

ACL 2007  
REPORTS  
FROM  
THE  
ORGANIZERS

ACL Secretary report

Dragomir Radev  
radev@umich.edu

A number of new policies were adopted, including a policy to add IR and Speech papers to the ACL conference and to include short papers not unlike the NAACL HLT conferences.

A call for new initiatives was sent to the membership. A number of interesting proposals have come through. The deadline is June 4.

A newsletter was sent out in the Winter.

Priscilla's pay raise was discussed and approved.

The ACL handbooks and related materials were moved from Delaware to the ACL web site.

The Webmaster, Ali Hakim, reorganized a number of sections of the site, including policies and conference organization.

A Wiki for ACL was created at the initiative of Peter Turney. Ali Hakim helped set it up.

Minutes from the most recent executive meetings (Sydney 2006 and Winter phone conference 2007) were distributed to the exec.

The Lifetime achievement award winner for 2007 was picked and invited to attend ACL 2007 to receive the award.

The archives of the ACL were reorganized and updated, thanks to the archivist, Eric Fosler-Lussier.

An initiative to organize all ACL resolutions on the web was started. This will be a time consuming endeavor.

=====

EACL Report, May 2007

=====

Alex Lascarides, Chair  
Giorgio Satta, Chair-elect  
Anette Frank, Secretary  
Mike Rosner, treasurer

1. ACL 2007 in Prague  
-----

There has been some confusion among the EACL board about ACL policies for joint conferences. We were in fact unaware that EACL and ACL are supposed to be jointly hosting ACL 2007 until May!

We are grateful that the role we should take when ACL is in Europe has now been clarified, and we hope to have procedures in place to ensure that we contribute more to early stages in organising a conference (e.g., offering feedback on the bids to host ACL and contributing to decisions about PC chairs).

Gertjan van Noord, Mike Rosner, Anette Frank and Alex Lascarides are on the organisation committee for ACL 2007.

2. EACL 2009  
-----

We have chosen to hold EACL 2009 in Athens, Greece, from March 30th to April 3rd. The structure is as follows:

March 30th: Tutorials and workshops  
 March 31st: Tutorials and workshops  
 April 1st: Main conference  
 April 2nd: Main conference  
 April 3rd: Main conference

This will be the main agenda item for our EACL board meeting, to be held in Prague. We hope to have a shortlist of names for general chair and PC chairs as an outcome of this meeting.

3. EACL Sponsorships  
 -----

EACL expects to sponsor students attending ESSLLI 2007 in Dublin, valuing 3000 Euros in total.

4. Financial report:  
 -----

To be supplied by Mike Rosner

The above figures are not yet finalised. We are still waiting on the final outcome of ACL 2004, which made a loss.

5. EACL officers: Elections  
 -----

Elections were finally held for new candidates on the board, three months late! We welcome Giorgio Satta as the new chair-elect, and Eric Gaussier, Kiril Simov and Josef van Genabith onto the advisory board. Gertjan van Noord and Felisa Verdejo have become members of the nominating committee.

The current EACL is as follows:

Chair:  
 Alex Lascarides University of Edinburgh  
 Chair elect:  
 Giorgio Satta University of Padua  
 Treasurer:  
 Mike Rosner University of Malta  
 Secretary:  
 Anette Frank DFKI Saarbrücken  
 Advisory Board:  
 Eric Gaussier Joseph Fourier University, Grenoble  
 Paola Merlo University of Geneva  
 Kiril Simov Bulgarian Academy of Sciences, Sofia  
 Josef van Genabith Dublin City University  
 Nominating committee:  
 John Carroll University of Sussex  
 Claire Gardent LORIA, Nancy  
 Gertjan van Noord University of Groningen  
 Felisa Verdejo Ciudad Universitaria Madrid

We would like to thank everyone who has come to the end of their tenure for all the work and effort they have put into ensuring that EACL runs smoothly: Gertjan van Noord (Chair, 2005--2007), Galia Angelova, Ido Daga and Felisa Verdejo (advisory board), and Philippe Blache and Donia Scott (nominating committee).

EACL Student Board  
 -----

The EACL has elected two new members of the Student Board, who will

take up their service in July 2007. They replace Jonathon Read and Violeta Seretan:

Vera Demberg, University of Edinburgh, UK and  
 Yanjun Ma, Dublin City University, Ireland.

We would like to thank Jonathon and Violeta for their sterling work over the last two years.

=====

EACL Treasurer

2006 Financial Report

Mike Rosner

This report describes ACL assets held in European accounts. A separate consolidated statement of European Chapter finances appears with European Chapter report.

The ACL accounts are described as follows:

1. CHF-CH-CURRENT 2006  
Swiss franc account held in Switzerland
2. EUR-CH-CURRENT 2006  
Euro current account held in Switzerland - no interest
3. EUR-CH-SAVINGS 2006  
Euro savings account held in Switzerland - interest bearing
4. EUR-MT-CURRENT 2006  
Euro savings account held in Malta - interest bearing

5. CONSOLIDATED ASSETS HELD IN EURO AND CHF BY YEAR  
 Assets held in CHF denominated bond funds (held in Switzerland) and EUR denominated bonds (held in Malta). These are shown by year in euro equivalent.

1. CHF-CH-CURRENT 2006				
=====				
1 EUR =	1.65614	CHF	CHF	EUR EQUIV
31/12/2005		1,295.73	1,295.73	782.38
INCOME				
	pubs+dues	213.30		
	interest	1.15		
	total income	214.45	214.45	129.49
EXPENSES				
	bank charges	293.05		
	total expenses	293.05	-293.05	-176.95
31/12/2006		1,217.13	734.92	

Mike Rosner, 22/05/2007

=====

2. EUR-CH-CURRENT 2006  
 =====

31.12.2005	EUR	3,145.48
------------	-----	----------

EXPENSES  
 EACL06 PC charges -2,709.69  
 -129.26  
 31.12.2006 306.53

Mike Rosner 22.05.2007

3. EUR-CH-SAVINGS 2006

31/12/2005 10,083.40

INCOME  
 Interest 95.90

EXPENSES  
 withholding tax -33.57

31/12/2006 10,145.73

4. EUR-MT-CURRENT 2006

31/12/2005 43,522.77

INCOME  
 TFR French ac 25,603.56  
 DUES 9,132.53  
 INTEREST 229.83

subtotal 34,965.92 34,965.92

EXPENSES  
 Office Expenses -1,000.00  
 EACL Foundation -101.87  
 Bonds -50,331.40  
 EACL06 PC MTG -1,680.74  
 Tutorials -2,440.16  
 Inv. Spkrs -2,627.51

Sponsorship  
 EACL06 -2,127.68  
 ESSLI -3,000.00  
 SIGNLL -288.43  
 Lexicom -500.00  
 CHARGES -117.67

subtotal -64,215.46 -64,215.46

31/12/2006 14,273.23

Mike Rosner 22/05/2007

5. CONSOLIDATED ASSETS HELD IN EURO AND CHF BY YEAR

YEAR	EUR EQUIVALENT
2000	152,479
2001	159,427
2002	172,111
2003	177,101
2004	223,449
2005	215,953

2006 202,166

Mike Rosner 22/05/2007

NAACL Treasurer's report

Chris Manning

April 2007

The Chapter's bank account held \$31,609.34 at the end of Feb 2007, and hasn't changed much since then. As usual, delays in conference accounting mean that this number is far from the Chapter's true financial situation. The biggest news is that I've basically finished accounting for HLT-NAACL 2004, which will yield a profit of approximately \$41,000 and a consequent transfer of about \$43,000 to the Chapter's bank account. The ACL Treasurer has not yet finished accounting for the ACL2005 conference, from which we get 50% of the profit/loss. I'm going to try to do HLT-NAACL 2006 in the latter half of the summer.

Major recent and pending expenses are: JHU 2006 (\$8,755), LSA 2007 (\$11,750), NAMCLO (\$2,000), and JHU 2007 (\$4,350). The chart below shows the NAACL account balance from since I became treasurer, projecting slightly into the future; last year's report shows the rest of NAACL's financial history. There is no reason why NAACL cannot continue to operate financially just as it has for the last few years. If there are opportunities for improvement, they seem to lie in: - Developing better policies and procedures for allocating money. This requires us to get more organized earlier. - Being more proactive in raising income, by exploiting commercial connections, and other fund raising opportunities.

Christopher Manning, NAACL Treasurer; April 21, 2007, Palo Alto, CA.

ACL Nominating Committee Report

Johanna D. Moore

Here is what we have so far (and have had in place for over a month).

For ACL Exec to replace KY Su:

- Yuji Matsumoto;
- Hwee Tou Ng.

For ACL Vice-President Elect:

- Jan Hajic;
- Ido Dagan.

For ACL Treasurer:

- Graeme Hirst;
- Alon Lavie.

All of the six of the above have agreed to be nominated.

ACL 2007 General Chair Report

John Carroll

Firstly, I would like to thank all the people involved with organising ACL this year. All of them carried out their roles to the full, were extremely competent in everything they did, and were responsive and constructive in discussions about any problems that cropped up.

