer evaluations and potential questions that may come up after rebuttal
5. Quality of the final program and final versions
6. Quality of interactions during conference (especially in the presence of parallel sessions)
The Life of a Micro Submission/Paper

- Submission process
- Review process
- Rebuttal process
- Post rebuttal online discussion and re-scoring
- PC meeting
- Post-PC meeting
- During conference
- After conference – ad infinitum
Paper Submission Process

- Double-blind submission and review process
- Conflicts marked by authors

Submission format the same as the final format; no extra page charges

- Main goal: eliminate fairness issues... Mitigates the concern “Does this paper fit in the final format?”
- Could be done better in the future
  - We did: 12-page 10-pt submissions, 12-page 9-pt final format
  - Better: 11-page 9pt submissions, 12-page 9pt final format.
Origin of Papers

- **Academia only**
  - Submitted: 160
  - Accepted: 20

- **Industry only**
  - Submitted: 0
  - Accepted: 9

- **Both**
  - Submitted: 50
  - Accepted: 10
Main Country of Papers

- USA
- China
- Spain
- Canada
- Korea
- France
- Switzerland
- Sweden
- Cyprus
- UK
- Israel
- Taiwan
- Singapore
- Iran
- Belgium
- Palestine
- Japan
- Italy
- India
- Germany
- Chile
- Brazil

Number of Papers

Accepted
Submitted

0 20 40 60 80 100 120 140 160 180
Optional Information with Submission

- Authors able to provide an appendix
  - Peripheral material that can aid reviewers, such as full proofs of theorems, details of the experiments conducted, or more experimental results
  - Main goal: Satisfy reviewer curiosity; handle out-of-scope issues

- Authors able to say “this paper was submitted to a past conference”

- Authors able to provide past reviews and responses to them (if the paper was submitted to a previous venue and rejected)
  - Main goal: proactive addressing of potential concerns that can come up post-rebuttal... Anticipate and address “I had reviewed this paper previously and the authors did...”
Papers with an Appendix

Number of Papers

Submitted
Accepted

Total
Papers with an Appendix
Papers with Past Reviews

<table>
<thead>
<tr>
<th></th>
<th>Number of Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Submitted</strong></td>
<td>230</td>
</tr>
<tr>
<td><strong>Accepted</strong></td>
<td>14</td>
</tr>
</tbody>
</table>

- **Total**: 244 papers
- **Papers with Past Reviews**: 14 papers
Papers with “Submitted to Past Conference” Checked

![Bar Chart]

- Total
- Papers with “Submitted to Past Conference” Checked

Number of Papers

- Submitted
- Accepted
PC-Authorized papers

Total | PC-Authored Papers

Submitted

Accepted

Number of Papers

0 50 100 150 200 250
Review Process

- Reviewers
  - The 44-member Program Committee
  - The 106-member External Review Committee (fully involved)
  - At least 98 more external/aiding reviewers

- Review form explicitly asked for three specific rebuttal questions
- External reviewers fully involved in the process, up until the PC meeting
Number of Reviews

- 1325 total reviews
  - 815 from PC members
  - 510 from external reviewers
  - All reviewers assigned by me

- 5.8 reviews/paper
- 216 papers received >= 5 reviews
- 159 papers received >= 6 reviews
- 44 papers received >= 7 reviews
Reviews Categorized by Expertise Score

- **Number of Reviews**
  - Accepted
  - Submitted

- **Expert in the subject area of the paper**
- **Knowledgeable, but not an expert in the subject area**
- **Some familiarity with the subject area**
- **Little familiarity with the subject area**

- **Li^le familiarity with the subject area**
Reviews Categorized by Novelty Score

- Surprisingly new contribution
- New contribution
- Incremental improvement
- Very incremental improvement
- Done before (not necessarily published)
- Published before

Number of Reviews

- Accepted
- Submitted

20
Can We Publish Your Review?

- Yes: Submitted 700, Accepted 100
- No: Submitted 600, Accepted 100

Number of Reviews
Rebuttal Process

- Authors’ chance to respond to reviewer concerns after initial evaluation by reviewers

- **Authors were able to see all initial evaluation scores and all comments of reviewers**

- All reviewers were required to read and update scores and reviews based on rebuttal
  - A majority of reviewers updated their reviews and scores, and provided post-rebuttal comments

