

## Program Committee Chairs Report

Jeff Bilmes, University of Washington - bilmes@ee.washington.edu

Jennifer Chu-Carroll, IBM T.J. Watson Research Center - jenc@us.ibm.com

Mark Sanderson, University of Sheffield - m.sanderson@shef.ac.uk

### 1. Schedule

December 16, 2005 Full Paper submissions due

February 17, 2006 PC meeting

February 23, 2006 Full Paper notification of acceptance

March 3, 2006 Short Paper submissions due

April 6, 2006 Short Paper notification of acceptance

April 17, 2006 Camera-ready full/short papers due

June 4-9, 2006 Conference

### 2. Overview remarks

(Note that this 2006 report is adapted from the structure of the 2004 post-conference report written by Dumais, Marcu and Roukos.)

The co-chairs represent the three main fields covered by HLT/NAACL 2006 - Jeff Bilmes (Speech), Jennifer Chu-Carroll (NLP) and Mark Sanderson (IR). We generally divided tasks such as suggestions and recruiting of reviewers, publicity, and assignment to area chairs by discipline. For those tasks that cut across the disciplines, for some tasks we found the work could be divided equally among us, such as determining the final schedule, for others we assigned someone to lead the work: Jennifer, organization of PC meeting; Jeff, review software; Mark, final report. We think that this arrangement worked well in general, but we do recognize that because NLP dominates the programme of the conference, Jennifer Chu-Carroll ended up with more work to do in each of the shared tasks.

### 3. Paper reviewing process

We think the paper reviewing process went very well. The quality of full-paper and short paper submissions was high and submissions were up on previous years. We agree with the 2004 PC chairs' view that this is largely due to the quality of makeup of the program committee, and recognition of HLT/NAACL as an outlet for good work at the intersection of NLP, IR and Speech. It is clear though that for IR and Speech, only work at that intersection (e.g. QA, language models for Speech) is being submitted in numbers to HLT/NAACL, more core IR or core Speech papers were not. Particularly for IR this can in part be attributed to ACM SIGIR's submission deadline being just 6 weeks later than HLT/NAACL's deadline.

As in 2004, reviewing was done using a two-tiered system, Senior Programme Committee (SPC) members and Reviewers. Twenty eight SPC were responsible for a topical area and coordinated the reviewing process (recruiting reviewers, assigning papers to reviewers, managing reviews and attending the PC meeting) in those areas. The SPC are listed at the end of this section. The same review committee handled both Long papers (8 pages) and Late-breaking papers and posters (4 pages). The Co-Chairs made an initial assignment of submissions to the SPC.

There was a face-to-face PC meeting for Long papers. Despite the different disciplines, at the review meeting, a common view on what was acceptable emerged quickly at the meeting. Final decisions for Late-breaking papers were made by conference call. Some Late-breaking

papers were selected to be presented orally and others as posters. Reviewing for both Full and Late-breaking papers was blind.

Long paper reviewing was carried out over the Dec/Jan holidays with a period of discussion in early February. Late-breaking papers were due after the decisions for Long papers were announced, thus allowing people to resubmit if desired (some were encouraged to do so by the initial reviewers). We found the reviewing process particularly the system of decentralizing management of reviews to the Senior PC worked very well.

Unlike the experiences of the 2004 PC chairs, we found the reviewing software to work well, like in 2004, we also used the START conference reviewing software and found it to fit the work flow of our reviewing process well. There was excellent customer support.

Based on the distribution of submissions from 2004, we selected twenty eight SPC. Some chairs covered the same nominal area, which we did to balance the anticipated load. There was a slight over recruitment of SPC chairs in IR, here load was balanced across long and short papers, where some chairs took more of a role in the short paper submissions and others in the long. In general the load balancing worked well, and we tried to assign no more than 15 Long Papers and 10 Short Papers to each SPC member.

After the PC meeting, SPC Chairs were invited to nominate a best paper candidate from the papers submitted in their field. A total of seven nominations were received. Within those nominations were four papers first authored by students. The nominated papers were split into two short lists for best paper and best student paper. All SPC chairs were canvassed for their opinions on the papers and votes from them were tallied. Best student paper had a clear winner, for best paper, there was a tie and as an additional complication, one the papers was co-authored by an SPC. Therefore, the tied papers were read by one of the program chairs and a conference call was held between them to decide on the eventual winner.