I have thanked them by name in the General Chair Preface in the conference proceedings (and I hope I have not inadvertently left anyone out). So in this report I will only highlight a few issues that warrant some thought for next year, or things that have been done differently than in previous years.

#### 1. Unreliability of email

One tutorial proposal got held up on a mail server and did not reach the tutorials chair -- until he was prompted to look for it by the proposer after the notification date. The problem was resolved by asking the proposer to submit a new version, which was passed for comments to the two reviewers who were most likely to have been given the proposal had it originally been received before the deadline.

Lessons learned: either use START for tutorial (and perhaps workshop) proposals, or stick with email submission but add a step to the process "Acknowledgment of receipt" with a published date a couple of days after the submission deadline, so that a proposer knows to expect an acknowledgment and when it should arrive by.

#### 2. Co-location of other, major workshops/conferences

As in previous years, decisions about co-location of major ACL SIG workshops/conferences (e.g. IWPT, EMNLP) have to be taken before those for standard workshops, since their schedules are on a longer timescale. The same goes for workshops that include shared tasks; an example of this is the SIGLEX-sponsored SemEval which submitted a workshop proposal early, before the call for workshops proposals went out.

This year, the discussion about co-location of EMNLP-CoNLL was made harder by the fact that the EMNLP organisers had talked to a number of people connected with ACL 2007 at different points over the previous months, and it took a while to piece together who had said what to whom. I suggest that in future all such communication be directed through the workshops chair (if they have been appointed by then) and the general conference chair.

Another issue is that the EMNLP schedule has a strong interaction with the main conference schedule since EMNLP wants to have their submission deadline at least 1 day after the main conference notification date. This dependency needs to be taken into account at an early stage in planning the conference, otherwise the main conference notification date could end up being too late for EMNLP to manage their submission process.

Another dependency: we set a recommended submission date for workshops to be after the main conference notification date -- so that workshop PCs would not have to deal with submissions that were also under consideration for the main conference.

This year, the only other major conference in our field with a similar reviewing period was AAI; there were a small number of double submissions which the PC chairs sorted out with their AAI counterparts.

#### 3. Birds-of-a-feather meetings

Birds-of-a-feather (BoF) meetings are short informal gatherings for researchers with a common interest, often used by SIGs for their business meetings, or for people in new research areas to plan future workshops.

There are several rooms available at ACL 2007 for this purpose, at lunchtimes during the entire conference/workshop stretch. Priscilla Rasmussen has been coordinating requests for BoF meetings and assigning rooms.

#### 4. The reviewing process

There were complaints from the authors of three submissions about the quality of reviews they received. In two cases their complaints had some substance. The authors of one submission compared the ACL reviewing process to that of AAI, in which authors have an opportunity to write a response to reviewers' comments before acceptance/rejection decisions are made. Future ACL conferences could perhaps consider adopting this model.

#### 5. Workshops

There was an email discussion between the organisers about whether participants should be allowed to register for overlapping workshops; also, how often it happened that people registered for a 2-day workshop and then instead mainly attended 1-day workshops in that 2 day period, thus saving a few dollars in registration fees.

This is probably the first ACL with workshops lasting only half a day. These slots were given either (i) to new, emerging areas which sounded attractive and exciting, but for which it was not clear that there would be sufficient submissions to fill a whole day, or (ii) to more established areas where a similar workshop had run fairly recently so the amount of new research might be limited. This idea needs to be evaluated to see whether it should be tried again in future.

#### 6. Submissions and final versions of papers

The Publications Chair will probably comment on the issues below, but I want to flag them as well:

\* Depending on how they are produced, PDFs can contain embedded information that identifies the authors; this is obviously undesirable for submissions. Perhaps in such cases the START system could be configured to warn submitters and ask them to upload a new version without this information?

\* There were negotiations about the maximum numbers of pages for workshop papers, at submission and for final versions. The previous limit of 8 pages is less valid now that so few hardcopies of workshop proceedings are sold.

\* The requirement for final versions of papers in US letter size caught out many authors. Perhaps the STRRT system could again be customised to check PDFs automatically for page size when they are uploaded?

John Carroll  
May 2007

=====

Report of the Program Chairs of ACL-2007

Antal van den Bosch (Tilburg University, The Netherlands)  
Annie Zaenen (PARC, USA)

Date: 21 May 2007

## 1. Planning and preparations

Planning and organisation of the reviewing process for the main session of ACL-2007 commenced, largely via email between the two program chairs and with additional communication with the ACL exec and general chair John Carroll, in August 2006. The following areas and area chairs were selected in the Fall of 2006:

Lexicon, lexical databases, ontologies, language resources:  
Tim Baldwin (University of Melbourne, Australia)

Summarization, generation:  
Kees van Deemter (University of Aberdeen, UK)

Pragmatics, dialog systems, discourse:  
Barbara Di Eugenio (University of Illinois at Chicago, USA)

Syntax, parsing, formalisms:  
Josef van Genabith (Dublin City University, Ireland)

Question answering, information extraction, information retrieval:  
Claire Grover (University of Edinburgh, UK)

Semantics, lexical semantics, formal semantics, logic, textual entailment:  
Diana McCarthy (University of Sussex, UK)

Machine learning, algorithms for NLP  
Dan Roth (University of Illinois at Urbana-Champaign, USA)

Phonology, morphology, FS technology, tagging, word segmentation  
Richard Sproat (University of Illinois at Urbana-Champaign, USA)

Speech, language modeling, spoken dialog systems  
Marc Swerts (Tilburg University, The Netherlands)

Machine translation  
Andy Way (Dublin City University, Ireland)

On top of the standard invitation and instructions, we asked the area chairs to urge their reviewers to pay attention to the originality of the submissions, to avoid the observed trend of going for safe and sound but somewhat incremental work.

As submissions came in near deadline time, it became clear for most areas that having about 25 reviewers would be too low for most areas to keep the burden of reviewers at six papers, so more reviewers were sought and added around the deadline date, to a total of 332 (i.e. over 33 on average per area). Some reviewers reviewed up to 11 papers but that was very rare; the average was 5.25, and the mode 7.

Florence Reeder provided mentoring services for 17 submissions. She was assisted in the mentoring service by Charles Callaway, John White, Bea Oshika, Ken Samuel, Deborah Dahl, Marilyn Kupetz, and Chrys Chrystello.

Analogous to the scheme used for COLING-ACL 2006, we adopted a bidding phase in which reviewers could voice preferences for reviewing particular papers. This went smoothly and was judged positively overall.

The schedule for the whole process was (and was kept at) the following:

Wed Jan 10	reviewers recruited by area chairs
Tue Jan 23	submissions arrive
Wed Jan 24	papers assigned to primary track/area
Thu-Mon Jan 25-29	reviewer bidding phase
Tue Jan 30	initial assignment of reviewers to papers using START
Thu Feb 1	final assignment of reviewers to papers by area chairs
Fri Feb 2	reviewing stage begins
Sun Mar 11	reviews due
Mon-Mon Mar 12-19	e-mail/START discussion amongst reviewers and area chairs on "disagreement" papers
Mon-Thu Mar 19-22	area chair discussion/decisions on accept/reject
Fri Mar 23	author notification
Fri May 4	camera ready due
Sun-Fri Jun 24-29	ACL-2007 conference

The first call for papers was released around September 5, 2006. A second call was released around December 1, 2006. The call was also in ACL newsletters sent around regularly by Priscilla Rasmussen, and was publicized on the conference web page.

## 2. Submission and Review Process

The submission deadline was January 23, 2007 (5pm US Eastern time, 10pm GMT). Notification of acceptance was set at March 23, 2007, after discussion with the organizers of EMNLP-CoNLL who wanted a little room between ACL's notification date and their submission deadline (March 26), and sufficient room for their reviewing process. Camera ready papers were due on May 4, 2007 and the final program was sent to the publications chair about a week later.

We decided not to have an area chair meeting. Communication with the area chairs (mostly through email) was unproblematic throughout the entire process. We were confident that by maintaining a steady level of detail, clarity, and speedy feedback in our email communication, the decision making process could be performed using the same channels. For the final decision making process, the two co-chairs met for two days at Tilburg University, and interacted during that time vigorously with the area chairs by email and sometimes by phone.

The program committee's selection of 131 papers was based on 588 submissions, a new record in an increasing trend. The original total number of submissions was 605. The difference of 17 papers is explained by 12 papers being withdrawn (mostly due to double submissions, where authors opted for presentation at another venue, while ACL reviewing was still ongoing), and by an unfortunate number of 5 papers which did not comply with the anonymization guidelines (after the authors were warned twice about this and were given a brief period to correct the anonymity of their submissions).

All submissions received three reviews. A significant effort was delivered by the area chairs who managed to get all of the reviews completed in time for the final decision making process.