- At least 160 papers discussed extensively online post rebuttal
- External reviewers fully involved in the discussion
Overall Merit Score Categories

1. Very poor (F) -- Fail. The paper is unacceptable.
2. Poor (D) -- Poor-quality paper
3. Average (C) -- Average-quality paper with many deficiencies that are difficult to overlook.
4. Good (B) -- Solid paper with some deficiencies
5. Very good (A) -- High-quality paper that adds insight to the field
6. Excellent (A+) -- Very high-quality paper that exemplifies and/or expands the field
Pre- and Post-Rebuttal OM Scores

- Pre-rebuttal score is the reviewer’s score at the time the rebuttal is exposed to authors

- Post-rebuttal score is the overall merit score at the time the entire review process is over. Affected by:
  - Authors’ rebuttal
  - Online discussion among reviewers
  - Reading of other reviews
  - Discussion during the PC meeting
  - i.e., everything that happens after rebuttal
Distribution of Review Scores Received by Submitted Papers

Pre-rebuttal scores vs. Post-rebuttal scores

Number of Reviews

Very poor | Poor | Average | Good | Very good | Excellent

25
Distribution of Review Scores Received by Accepted Papers

Number of Reviews

Pre-rebuttal scores

Post-rebuttal scores

Very poor
Poor
Average
Good
Very good
Excellent
Magnitude of Change in Average OM Score After Rebuttal vs. Acceptance

- Lowered by more than 1
- Lowered by 0.5 to 1
- Lowered by up to 0.5
- No change
- Raised by up to 0.5
- Raised by 0.5 to 1

Number of Papers

- Rejected
- Accepted

Number of Papers

- Lowered by more than 1
- Lowered by 0.5 to 1
- Lowered by up to 0.5
- No change
- Raised by up to 0.5
- Raised by 0.5 to 1
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 1 – 1.5

- Lowered by 1
- No Change
- Raised by 1

Number of Papers

- Rejected
- Accepted
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 1.5 – 2

![Bar Chart]

- Lowered by 2
- Lowered by 1
- No Change
- Raised by 1
- Raised by 2

Number of Papers

- Rejected
- Accepted
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 2 – 2.5

- Lowered by 2
- Lowered by 1
- No Change
- Raised by 1
- Raised by 2

Number of Papers

- Rejected
- Accepted
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 2.5 – 3

Number of Papers

Rejected
Accepted

Lowered by 2
Lowered by 1
No Change
Raised by 1
Raised by 2

0 5 10 15 20 25 30 35 40
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 3 – 3.5

Number of Papers

- Lowered by 2
- Lowered by 1
- No Change
- Raised by 1
- Raised by 2
- Raised by 3

- Rejected
- Accepted
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 3.5 – 4

![Bar chart showing the number of papers with score changes.]

- Lowered by 2
- Lowered by 1
- No Change
- Raised by 1
- Raised by 2

<table>
<thead>
<tr>
<th>Number of Papers</th>
<th>Rejected</th>
<th>Accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowered by 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowered by 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Change</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raised by 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raised by 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 4 – 4.5

- Lowered by 1
- No Change
- Raised by 1

Number of Papers

- Rejected
- Accepted
Score Changes for Papers w/ Avg. Pre-Rebuttal OM 4.5 – 5

- Lowered by 1
- No Change
- Raised by 1

Number of Papers

- Rejected
- Accepted
Paper Overall Score Distribution, Pre- and Post-Rebuttal

<table>
<thead>
<tr>
<th>Average Overall Merit</th>
<th>Number of Papers</th>
<th>Pre-Rebuttal</th>
<th>Post-Rebuttal</th>
<th>Rejected</th>
<th>Accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00 - 1.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.50 - 1.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00 - 2.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.50 - 2.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.00 - 3.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.50 - 3.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00 - 4.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.50 - 4.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Paper Avg. “Novelty” Score Distribution

Number of Papers

Rejected
Accepted
Paper Avg. “Expertise” Score Distribution

![Bar chart showing the distribution of rejected and accepted papers based on their expertise scores. The x-axis represents different score ranges: 1-1.5, 1.5-2, 2-2.5, 2.5-3, 3-3.5, 3.5-4. The y-axis shows the number of papers. The chart indicates that a significant number of papers fall into the 1.5-2 and 2-2.5 score ranges, with a notable increase in accepted papers as the score increases.]
Paper Avg. “Importance” Score Distribution