The Best Paper Award for HLT/NAACL 2006 went to: Probabilistic Context-Free Grammar Induction Based on Structural Zeros by Mehryar Mohri (Google Research) and Brian Roark (OGI and OHSU)

The Best Student Paper Award went to Prototype-Driven Learning for Sequence Models by Aria Haghighi and Dan Klein (UC Berkeley).

We invited two Keynote Speakers who together we thought best encompassed the different disciplines of the conference. The speakers were Joshua Goodman (Microsoft Research), Email and Spam and Spim and Spat; and Diane Litman (University of Pittsburgh) Spoken Dialogue for Intelligent Tutoring Systems: Opportunities and Challenges.

The Senior Program Committee, Affiliations were:

Johan Bos     University of Roma "La Sapienza"  
Jamie Callan     Carnegie Mellon University  
Joyce Chai     Michigan State University  
Jason Eisner     Johns Hopkins University  
Mark Gales     Cambridge University  
Fredric Gey     University of California Berkeley  
Roxana Girju     University of Illinois Urbana-Champaign

Mark Hasegawa-Johnson University of Illinois Urbana-Champaign  
Julia Hirschberg Columbia University  
Alon Lavie Carnegie Mellon University  
Wei-Ying Ma Microsoft Research Asia  
Mehryar Mohri Courant Institute of Mathematical Sciences & Google Research  
Marius Pasca Google  
Gerald Penn University of Toronto  
Dragomir Radev University of Michigan  
Owen Rambow CCLS, Columbia University  
Steve Renals University of Edinburgh  
Stefan Riezler Google  
Rohini Srihari SUNY Buffalo  
Amanda Stent SUNY Stony Brook  
Michael Strube EML Research  
Christoph Tillmann IBM T.J. Watson Research Center  
Peter Turney National Research Council Canada  
Ellen Voorhees NIST  
Ralph Weischedel BBN Technologies  
Fei Xia University of Washington  
ChengXiang Zhai University of Illinois Urbana-Champaign  
Ming Zhou Microsoft Research Asia

#### 4. Summary of paper quality and acceptances

The number of submissions increased substantially over the 2004 conference, and the quality of submissions was excellent. We received 257 submissions for full papers, of which 62 were accepted (25% acceptance rate). We received 127 submissions for the late-breaking papers track, of which 52 were accepted (41%).

#### 5. Publications

In the 2004, the PC Chairs highlighted problems with coordinating responsibilities and information with the publication chairs. Assuming that changes were put in place since 2004, we found this process to run well in 2006, although there was some confusion between the chairs and others in the organizing committee who was responsible for some of the tasks, such as setting the schedule and recruiting of session chairs.

#### 6. Areas for Improvement

HLT/NAACL is still in its infancy, but it is certainly pleasing to have a strong growth in submissions from the last running of this conference from three years ago. We used the knowledge of all three chairs to solicit submissions from as many parts of our three communities as possible. While there were a many interdisciplinary papers submitted and accepted, it is certainly the case that much of the conference's program was dominated by NLP-centered topics with little direct interest to IR and Speech research. We think this make up of the conference also reflects the view of the conference in the three research communities: it isn't viewed as a forum for core IR or Speech papers.

However, if it is the goal of HLT/NAACL to improve the numbers of IR and Speech papers submitted, quite how this situation can be improved isn't entirely clear to us. For example, within the IR field, the 2006 PC chair recruited well known members of the IR community to the SPC covering a wide range of topic areas and advertised the conference extensively. However, because of the dates of the conference, submissions were being solicited in the same time-frame of SIGIR, which harmed submission numbers and the perception of the conference in IR as a forum for "NLP-related IR" research persists.

It has been noted that acceptance rates in the IR and Speech areas appear to have been tougher than in the NLP areas. However, both Speech and IR PC chairs have reviewed the papers that failed and concluded that the papers were just not good enough to be acceptable to a conference such as HLT/NAACL.