At decision making time we made the following steps. We asked all area chairs to divide their papers into `accepts`, `borderline`, and `reject` papers. We tentatively accepted all `accepts`, totaling 100 papers. We then used the START system to rank all remaining papers according to a formula that used the overall recommendation score as a basis, and added the originality scores to break the ties (using the START formula "Recommendation + (0.2 \* Originality)"). We ranked all `borderline` cases, and took a temporary threshold of 4.27. This threshold marked 130 papers, which was the total number of submissions that were recommended with priority by the area chairs (the `accepts` plus any `borderline` papers that they particularly recommended). We dropped `borderline` papers above that threshold that were not recommended, and added papers from below the threshold that did, resulting in a list of 131 papers to be accepted. We proposed this list to the area chairs, who all agreed.

In the week leading to the notification date we decided to plan for four parallel sessions. During decision making time it became clear that even then we would not be able to accommodate more than about 110 papers if we would use 30 minute slots. We considered a lottery that would reduce some papers to posters, but finally we decided to reduce the time slots to 25 minutes, which was just enough to cover the 131 accepted papers. In sum, we managed to get the acceptance rate up to 22.3% and to accept all the papers on the area chairs' priority

lists. Yet, to accept all the papers that were deemed worthy of being presented at the ACL we would have had to accept around 150 papers. See the table below for the distribution of the admissions and accepted papers per area.

area	submissions		
	#	#accept	%accept
Discourse, Dialogue and Pragmatics	57	16	28
Syntax, Parsing and Formalisms	63	15	24
Lexicon, ontologies and resources	63	11	17
Morphology, etc.	60	13	22
Machine Learning, algorithms for NLP	67	16	24
IE, IR, QA	70	15	21
Machine Translation	65	15	23
Speech, Language Modeling	41	9	22
Generation and Summarization	48	10	21
Semantics etc.	54	12	22
Total	588	131	22.3

The geographical distribution of the first authors of all submissions is shown below. We received submissions from 45 countries. Asia & the Pacific region accounted for 35% of all submissions, closely followed by Europe also with 35%; 27% came from the North American continent, 2% from the Middle East, and less than 1% from South America and Africa.

Continent	#subm	#acc
-----------	-------	------

Asia & Pacific=====

11 countries

205/584 = 35.1% of submissions; 22/131 = 16.8% of accepted

Australia	18	4
China	74	6
Bangladesh	1	
Hong Kong	10	
India	11	
Japan	53	3
Republic of Korea	6	
Singapore	10	8
Sri Lanka	1	
Taiwan	20	1
Thailand	1	

Europe=====

23 countries

204/584 = 34.9%; 41/131 = 31.3%

Austria	1	
Belgium	4	1
Czech Republic	6	1
Denmark	3	
Estonia	2	
Finland	3	1
France	21	4
Germany	32	6
Greece	3	
Hungary	2	
Iceland	2	
Ireland	10	5
Italy	10	3
Netherlands	11	3
Poland	1	
Portugal	3	

Romania	3	
Russian Federation	1	
Spain	20	
Sweden	8	1
Switzerland	5	2
Turkey	4	
United Kingdom	49	14

South America =====

3/584 = <1%

Argentina	1	
Brazil	1	
Mexico	1	

North America =====

157/584 = 26.9%; 61/131 = 46.6%

Canada	11	4
United States	146	57

Middle East =====

2 countries

11/584 = 1.9%; 6/131 = 4.6%

Iran	1	
Israel	10	6

Africa =====

2 countries

5/584 = <1%

Egypt	3	
South Africa	1	
Tunisia	1	

### 3. Conference management system

As in previous ACL conferences, we used the START system to manage submissions, bidding, and reviews. Things generally went well and we got good support from Rich Gerber. Most minor problems were dealt with quickly, and suggestions for minor improvements were either immediately implemented or noted for later implementation.

### 4. Issues

The most salient issues that came up during the process and during reflective discussions were the size of the whole process, conflicts of interest, and dealing with dual submissions (both simultaneous and serial submissions).

As for the size of the process, some flexibility was required from the program and area chairs when it turned out that we would need to process near 600 submissions, hence would need to have the reviewing capacity for close to 1,800 reviews (in the end, 1,764 reviews were written). Near future ACL conferences should expect similar numbers of submissions. With ten areas it would have been optimal if the area chairs would have been instructed to recruit about 30-35 reviewers from the start. For the future this should be the norm, unless more areas are discerned. We went for the ten areas listed, roughly following the EACL-2006 and ACL-2005 division of areas. This division worked well, although we observed some overlap notably between the three areas of semantics and lexical semantics, information extraction and question answering, and discourse and dialog.

For the program chairs the large size of the program also meant a high frequency of relatively small issues to handle: interactions with authors about their submission, START issues (as all ACL events were handled though START, where co-chair Van den Bosch handled the initialization of START pages for all fifteen workshops and co-located events) or merely side issues such as requests for more information.

The issue of conflicts of interest arose to a mild degree due to the fact that most area chairs were co-authors of papers; one area chair co-authored five submissions. Inevitably these papers were handled by other area chairs, who had to carefully select reviewers for these submissions. Although we have not seen evidence of this process failing, there is an increased risk of errors. We believe that area chairs should be made aware of this issue. At the same time we have the impression that if area chairs would be advised against submitting papers, it would be more difficult to recruit competent area chairs. The program co-chairs did not co-author any submission.

Dual submissions present a problem as authors are not always forthcoming with information, and as they tend to have different ideas about what constitutes an almost identical paper. After some discussion we resorted to having the authors of double submission tick an option on the submission page to declare that the paper was submitted elsewhere or was going to be submitted elsewhere during the reviewing period. If ticked, submitters were also asked to fill in a text box stating to which other conference or journal the paper was submitted. Although as many as 43 authors provided this information, we do not think that the problem has been solved completely satisfactorily. At decision time we discovered that two papers that were shortlisted for acceptance had already been presented and published elsewhere - these papers were subsequently rejected. One more paper was already scheduled for presentation at a workshop but was retracted from that workshop in favor of ACL.

We discussed the problem with the area chairs, who did not express worries so much towards simultaneous double submissions (on which the ACL guidelines should be upheld), but rather at serial double submissions to other events before and after ACL. The general opinion was that a better coordination with the programme chairs of conferences such as NODALIDA, NAACL, EMNLP, IWPT and CoNLL is necessary, so that (information on) reviewers and reviews may be shared, of course without compromising independence and fairness to the authors who want of course as many chances as possible, and who in fact may improve a rejected paper before resubmitting it elsewhere. Information on papers and reviewers was indeed shared with Jason Eisner, program chair of EMNLP-CoNLL.

Another proposal that came up in the course of the discussion with the area chairs was to ask experimental papers to make the code, data and scripts available to the reviewers.

#### 4. Best Paper Award

This year's conference continues the tradition of recognizing one of the submitted papers with the Best Paper award. It will be selected by the ten area chairs, the two program chairs, and the general chair, based on a shortlist of papers proposed by the area chairs.

#### 5. Acknowledgments

We thank John Carroll, General Conference Chair, the Local Arrangement Committee headed by Eva Hajicova, and the ACL executive, especially Dragomir Radev, for their help and advice, and last year's co-chairs, Claire Cardie and Pierre Isabelle, for sharing their experience. Our

sincere thanks go to Su Jian for putting together the proceedings. We thank Femke Wieme and Lauraine Sinay of Tilburg University for checking paper formatting issues.

=====  
ACL 2007 Local Organizer Committee Report  
Eva Hajicova, Jan Hajic, Anna Kotesovcova (as of May 22, 2007)

The LOC has now 12 members, all of whom are either staff members or PhD students of the organizing institute (UFAL). Thanks to the devotion and effectiveness of the members of LOC, the organization proceeds without serious problems. The unexpected problems are being solved on the run, with the kind help of the responsible chairs/colleagues and, of course, of Priscilla.

The accommodation reserved up to now in the TOP hotel reaches the number of 809 beds in total, and there are 74 beds reserved already in the students dormitory. We know that some tens of people have reserved beds in centrally located hotels in Prague so that the expected number of participants is at least 800.

Additional events - ACL Exec Meeting, PC Dinner, CL Editorial Board Meeting, Student lunch, ACL Business Meeting, EAACL Business Meeting and ACL 2007 Debrief Meeting are prepared.

The social and cultural programme (welcome party, banquet, concert) as well as the ACL EXEC+ dinner is also already arranged for.

=====  
ACL-07 Tutorials Chair Report  
Joakim Nivre

In response to the call for tutorial proposals, 19 proposals were received before the deadline of December 15, 2006. They were grouped into five clusters and the proposals in each cluster were reviewed by one international expert in addition to the tutorials chair. The clusters were as follows, with the number of proposals in brackets and the expert reviewing the proposals in each cluster following:

Core NLP (2), Chris Manning  
Crossing borders (3), Robert Dale  
Machine learning (3), Walter Daelemans  
Corpora (3), Nancy Ide  
Dialogue (1), Diane Litman

Based on the expert opinions, the tutorials chair assisted by the general chair selected five proposals, taking the following criteria into account:

- Quality: Preference was given to proposals that were ranked highly by the experts, taking both the content of the proposal and the competence and experience of the presenters into account.
- Diversity: Preferably not more than one (or exceptionally two) proposal(s) should be selected from each cluster.
- Novelty: Tutorial topics featured at recent ACL events were dispreferred (unless the content was clearly novel and different).