![Bar chart showing the distribution of accepted and rejected papers across different score ranges.]
Paper Avg. “Writing Quality” Score Distribution

Number of Papers

1-1.5 1.5-2 2-2.5 2.5-3 3-3.5 3.5-4 4-4.5 4.5-5

Rejected
Accepted
Paper Avg. “Soundness of Ideas” Score Distribution
Paper Avg. “Soundness of Evaluation” Score Distribution

Number of Papers

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Rejected</th>
<th>Accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-2.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-4.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Program Committee Meeting

- **1.5-day meeting**
  - Main goal: avoid hasty decisions. Improve decision quality. ...
  - Extra day enables mulling over decisions. More time allows for more and higher-quality discussions. No “lost” PC members.

- August 25-26, Hilton O’Hare airport
  - All PC members present the entire first day
  - 42/44 PC members present the second day

- 82 papers discussed in (somewhat of a) rank order
  - Not average overall merit score rank

- **Goal: Consensus across the entire PC for accept/reject decision**
  - Entire PC voted when consensus was not clear
Metrics for Paper Discussion Order (I)

- Determined after many different ratings are considered
  - Average Overall Merit (OM)
  - Average OM weighted with expertise (2 ways)
  - Average OM weighted with reviewer generosity
  - Average OM weighted with reviewer expertise and generosity (2 ways)

- Total 6 OM metrics per paper
- 6 different versions of each OM metric
  - [pre-rebuttal, post-rebuttal] x [all reviewers, reviewers with expertise > 1, reviewers with expertise > 2]
- 36 metrics and ranks for each paper
  - A kitchen-sink metric averages all these
  - A correlation analysis would be interesting
Metrics for Paper Discussion Order (II)

- No single best metric
- Plus, many things metrics cannot capture
- We took all the rankings with a *large* grain of salt

- Final discussion order determined by examining all metrics and each paper individually
  - 9 groups of papers formed; 7 of which were discussed

- Bottom line:
  - Top 72 papers in terms of post-rebuttal rank (with non-experts excluded) ended up on the discussion list
  - +7 more with at least one A score but lower averages
  - +3 more papers brought up by reviewers
New Metrics to Rank Papers
Problem with Average Overall Merit

All scores are given the same weight

• Varying expertise
  – Different reviewers have different expertise

• Varying “generosity”
  – Some reviewers more harsh/generous than others
Expert Mean Score

Taking expertise of each reviewer into account

- Idea: Give more weight to scores from experts
- Weight each score by expertise

$$\text{Expert Mean of a Paper} = \frac{\sum \text{Expertise} \cdot \text{Score}}{\sum \text{reviews} \cdot \text{Expertise}}$$

“score” refers to the overall merit
Generosity Mean Score

Take generosity of each reviewer into account

• Idea: A review with a low score compared to other reviews for the paper is considered “less generous”

Generosity of a Reviewer = \[ \text{Mean}_{(\text{all papers})} \left( \frac{\text{Score by the reviewer}}{\text{Mean score for the paper}} \right) \]

Generosity score of a Review = \[ \frac{\text{Review score}}{\text{Generosity of the reviewer}} \]

\[ \sum_{\text{all reviews}} \text{Generosity score of the review} \]

Generosity Mean of a Paper = \[ \frac{\sum_{\text{all reviews}} \text{Generosity score of the review}}{\text{Number of reviews}} \]
Expert Generosity

• Take expertise and generosity into account
• Idea: An expert review with low score is likely not harsh compared to a non-expert review with high score
• Two flavors of expert-generosity
  – Pre-expert generosity
  – Post-expert generosity
Pre-expert Generosity Mean

- Account for generosity before expertise
- Similar to expert mean
- Use generosity score instead of overall merit

\[
\text{Expert Mean of a Paper} = \frac{\sum_{\text{reviews}} \text{Expertise} \times \text{Score}}{\sum_{\text{reviews}} \text{Expertise}}
\]

\[
\text{Pre-expert Generosity Mean} = \frac{\sum_{\text{reviews}} \text{Expertise} \times \text{Generosity Score}}{\sum_{\text{reviews}} \text{Expertise}}
\]
Post-expert Generosity Mean