## 7. Profiles of Submissions

Below we summarize the makeup of the Full and Short paper submissions, in terms of international representation and in terms of topic distributions.

Full Submitted	Full Accepted	Short Submitted	Short Accepted	Area
11	2	4	1	Speech Acoustic Modeling
18	4	17	7	NLP Discourse/Dialogue/Multimodality
19	5	7	2	IR General Information Retrieval, including Web Search
21	4	9	7	NLP Generation/Summarization
39	9	15	7	NLP Information Extraction
12	1	14	3	NLP/Speech Language Modeling
18	3	9	4	IR Language Processing for Information Retrieval
7	2	3	2	Speech Lexical/Pronunciation Modeling
26	10	15	5	NLP Machine Translation
37	12	12	7	NLP Parsing/Grammar/Morphology
20	4	9	2	NLP/IR Question Answering
33	7	12	4	NLP Semantics

### Submitted full paper keywords

#	Keywords	#	Keywords
3	Acoustic Modeling	2	Multi lingual speech recognition
21	Corpora	2	Multimodal representations and processing
1	Cross language information retrieval	3	New Approaches to ASR
6	Dialogue systems	23	NLP applications
13	Discourse	28	Parsing
9	Discriminative Training	1	Phonology
10	Evaluation	23	Question answering
2	Formal Models for IR	1	Rich transcription
20	Information extraction	5	Semantics
3	Information retrieval models	1	Sentiment analysis
3	Interactive IR	9	Speech based interfaces
1	Language identification	1	Speech recognition
6	Language modeling	8	Text alignment
5	Language resources	5	Text generation
1	Learning techniques for language processing	1	Text mining
2	Lexical and knowledge acquisition	15	Text summarization
3	Lexical/Pronunciation Modeling	3	Topic Detection and Tracking
13	Machine translation	1	Web information retrieval
2	Mathematical models of language	3	Word sense disambiguation
2	Morphology		

### Full papers, countries (of contact author only):

3	Australia	1	Jordan
---	-----------	---	--------

2	Bangladesh	2	Netherlands		
1	Brazil	1	Nigeria		
5	Canada	1	Pakistan		
6	China	1	Poland		
4	Czech Republic	1	Portugal		
1	Denmark	3	Republic of Korea		
2	Finland	1	Russian Federation		
5	France	2	Singapore		
11	Germany	6	Spain		
2	Hong Kong	Special Administrative Region of China		1	Sweden
5	India	1	Taiwan Province of China		
2	Iran (Islamic Republic of)	1	Thailand		
2	Ireland	3	Turkey		
3	Israel	18	United Kingdom		
5	Italy	147	United States		
12	Japan				

Short Paper Keywords

3	Acoustic Modeling	1	Multi lingual Retrieval		
11	Corpora	1	Multimodal representations and processing		
5	Dialogue systems	15	NLP applications		
11	Discourse	1	New Approaches to ASR		
1	Discriminative Training	9	Parsing		
5	Evaluation	1	Psycholinguistics		
1	Evaluation in IR	12	Question answering		
6	Information extraction	2	Sentiment analysis		
2	Information retrieval models	6	Speech based interfaces		
2	Language identification	2	Speech summarization		
7	Language modeling	1	Text alignment		
3	Language resources	5	Text summarization		
1	Learning techniques for language processing	1	Topic Detection and Tracking		
8	Machine translation	1	Web information retrieval		
2	Morphology				

Short papers, countries (of contact author only):

1	Brazil	17	Japan		
9	Canada	1	Netherlands		
3	China	1	Portugal		
1	Finland	2	Republic of Korea		
4	Germany	1	Singapore		
1	Greece	3	Spain		
4	Hong Kong	Special Administrative Region of China		1	Sweden
1	Hungary	2	Switzerland		
5	India	1	Taiwan Province of China		
1	Iran (Islamic Republic of)	1	Thailand		
2	Ireland	7	United Kingdom		
2	Israel	54	United States		
1	Italy				

Roll Ups by Area (NLP, IR, Speech)

Area	Full (accepted)	Short (accepted)
Speech	4	3
NLP/Speech	1	3
NLP	46	37
NLP/IR	4	2
IR	8	6