This resulted in the following five proposals being accepted for presentation:

T1 (p.m.): Bayesian Nonparametric Structured Models  
Percy Liang, Dan Klein

T2 (a.m.): Usability and Performance Evaluation for Advanced Spoken Dialogue Systems  
Kristiina Jokinen, Michael McTear

T3 (p.m.): Textual Entailment  
Ido Dagan, Dan Roth, Fabio Massimo Zanzotto

T4 (a.m.): From Web Content Mining to Natural Language Processing  
Bing Liu

T5 (a.m.): Quality Control of Corpus Annotation Through Reliability Measures  
Ron Artstein

After the deadline of notification, January 15, 2007, it was discovered that one additional proposal had been sent before the deadline but had not reached its final destination because of a temporarily malfunctioning server that did not relay the message properly. After discussion with the general chair it was decided that the proposal should be reviewed as if it had been received in time, since the program could still accommodate one additional tutorial. This review, which involved two of the international experts, resulted in the proposal being declined. However, the incident shows that e-mail without explicit acknowledgment is not a safe method for submission of tutorial proposals. In the future, it is therefore recommended that tutorial proposals are submitted through the same conference system as all other submissions.

=====  
Report of ACL 2007 Publications Chair

Su Jian

As the publications chair for this conference, I have used the publication tool ACL PUB built by ACL2005 publications chairs Jason Eisner and Philipp Koehn.

1. Main conference publications

Same CoLing/ACL 06, participants do not get hardcopy proceedings, unless they explicitly request during registration and pay extra for them. On the other hand, we do prepare the printed companion volume, as Priscilla find it hard to sell the proceedings to libraries and others post-conference and then have to explain why they are not getting the full 2-volume set. Also it does form part of the written record of the conference.

A tool has been given to start to integrate to take away the author ID information, which could be ready for the next meeting.

2. Workshop publications

All 16 workshops have printed proceedings. Besides ACL PUB, a dummy data set was also created and distributed to the other book chairs, basically the organizers of WS, P&D, SRW, collocated conference in advance, so that they could have a dryrun to avoid the last minute panic. They've been given further messages on how to assemble the proceeding, feedback on their proceeding issues after the proofreading.

All proceedings covers were designed by the local organizing committee.

3. CDROM

All papers in the main conference and workshops are available on CDROM, which are distributed to all participants. This time it appears only EMNLP-CoNLL encounter problem on converting pdf to ps, which could be solved finally. If no more issue appears on the conversion to ps files, we might still keep to ps files to make it easier to the unix users.

All papers are also expected to be available at the ACL Anthology by the time if not earlier of the conference.

4. Technical issues

Deviant paper formatting (e.g. paper size, font embedding, etc.) has always been a problem, and it is no exception this time. 99 out of 131 papers have been communicated with the authors to make the necessary amendments. Among which about 26 papers used A4 instead of Letter sized required, although it has been emphasised again and again that Letter size is required.

The similar thing happens to almost all the other proceedings as well.

5. Suggestions and discussions

(1) It'll be good to make the report of the previous years' publication chair available to the current one. So that he or she could benefit from the experience. This might be extended to the other book chairs as well.

(2) It needs to remind the other book chairs in advance that they're responsible to the quality of their proceedings. Some of them tends to overlook the paper formatting issues totally.

Su Jian  
June 2007

=====  
Report for ACL 2007 Student Research Workshop

Student Chairs:  
Violeta Seretan (violeta.seretan@lettres.unige.ch),  
Chris Biemann (biem@informatik.uni-leipzig.de)  
Faculty Advisor:  
Ellen Riloff (riloff@cs.utah.edu)

1. Program Committee

The co-chairs of the ACL 2007 Student Research Workshop are Violeta Seretan (University of Geneva, Switzerland) and Chris Biemann (University of Leipzig, Germany). Ellen Riloff (University of Utah, USA) is the Faculty Advisor. The program committee was formed by the co-chairs by asking previous SRW reviewers, previous SRW participants and other researchers from the community. The final program committee consists of 52 reviewers, of which 26 were students or young researchers and 26 were senior researchers.

2. Paper Submission and Acceptance

We received 52 submissions from 22 countries (see Table 1). All papers were assigned 3 reviewers (at least 1 senior and 1 student reviewer). We accepted 16 papers, of which 9 are regular (oral) presentations and 7 are posters.

Country/Region	submissions	accepted
Brazil	1	-

China	6	-
Czech Republic	1	-
France	6	-
Germany	3	-
Hungary	1	1
Italy	1	-
India	2	-
Iran	1	-
Ireland	3	2
Nepal	1	-
New Zealand	1	1
The Netherlands	2	2
Palestine	1	-
Poland	3	2
Russia	1	-
Spain	2	-
Sweden	3	1
Switzerland	2	1
UK	6	4
Ukraine	1	-
USA	4	2
-----		
TOTAL	52	16

Table 1: Submission and acceptance by countries

### 3. Presentation Format

The Student Research Workshop posters are included together with the main conference poster session on Day 1 of ACL. We made sure that the student posters are presented next to the main conference posters, and can be kept hanging the whole conference.

The regular (oral) presentations are held as a parallel track on site on Day 2. Each oral presentation consists of 15 minutes of talk, and 5 minutes each for panelist feedback and for general audience questions.

### 4. Panelists

The co-chairs asked senior conference attendees to be on the panel to provide feedback to student authors. All papers received one or two panelists. These 19 panelists were selected for their knowledge in the area and availability during the workshop.

### 5. Funding

We submitted our request to the National Science Foundation in January 2007 and received the award in April 2007. The grant totaled \$22,200, of which \$18,000 was budgeted for student travel and \$4,200 was budgeted for administrative costs associated with running the workshop (e.g., facility rentals, workshop proceedings, student lunch). We will be able to provide funds to every SRW participant. Because the cost of traveling to Prague varies depending on the student's location, the level of funding is determined based on the cost of travel. We will award a guaranteed \$800 in funds to students from Europe, \$1,800 to students from North America and \$2,000 for students from Australia/New Zealand. The rest of \$2,000 will be used to cover the student co-chairs registration and will serve as a buffer for extra expenses. The remaining funds will be allocated to those students that did not manage to cover all costs with their allotted funds.

### 6. Organization and Planning

The Workshop was publicized by sending CFPs to mailing lists of computational linguistics and related fields, as well as direct emails to professors at various departments. The availability of funding

appears to be an important incentive for submissions, and we found it was important to include some funding information on the CFPs. The Workshop webpage was placed prominently on the main conference website. In addition, the ACL Newsletters helped to disseminate information on the Student Research Workshop. We are grateful to the main conference organizers for the support. The entire submission and review process was managed by the START system. This system proved immensely helpful for managing the 52 submissions and 52 reviewers. Before sending out acceptance and reject letters, we double-checked with the main conference organizers to avoid double acceptance. We rejected two papers that have been published at other events without considerable changes.

### 7. Suggestions and Considerations

a) We believe that the success of the Student Research Workshop depends on the quality of the reviewer and panelist feedback to students. We were happy to find 52 reviewers and 19 panelists who are supportive of this educational goal. The community was very responsive, which shows that the Student Research Workshop is widely accepted and recognized. Also for students whose work could not be accepted, the elaborate reviews will be of great help. We recommend that future Workshop organizers continue the tradition of concentrating their efforts on assembling good reviewers and panelists.

b) For many Workshop presenters, this is their first major conference attendance. Therefore, we thought it would be beneficial for students if we could arrange their poster/regular sessions early during the conference, such that they can begin networking and get the most out of the duration of the conference. We suggest that future Workshop planners communicate with the main conference organizers in the early stages of planning to ensure that the logistics for this situation works out. Both oral and poster presentations were scheduled and located to make the SRW look like a part of the main conference rather than a separate event. We believe this is beneficial for the students, as they get more attention from the general audience.

c) This year, the submission deadline as well as the notification of acceptance for the SRW was set to be at the same time as that of the main conference. We did receive papers that indicated double submissions, but not to ACL main session and the SRW. Although we clearly stated that double submission has to be labelled as such, some students did not consider previously published papers as double submissions and found it hard to understand why their papers were rejected once we found out by manually checking the web. In the future, this point should be made even more explicit in the Call for Papers.

d) We set the camera-ready deadline one week before the deadline for the main conference. As nearly all papers needed adjustments, which took until two weeks past the deadline to finish, this proved to be a necessary means that is highly recommended to the next organizers.

e) In the Call for Papers, we aimed at early stage Ph.D. work and suggested that advanced students should submit to the main conference. We believe this is a necessary contrast to NAACL's Doctoral Consortium, and suggest to do so in subsequent years. However, about half of the submissions were at too early a stage and could not even fill 4 out of a maximum number of 6 pages, which imposed extra work on reviewers, as most of these papers were clear rejects both formally and by content. We suggest to provide a minimum page number in the Call for Papers to avoid this.

f) Since our funding came from the National Science Foundation, in order for students to be reimbursed for their airfare their flights had to comply with the Fly America Act. This essentially means that the flights have to be on a U.S. carrier or code-shared by a

U.S. carrier, unless no such options are available (there are a few exceptions, primarily when the travel time would be dramatically increased). This caused some confusion among students, especially those traveling entirely within Europe, because they weren't sure if there were any "acceptable" flights or because the acceptable flights were a lot more expensive than an "unacceptable" flight. Our solution was to run their flight plans by a travel agent at the University of Utah who determined whether their flights were acceptable or not. However, this solution introduced lag time before the students could book their flights and it was a bit cumbersome for the travel agent.