• Account for expertise before generosity

\[
\text{Expert - generosity of a Reviewer} = \text{Mean} \left( \frac{\text{Score by the reviewer}}{\text{Expert mean score}} \right)
\]

\[
\text{Expert - generosity score of a Review} = \frac{\text{Review score}}{\text{Expert - generosity of the reviewer}}
\]

\[
\text{Post - expert Generosity Mean} = \frac{\sum_{\text{all reviews}} \text{Expert-generosity score of the review}}{\text{Number of reviews}}
\]
### Case Study (for Generosity)

<table>
<thead>
<tr>
<th></th>
<th>Reviewer 1 (Generous)</th>
<th>Reviewer 2 (Neutral)</th>
<th>Reviewer 3 (Harsh)</th>
<th>Mean Average Score</th>
<th>Mean Generosity Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper 1 (Good)</td>
<td>6</td>
<td></td>
<td>3</td>
<td>4.5</td>
<td>4.57</td>
</tr>
<tr>
<td>Paper 2 (Average)</td>
<td>5</td>
<td>3</td>
<td></td>
<td>4</td>
<td>3.38</td>
</tr>
<tr>
<td>Paper 3 (Bad)</td>
<td>2</td>
<td>1</td>
<td></td>
<td>1.5</td>
<td>1.71</td>
</tr>
<tr>
<td>Generosity</td>
<td>1.29</td>
<td>1.04</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Accounting for generosity increases the gap between the good paper that received a harsh review AND the average paper that received a generous review.
Post PC-Meeting

- **Summary statements for discussed and rejected papers**
  - Written by myself or an expert PC member
  - Goal: provide insight/visibility for authors into the discussion of the paper in the PC meeting... Provide major reasons for rejection.
  - Can be done more methodically in the future: assign a scribe for each paper

- **All papers shepherded**
  - Goal: improve quality of final program
  - Did achieve its purpose in many cases
Goal:
- Improve exposure of papers in the presence of parallel sessions
- Enable better technical interactions
- Enable better “Best *” decisions

Lightning session

Poster session
Best Paper Award Process

- 8 papers chosen as candidates based on
  - Rankings (based on 36 different metrics)
  - PC member vetting
  - Reviewer vetting

- 7-person Best Paper Award Committee
  - Not conflicted with any of the 8 candidate papers
  - Can select another paper as the winner
  - Can select 0-N papers
  - The same committee will select the Best Lightning Session Presentation and Best Poster Award winners
    - Sessions to aid the selection process
Feedback

- Any type of feedback on any part of the process is very much appreciated

- Goal: Living document that continually improves the process and institutes best practices

- Methods of providing feedback
  - Online feedback form
    - http://www.surveymonkey.com/s/Z8J6FXT
  - In person
  - Via email
  - Snail mail!
  - ...
Micro-45 Survey

Link on Website
www.microsymposia.org/micro45
Thanks

- Literally hundreds of people
- PC, ERC, Steering Committee
- All reviewers
- All submitting authors
- All presenters
- All attendees
- Many others (see my note)

- Strong and thriving community effort
Thanks

- These slides were prepared in part by
- Vivek Seshadri, Carnegie Mellon
- Chris Fallin, Carnegie Mellon
- Justin Meza, Carnegie Mellon
Thank You.

Onur Mutlu
PC Chair
December 3, 2012
Vancouver, BC, Canada
Vancouver Aquarium

Tonight – Buses leave from 6:45pm to 7:45pm, return from 10pm to midnight.

Extra Tickets Available
Micro 2012
Program Chair’s Remarks

Onur Mutlu
PC Chair
December 3, 2012
Vancouver, BC, Canada
Additional Data
Reviews Categorized by Magnitude of Score Change

- Submitted
- Accepted

Number of Reviews

- Raised by 1
- Raised by 2
- Raised by 3
- Lowered by 1
- Lowered by 2
- Lowered by 3
- No change
Reviews Categorized by Magnitude of Score Change (Detail)

Number of Reviews

- Raised by 1
- Raised by 2
- Raised by 3
- Lowered by 1
- Lowered by 2
- Lowered by 3
- No change

Submitted: 215
Accepted: 1003
Papers Categorized by Post-Rebuttal Avg. Score Change Direction

Number of Papers

Moved up
Moved down
No change
Papers Categorized by Post-Rebuttal Avg. Score Change Direction

(number on top indicates average pre-rebuttal score across all papers in the respective group)