Report from ACL/EACL-07 Workshop chair

Simone Teufel

The committee consisted of myself, Katja Markert, Beth Ann Hockey and Dekai Wu. We received 27 applications for workshops, which was more than expected. The committee chose 15 Workshops of different lengths, cf. below. The choice was very hard to do, as the quality of proposals and committees was very high. One factor this year, in addition to the usual ones, was the wish to be broad, i.e., to give research which is CL-related but not core-CL a place to present close in time to the core conference (e.g. embodied NLP, cognitive aspects...). We also wanted to allow old favourites which always have a wide participation, and encourage empirical, task-based competitions.

Two of the workshops were the result of a merger (WS9 and WS10).

The 15 chosen workshops are:

2 day:

WS1 SemEval

1.5-day:

WS9 Joint Workshop on Entailment and Paraphrase and 3rd PASCAL Recognizing Textual Entailment (RTE-3) Challenge  
WS10 Workshop on Linguistic Annotation (the LAW)

1 day:

WS2 Statistical Machine Translation  
WS3 Comp. Approaches to Semitic Languages: Common Issues and Resources  
WS5 BioNLP'07  
WS7 Deep Linguistic Processing  
WS8 SIGMORPHON Computational Research in Morphology and Phonology,  
WS12 NLP for Balto-Slavonic languages, Special Focus on IE  
WS13 Grammar-based approaches to spoken language processing  
WS15 Cognitive Aspects of Computational Language Acquisition

.5 day:

WS4 Language Technology for Cultural Heritage Data  
WS6 A Broader Perspective on Multiword Expressions  
WS11 4th ACL-SIGSEM Workshop on Prepositions  
WS15 Embodied Language Processing

This year, workshops were allowed to have posters. This was due to the fact that three of the workshops (WS1, WS2 and WS9) were in the form of competitions, and it had been planned that all participating systems should be given some space for presentation. Due to high submission rates and quality of submissions in other WS, these also opted for poster presentations.

The organisation of posters for the workshops was the most time-consuming aspect of my work, because so many workshops decided they wanted to have posters, and the decision came very late (during or after the reviewing period). I suggest that this should be changed for future ACLs; if posters will be allowed in following ACLs, I would recommend that clear rules eg.

about finances and deadlines for ordering posters, should be established beforehand.

To my knowledge, all WS organisers kept the deadline for proceedings to go to the publications chairs.

Special thanks to the local organisers, who provided fast, professional, unburocractic support for the WS organisation.

Simone Teufel

May 2007

Exhibits Chair Report for 2007

Jaroslava Hlavacova, Exhibits Chair (publishers)  
hlava@ufal.mff.cuni.cz

I sent the Call for Exhibits to approximately 25 publishers, at least twice (middle of February, middle of March, beginning of April). The great majority did not respond. From the rest, only two publishers were interested and booked an exhibition booth - both Czech minor publishers LEDA and Karolinum. Some of publishers ordered insertion of flyers into conference bags. Oxford University Press and Cambridge University Press will send only books without anybody to take care about them.

At the last moment, two European projects (Euromatrix and Companions) ordered a booth for their presentation.

--

Pavel Pecina, Exhibits Chair (companies)  
pecina@ufal.mff.cuni.cz

Four of all twelve sponsors are interested in an exhibition booth: Google as a Gold Sponsor plus three Silver sponsors (Textkernel, Newstin, and Powerset)

who paid for the booth separately.

ACL 2007 Sponsorship Report

The Sponsorship Chairs

Gabor Proszeky: proszeky@morphologic.hu  
Jan Hajic: hajic@ufal.mff.cuni.cz  
Jun'ichi Tsujii: j.tsujii@manchester.ac.uk  
Martha Palmer: Martha.Palmer@colorado.edu

The following sponsors have been recruited by the sponsorship chairs.

The publicity document is attached below. The total amount of sponsorship, not counting NSF, is \$24,500.00

GOLD Sponsors \$5,000.00: Google

Microsoft Silver Sponsor, \$4,000

SILVER Sponsors \$2,500.00: TextKernel, Newstin, Powerset

BRONZE Sponsors \$1,000: BBN, IBM (for the Best Student Paper Award), MorphoLogic, Xerox-Grenoble, Language Weaver, CSKI, Kilgray, ELRA.  
The Student Research Workshop is being generously supported by:

\* National Science Foundation.

For a complete list of sponsors, please visit this webpage  
<http://ufal.mff.cuni.cz/acl2007/sponsorship/>

=====

POSTERS and DEMOS  
Sophia Ananiadou

This year we had 113 submissions out of which 61 were selected for presentation, resulting in a 54% acceptance rate.

The criteria for acceptance of posters were to describe original work in progress, and to present innovative methodologies used to solve problems in computational linguistics or NLP. 48 posters were accepted. For demonstrations the criterion for acceptance was the implementation of mature systems or prototypes in which computational linguistics or NLP technologies are used to solve practically important problems. 13 demonstrations were accepted.

=====

Local Demos Arrangements

Miroslav Spousta  
(member of ACL'07 LOC responsible for local demos arrangements)

Demo session at the ACL 2007 conference will take place at the Top Hotel, Prague on Tuesday, June 26. Currently, there are 13 demos accepted for the conference, divided into two groups (7+6) -- morning and afternoon.

Accepted demos list as well as instruction for demo presenters are on the ACL'07 web page (<http://ufal.mff.cuni.cz/acl2007/demos/>, <http://ufal.mff.cuni.cz/acl2007/instructions/demo/>). We will provide table, chairs, LCD 19'' and poster easel for each presenter.

I cooperate with Zdenek Zabokrtsky (LOC) and Sophia Ananiadou (Demos/Posters Chair) on the scheduling and arranging of the demo session.

=====

Local preparation of posters sessions  
for ACL 2007 and associated events

Zdenek Zabokrtsky

There are four groups of events which 'consume' poster easels: (1) ACL poster sessions, (2) ACL demo sessions, (3) poster session of Student Research Workshop, and (4) associated events, namely EMNLP and 12 workshops (the organizers of the remaining three workshops did not announce any intention to have poster sessions). Besides that, there are also poster easels reserved for organizers' needs.

As for the ACL poster and demo sessions, I'm cooperating with Sophia Ananiadou, who is the ACL head for poster and demo sessions. According to the information which I received from her, there are 48 poster contributions divided into five poster sessions and 13 demo contributions divided into two demo sessions. The schedule of poster and demo sessions is available at <http://ufal.mff.cuni.cz/acl2007/demos/>.

There will be 7 contributions in SRW poster sessions. The poster easels for SRW will be located close to the easels of the main ACL poster sessions.

I am also responsible for allocation of poster easels for EMNLP and all workshops poster sessions. The posters are scheduled according to the information which I received from Simone Teufel and Jason Eisner.

The detailed schedule of all 'poster consuming events' is available to ACL organizers at [http://ufal.mff.cuni.cz/~zabokrtsky/acl/poster\\_schedule.html](http://ufal.mff.cuni.cz/~zabokrtsky/acl/poster_schedule.html) (for internal purposes only).

Instructions for all poster presenters are already placed on the official ACL'07 web: <http://ufal.mff.cuni.cz/acl2007/instructions/poster/>.

Zdenek Zabokrtsky  
(member of ACL'07 LOC responsible for arrangement of poster sessions)

=====

Mentoring Service (Florence Reeder: [freeder@mitre.org](mailto:freeder@mitre.org))

- o 17 requests for mentoring - all returned successfully
- o 8 mentors - many new
- o 3 new mentor volunteers for next year
- o Native languages: Chinese, Japanese, Russian, Arabic, Czech and French
- o Mentors: Charles Callaway, John White, Bea Oshika, Ken Samuel, Deborah Dahl, Marilyn Kupetz, Chrys Chrystello and myself.

=====

ACL 2007 Publicity Chairs Report

Jiri Mirovsky, Pavel Schlesinger, Pavel Stranak

(as of May 22, 2007)

We have compiled local information and prepared some maps. These have been published on the conference website. Four newsletters have been published so far. We are in the process of preparing the participants' handbook.

=====

Computational Linguistics Journal Editorial Report  
Robert Dale

1 SUMMARY

All is stable and going well, but big changes may be in the offing.

2 HIGHLIGHTS

- 2.1 Last Words column now running in every issue.
- 2.2 Survey Articles category being redefined to clarify its aim.
- 2.3 Stephen Wan has stepped down after two and a half years of excellent service as editorial assistant, to be replaced by Mary Gardiner, who carried out the role (also excellently!) in 2003-2004.
- 2.4 Mary is currently undertaking a detailed exploration of online 2.review management systems so we can improve our processes.
- 2.5 Blackwell have approached us in an attempt to entice us away from MIT Press.
- 2.6 Talk of Open Access is all the rage (and being taken into account in 2.4 ).

Important: Note that my five year tenure as editor ends mid 2008. I'm happy to serve another five years, as both Julia Hirschberg and James Allen did, but this would be at the Executive's pleasure.

### 3 STATISTICS

Time to first decision for new submissions:

For 2001 papers: 110 days  
For 2002 papers: 127 days  
For 2003 papers: 129 days  
For 2004 papers: 131 days  
For 2005 papers: 146 days  
For 2006 papers: 125 days[\*]

[\*] The 2006 number excludes special issue papers (we do not have all the data), proposals for survey articles, and one paper on which we do not have a decision.

Excluding special issue submissions, we had: 69 submissions consisting of 52 New (+2 same year revision of a submission in 2006) and 15 resubmissions. Special issue submissions were another 31 New submissions.

Here's the traditional 'disposition by first decision' table (2006 excludes special issue submissions):

Decision	2006	2005	2004	2003	2002	2001	2000	1999
Submitted	52	62	53	65	65	57	64	47
Accept	2	13	11	16	23	18	15	13
Reject	21	29	22	20	20	12	11	9
Resubmit as squib	1	0	3	2	2	2	1	3
Revise and resubmit	27	37	23	25	18	22	27	12
Withdrawn	0	0	1	0	2	3	3	2
No decision	1	1	2	2	0	0	7	8

At the time of writing, for 2007 we have the following: 22 Submissions total, consisting of 12 New and 10 resubmissions.

Average time to date to decisions for New submissions: no new submissions have decisions yet.

2007 Decisions:  
Submitted 22  
Accept 5  
Reject 1  
Revise and resubmit 0  
No decision yet 16

(At the time of this report in 2006 we had 31 submissions for that year, and in 2004 we had 42.)

=====  
Computational Linguistics  
Book Review Editor's report  
Graeme Hirst  
May 2007

### BOOK REVIEWS PUBLISHED

In 2006 we published 16 book reviews in the journal, plus a few brief notices.

Most reviews are published in a timely manner -- that is, within 12 months of receipt of the book. This allows six months for the

reviewer (most take less) and five months for journal production.

### MATERIAL REVIEWED

I am continuing to be fairly strict in deciding if a book is to be reviewed, but try to include all books that are in "core" computational linguistics, as well as a variety of books from adjacent and overlapping disciplines that are likely to be useful in CL. We do not review technical reports, doctoral theses, conference proceedings, or workshop proceedings, except if revised for publication as a book by a recognized publisher. Sometimes it has proved to be impossible to find a reviewer for a book who has sufficient expertise in the subfield and yet does not have a conflict of interest due to an association with one of the authors, editors, or contributors of the book.

### PRODUCTION MATTERS

I am indebted to Nadia Talent for long hours of reading out loud with me to check the galleys.

=====  
Computational Linguistics  
Squibs and Discussions: report for 2006  
Pierre Isabelle

At the beginning of year 2006, there was only one submission in the squibs pipeline. In the course of year 2006, 19 additional papers were (re-) submitted. This was an all-time record for squibs. At the end of 2006, there were only two submissions left in the pipeline. Thus, 18 decisions were made over the year. The results were as follows:

- \* 2 papers accepted
- \* 7 papers rejected
- \* 8 invitations to rework the (re-)submission
- \* 1 withdrawn

The mean time taken for these decisions was 107 days.

Thus far, only 3 papers have been (re-) submitted in 2007.

-- Pierre Isabelle  
Squibs editor for Computational Linguistics  
23 May 2007

=====  
ACL ANTHOLOGY Report, July 2007  
Steven Bird

The ACL Anthology is a digital archive of research papers in computational linguistics, sponsored by the CL community, and freely available to all. It includes the Computational Linguistics journal, and proceedings of many conferences and workshops including: ACL, EACL, NAACL, ANLP, TINLAP, COLING, HLT, MUC, and Tipster. Conference proceedings are published in the anthology around the same time as the conference. CL articles are published in the anthology one year in arrears (but individual subscribers can access recent issues electronically via the MIT Press website).

The anthology now contains over 12,500 papers (up from 11,000 papers twelve months ago), along with full-text search. Most of the papers are also indexed by CiteSeer and Google Scholar, helping the citation counts of ACL authors. e.g. the following Google Scholar search reported nearly 8,000 results:  
<http://scholar.google.com/scholar?q=site%3Aacl.ldc.upenn.edu>. The ACM

Digital Library is creating rich metadata and doing full citation linking for all anthology materials.

ADDITIONS OVER LAST 12 MONTHS: Proceedings from HLT-NAACL-07, ACL-06, EACL-06, HLT-NAACL-06, ACL-04, COLING-04, HLT-NAACL-04; CL Vol 31 (2005).

CONVERSION OF LEGACY MATERIALS: The 2004 conferences all used idiosyncratic directory layouts, filenames and HTML formats, and were converted manually by a student assistant at Melbourne University.

MAILING LIST: A new mailing list has been set up for announcements concerning new materials added to the Anthology: <http://groups.google.com/group/acl-anthology>

FUTURE MATERIALS: The ACL publication software generates conference CD-ROMs using the same directory layout and file-naming conventions as the anthology, streamlining the online publication process. BibTeX files are automatically generated and made available to users. The journal and any SIG workshops not held in conjunction with an ACL meeting will continue to require manual processing.

DIGITAL OBJECT IDENTIFIERS: DOIs are akin to ISBN numbers, but apply to individual papers. They are now the standard way to uniquely identify an academic paper, and web services will be available for resolving DOIs to papers (e.g. <http://dx.doi.org/>).

MIT Press assigns DOIs to CL articles, and the ACM is in the process of assigning DOIs to each anthology item. Before the ACM could do this we had to join CrossRef and get a "DOI Prefix" (our prefix is 10.3115). The nominal cost for DOI assignment is \$1 per article; the ACM will cover the cost for past materials, while the ACL will cover the cost of DOI assignment for anthology materials from 2006 onwards.

IJCNLP: From 2008 onwards, IJCNLP proceedings will appear in the Anthology.

#### ONGOING ACTIVITIES

HIGHER-QUALITY BIBLIOGRAPHIC METADATA: The ACM Digital Library is creating high-quality bibliographic metadata for each individual paper, in conjunction with registering each paper with a DOI. It should be possible to extract that metadata and improve the quality of metadata on the Anthology site (e.g. removing OCR errors in the spelling of author and paper names).

PUBLICATION INSTRUCTIONS: The instructions for the publication software need to be updated to cover two further tasks: (i) obtaining the workshop identifiers from the Anthology editor, and (ii) uploading the materials to the anthology by FTP. Conferences and workshops not held in conjunction with a regular ACL meeting are not automatically included in the Anthology. Organizers of such events should consider using the ACL publication software and contacting the Anthology editor to ensure timely incorporation of the proceedings in the Anthology.

TIMING: Conference and workshop organizers have a variety of opinions about exactly when proceedings should appear in the Anthology (e.g. before, during, or after the event). I recommend that the ACL Executive establish a standard practice here. Journal papers appear 1-2 years after they are published, as decided by the Executive when the Anthology was founded. However, the ACM Digital Library publishes them much sooner (within a few months of publication?), for free access. I presume this practice has not hurt subscriptions, in which case I propose that the Anthology do the same.

IJCNLP: The 2005 conference proceedings were not included in the anthology because they were published by Springer. Springer agreed to make the proceedings freely available online after one year, but this

has not yet happened. Would the IJCNLP-05 organizers be able to supply the materials?

ACM DL: Our ACM Digital Library contact, Bernard Rous, has asked to receive CD-ROMs of ACL conferences as they are published, so that he can initiate the process of assigning DOIs. His address is:

Bernard Rous, Electronic Publishing Program Director,  
ACM, 2 Penn Plaza Suite 701, New York NY 10121-0701

TEXT EXTRACTION: There is an initiative to extract plain text from the ACL Anthology materials, involving Dragomir Radev, Min-Yen Kan and others. Most of the Anthology has been converted, and can be found at <http://wing.comp.nus.edu.sg/~min/dAnth/acl/>. This will facilitate the application of NLP techniques to our own publications.

TOPICAL INDEXING: The existence of persistent URLs makes it easy for individuals and special interest groups to set up annotated bibliographies with pointers to papers in the anthology. Moreover, the community's own text categorization techniques ought to be applied to its own text collection. The anthology site should link to any well-curated, comprehensive categorizations of its content, so that members of the CL community can benefit from them. The new ACL Wiki would be a convenient place for members to maintain topical indexes of ACL papers.

LONG-TERM MAINTENANCE: The Anthology "project" is almost concluded: (a) materials from the ACL's hardcopy and microfiche eras are now all digitized; (b) born-digital materials published in ad hoc formats have been manually converted; and (c) ACL's publications software supports publication via the Anthology. Final steps are to (d) update index pages with high-quality bibliographic metadata provided by the ACM; (e) host the primary copy of proceedings on the ACL website; and (f) assign well-documented tasks to a website manager and to the publications chair of ACL conferences, under the oversight of the ACL Secretary or designee. Once these steps have been carried out, the Anthology will be fully incorporated into the ACL's operation.

=====

#### NLSR Report - Thierry Declerck

In the reporting period, the ACL NLSR continued its main activity, which consists in updating the listing of Natural Language Processing Tools on the base of voluntary submissions made by the authors of the software (see [registry.dfki.de](http://registry.dfki.de)). One nasty episode: the submission form has been misused by senders of spam, and we had to update the submission form for stopping this nuisance. We plan in the future to include the services offered by the ACL NLSR to actual initiatives dealing with grid technologies (see for example the Language Grid project in Japan: <http://langrid.nict.go.jp/>), or with infrastructures for Language Resources (see the European infrastructure initiative CLARIN (Common Language Resources and Technology Infrastructure, <http://www.mpi.nl/clarin/>)).

=====

#### ACL Archives

Eric Fosler-Lussier

The ACL archive was reorganized in 2006-2007; the new archive can be found at <http://www.aclweb.org/archive>. Highlights of the reorganization include:

- Document labeling by both year and category;
- Introduction of new material including;

- Missing executive committee reports;
- Archived versions of the policies and procedures documents;
- Updated officers and editorial board listings;
- Expanded conference bids section;
- New archival policy: all documents stored in the archive must not point to links outside aclweb.org (in order to prevent loss of information).

As always, suggestions for items to include in the archive, corrections, or alternative categorizations are welcome at (the archivist)[one word] -at-aclweb.org.

Eric Fosler-Lussier, ACL Archivist

=====

ACL Wiki for Computational Linguistics - Report for 2007

Peter Turney  
peter.turney@nrc-cnrc.gc.ca

The ACL has created a new wiki specifically for Computational Linguistics, at <http://aclweb.org/aclwiki>. The purpose of this wiki is to facilitate the sharing of information on all aspects of Computational Linguistics. The ACL Wiki was created to fill a role that Wikipedia cannot fill. The wiki includes links to Computational Linguistics blogs, conferences, competitions, people, organizations, course descriptions, corpora, datasets, and introductory articles on topics such as Computational Lexicology, Computational Semantics, and Word Sense Disambiguation.

The ACL Wiki was initiated by Steven Bird, Ali Hakim, Dragomir Radev, and Peter Turney. On October 18th, 2006, MediaWiki software was installed on the ACL web server and the ACL Wiki began running. The new wiki was announced on the ACL mailing list on October 30th, 2006. This report was written when the wiki had been open to the public for about half a year. At that time, there were 1,755 pages in the wiki database. This figure includes minimal pages that have relatively low content, but there were at least 100 high-content pages. In this half-year period, there have been more than 100,000 page views and more than 2,000 page edits, an average of more than 500 views and 10 edits per day. Only registered users can edit pages and there are about 260 registered users. Registration requires only a valid email address. Although it has only been running for half a year, it is clear that the ACL Wiki has already attracted the support and involvement of the Computational Linguistics community.

=====

COLING/ACL 2006

Nicoletta Calzolari : COLING/ACL 2006 General Chair

COLING/ACL 2006 was the third joint conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics. The conference was held in Sydney, Australia, from 17-21 July 2006, with tutorials on July 16, workshops on July 22-23, and co-located events on July 15-16 and July 22-23.

In this joint conference we tried to maintain the spirit of both COLING and ACL, but the combination had its own personality, in a mixture that was more than the simple sum of the two. Part of its character was due to the location, for the first time -- for both conferences -- in Australia. For this reason we decided to have a member of AFNLP (the

Asian Federation of Natural Language Processing) on the Advisory Board and to give particular attention and visibility to the Asia-Pacific context, communities and languages. We sincerely thank both the AFNLP-Nagao Fund for providing financial support for those presenting Asian NLP research, and ALTA (the Australasian Language Technology Association) for their local support.

I profit of this report to express my gratitude again to all those without whom this conference would not have existed, and I think I can do that on behalf of all participants.

My biggest thanks go to all the Chairs, for their invaluable effort and dedication which made this Conference possible. First of all the two Program Chairs: Claire Cardie and Pierre Isabelle, who did a tremendous job, managing so many submissions and taking care of both regular papers and posters, and the two Local Arrangements Chairs: Robert Dale and Cecile Paris, who succeeded in keeping so many details under control, in such a smooth way as if everything were natural and effortless for them. And all the others, for their precious, competent and hard work.

I warmly thank the Advisory Board -- composed of four ICCL, four ACL, and one AFNLP members -- to whom we resorted for suggestions on important and sometimes delicate issues: Sandra Carberry, Eva Hajicova, Aravind Joshi, Martin Kay, Kathleen McCoy, Martha Palmer, Priscilla Rasmussen, Benjamin T'sou, Jun'ichi Tsujii.

I express my gratitude to all the sponsors for their great support to the conference.

I thank all the organizers of the so numerous surrounding workshops, tutorials, and other co-located events M-^V conferences, workshops, summer school -- adding value to the main conference, creating altogether probably the biggest ever happening in Computational Linguistics.

My thanks to the area chairs, the reviewers, the invited speakers, the authors of the various presentations, in particular the students who enter with enthusiasm in such an exciting field, all the participants who in many cases made a long trip to be present at COLING/ACL 2006, and all those who contributed in many ways to a success of the conference.

And I finally thank both ICCL and ACL for having decided to join forces again in such a great enterprise. COLING/ACL 2006 was, I'm sure, an exciting, stimulating and inspiring event for all those who attended.

=====

Report on the NAACL Conference Human Language Technologies, 2007  
Submitted by Candy Sidner, General Chair for the conference

NAACL's annual conference for 2007, Human Language Technologies 2007 was held in Rochester, N.Y. from April 22 through April 27, 2007 at the Hyatt Regency Hotel. Just under 500 people attended the conference.

There were 298 papers submitted to the meeting and 72 accepted, with 150 late breaking (short) papers submitted and 55 accepted. In addition, 16 demos were also given during the main conference. The 21 member organizing committee was chosen with representatives of all

three sub-areas for NLP, IR and Speech. The program committee chairs were Tanja Shultz, ChengXiang Zhai, Matthew Stone. The program chairs were supported by a senior program committee of 30 people and a program committee of 312 people to review the long and short papers. The full organizing committee, senior program committee and program committee can be found in the conference proceedings, which have been online at <http://acl.ldc.upenn.edu/N/N07/N07-1000.pdf> since shortly before the conference.

On April 22, the Doctoral consortium took place with 10 student papers and panels accompanying the presentations of each pair of papers. Also on April 22, four tutorials were held, and on April 26/27, there were 5 workshops.

The conference had several sponsors: the Eastman Kodak company, Microsoft Research, Powerset, Thomson, the Association For Machine Translation in the Americas, IBM, and Language Weaver.

Our organizing committee had some significant difficulties with the publication software. A full report will be provided after Memorial Day. The pc chairs also thought that having three chairs was somewhat unwieldy. One option would be to have a pc general chair and 2 associate chairs (representing whichever 2 areas were not the area of the pc general chair).

=====

Report on the state of the preparations of ACL08  
Prepared by: Chris Brew

General Conference Chair: Kathy McKeown  
Local Organizer: Chris Brew  
Other local arrangements people:: Detmar Meurers, Michael White, Donna Byron, Eric Fosler-Lussier

1. Dates: June 15-20, 2008

Tutorials: Sunday June 15, 2008  
Main conference: Monday-Wednesday, June 16-18, 2008  
Workshops: Friday Saturday June 19-20, 2008

2. Location: Hyatt Regency Hotel, Short North, Columbus OH

This is a conference hotel attached to the greater columbus convention center, somewhat airport like, but within comfortable walking distance of downtown and of small restaurant district near downtown.

3. Rooms reservation:  
Tutorials: 3 rooms (capacity flexible)  
Main Conference: 3 rooms (capacity 400, 300, 200) go together for plenary  
Workshops: 6 rooms (capacity flexible, with partitions)

large corridors suitable for poster use.

A bit bare, should consider asking for furniture to make more comfortable.

ACL-Exec dinner not determined, can work on this.

4. Accommodation:  
(a) single/double/triple/quadruple in conference hotel for USD 140/150/160/170 respectively. Block booking made. Kathy McKoy, Priscilla have details of numbers

and have signed off on commitment  
(b) 120 x USD 80 rooms at Red Roof in just across from conference hotel

Thinking is that main resource for cheaper rooms is triples/quads in conference

hotel. Student dorms considered but ugly in US/CA shared dorm mode and in any case far from conference hotel, requiring bus-ride.

5. Social programme:

welcome reception: Sunday, 15 June, in the Statehouse atrium. Beautiful space. Provisional booking, very unlikely to be needed by Governor on a Sunday night, but he would have precedence. We would know in good time to find a backup.

banquet plans: Tuesday 17 June, pavillion of Columbus Zoo, ample capacity, at least 700, would need transport. price: approx. 80. USD. Downtown options much less attractive

6. Conference fee: comparable to NAACL 07, which was USD 425 regular

=====

ACL/IJCNLP-2009 Report, June 2007  
Bonnie Dorr

The 47th Meeting of the Association for Computational Linguistics and 4th Bi-Annual Meeting of the IJCNLP, will take place in Asia in 2009. As the future past president in 2009, I am responsible for chairing the coordination group for ACL/IJCNLP-2009, which consists of members from both ACL and AFNLP:

Bonnie Dorr (chair of ACL/IJCNLP-2009 coordinating committee)  
Yuji Matsumoto (co-chair of ACL/IJCNLP-2009 coordinating committee)  
Kathy McCoy  
Martha Palmer  
Dragomir Radev  
Priscilla Rasmussen  
Keh-Yih Suh  
Benjamin Tsou  
Jun-ichi Tsujii

This group has posted the call for bids, at this URL:

<http://www.umiacs.umd.edu/~bonnie/acl-ijcnlp-09-cfb.html>

We received three very solid bids in February of this year. The group has reached a consensus for a recommendation to the two Executive Boards (ACL and AFNLP) and we expect to be in a position to announce the site of ACL/IJCNLP-2009 at the 2007 ACL this year.

=====

ACL-2010 Report, June 2007  
Steven Bird

The 48th Meeting of the Association for Computational Linguistics, will take place in Europe in 2010. According to our policy on joint conferences [1], I am responsible for chairing the coordination group for ACL-2010. This group will oversee the planning and organisation of the conference, and carry out the tasks described in [1].

The initial membership of this group is as follows, and will be updated as officers change.

- \* Giorgio Satta (as future EACL president)
- \* Anette Frank (as EACL secretary)
- \* Mike Rosner (as EACL treasurer)
- \* Alex Lascarides (as additional EACL board member; current president)
- \* Steven Bird (as ACL past-president in 2010)
- \* Drago Radev (as ACL secretary)
- \* Kathy McCoy (as ACL treasurer)

- \* TBA (additional ACL board member; to be newest member of the ACL executive in 2010)
- \* Priscilla Rasmussen (as business officer of the ACL)

Although its early to be thinking about ACL 2010, I would like to circulate the call for bids at the Prague meeting. This will permit potential hosts to start discussing it with colleagues and with the organisers of this year's meeting. The coordination group will be compiling the call for bids in June.

I am also in the process of securing [www.acl2010.org](http://www.acl2010.org) for this conference.

[1] <http://www.aclweb.org/policies/joint-conference.html>

=====

SIGDAT - 2007 Summer Report

David Yarowsky

SIGDAT is ACL's special interest group for linguistic data and corpus-based approaches to NLP.

In 2007, SIGDAT is organizing a 3-day Joint Conference on Empirical Methods in Natural Language Processing and Conference on Computational Natural Language Learning (EMNLP-CoNLL 2007) with our sister sig SIGNLL. The meeting is scheduled immediately after ACL-07 in Prague on June 28-30. Jason Eisner is program chair, Jan Hajic is local arrangements chair, and Eric Ringger is publications chair.

The conference appears to be highly successful: Over 400 submissions were received, and 66 papers were accepted for oral presentation and 44 were accepted as posters, yielding an total acceptance rate of approximately 25% (and 15% acceptance for full oral presentation). The proceedings exceeds 800 pages, and essentially the entire conference will be held in parallel sessions. In terms of scale on several dimensions, EMNLP is now regularly at a similar size to recent NAACL/HLT and EAACL meetings.

In 2006, SIGDAT organized a 2-day Conference on Empirical Methods in Natural Language Processing (EMNLP-2006), held immediately after ACL-06 in Sydney on July 22-23. Dan Jurafsky and Eric Gaussier were program chairs. As with this year, over 400 submissions were received, and 43 full papers accepted and 30 posters accepted, yielding an acceptance rate under 20%. The proceedings exceeded 600 pages.

As one of ACL's first SIGS, SIGDAT was formed prior to the requirement that SIGs have a constitution. SIGDAT is taking steps to create a constitution and further normalize our structure before the end of the year, consistent with ACL policy.

We will also actively pursue the question of the role of EMNLP, in particular in respect to its scheduling in conjunction with other ACL events, as it continues to grow.

- David Yarowsky  
Secretary-Treasurer

=====

REPORT ON SIGDIAL ACTIVITIES: July 2006 to June 2007

David Traum, SIGdial President

SIGdial is the ACL and ISCA Special Interest Group on Discourse and Dialogue which was formed in November 1997. More information about SIGdial can be found on the webpages: <http://www.sigdial.org> including

an actively updated calendar of upcoming events, resources, and previous reports. Members can join from the webpage, which includes participation in a low-volume, moderated mailing list (mainly conference and job announcements). SIGdial currently has 447 members from 34 countries (up 52 members and up 1 country from last year).

SIGdial has held an annual workshop on discourse and dialogue since 2000. The last workshop was held in July 15-16, 2006, in Sydney, Australia. The chairs were Jan Alexandersson and Alistair Knott. The next workshop will be in September 2-3, 2007, at Antwerp, Belgium, immediately following Interspeech 2007. The workshop is also recognized by ISCA as a satellite event for Interspeech 2007. The chairs are Harry Bunt and Tim Paek. More information on SIGdial workshops can be found here: <http://www.sigdial.org/workshops/>.

SIGdial also endorses a number of other dialogue-related workshops and events that are open to the general community. The SIGdial Endorsed events for the previous and upcoming year are:

- July 7-9, 2006: Workshop on Constraints in Discourse (Maynooth, Ireland)
- August 7-11th, 2006 ESSLII 2006 Workshop on "Coherence in Generation and Dialogue" (Malaga, Spain)
- September 11th-13th 2006: Brandial 2006: The 10th Workshop on the Semantics and Pragmatics of Dialogue (Potsdam, Germany)
- September 17th, Interspeech 2006 - ICSLP Satellite Workshop Dialogue on Dialogues: Multidisciplinary Evaluation of Advanced Speech-based Interactive Systems (Pittsburgh, PA)
- September 1, 2007: Young Researchers' Roundtable on Spoken Dialog Systems (Antwerp, Belgium)

=====

ACL SIGGEN report 2006-2007

Charles Callaway, Roger Evans, David McDonald, Jette Viethen, Michael White  
SIGGEN board

SIGGEN is in good health. SIGGEN members have organised a range of meetings, with more in the pipeline - the SIGGEN board would particularly like to thank all the organisers of these events for their hard work on behalf of the community.

Events this year:

- INLG'06 (Sydney, August 2006 - part of ACL'06)
- Workshop on Multimodal Output Generation (Aberdeen, January 07)
- Workshop on Shared Tasks and Comparative Evaluation in NLG (Arlington, April 07)
- ENLG'07 (Schloss Dagstuhl, June 2007 - prior to ACL'07)

Future events:

- Workshop on Using Corpora for NLG: Language Generation and Machine Translation (Copenhagen, September 07 - part of MT Summit)
- INLG'08 (Salt Fork, Ohio, June 2008)

There has been a lot of discussion and engagement in the SIGGEN community over the last year, in particular associated with proposals to set up 'shared task' evaluations which have become a feature in many other areas of NLP. This has resulted in the establishment of the first Shared Task and Evaluation Campaign for NLG, organised by Anja Belz (Brighton), Ehud Reiter and Albert Gatt (Aberdeen) and Jette

Viethen (Macquarie), taking place from May-September 2007.

Electronic communication has been quite modest. The SIGGEN website has been well maintained, but the mailing list has seen fairly low volumes of traffic, and the wiki has also not quite taken off yet, although it does have quite a lot of useful content. The SIGGEN board aim to improve the SIG's electronic profile over coming months, reviewing the role and delivery of the wiki, and exploring the establishment of a repository for NLG papers not submittable to the main ACL repository.

Elections for three committee posts were held in 2006, resulting in the appointment of Michael White, Roger Evans and Jette Viethen (student representative). The board thanks the outgoing members (Tilman Becker, Irene Langkilde-Geary and David Reiter) for their work on behalf of SIGGEN. The board has responded to the ACL request to improve compliance with SIG guidelines, and are now fully compliant apart from board and election structure - this requires consultation with the membership and possible constitutional changes and will be pursued over coming months. A call for proposals to host INLG'08 was issued, resulting in a successful bid from Michael White and Crystal Nakatsu (Ohio State) and David McDonald (BBN) to host in Ohio. Budgets and a contract with the venue have been negotiated. The SIGGEN shadow account shows a balance of \$928, but figures for INLG'06 have not yet been taken into account.

Charles Callaway, Roger Evans, David McDonald, Jette Viethen, Michael White  
SIGGEN board

=====  
ACL\_SIGLEX\_2007.txt

Rada Mihalcea

SIGLEX is the Special Interest Group on the Lexicon, providing an umbrella for research interests on lexical issues ranging from lexicography and the use of online dictionaries to computational lexical semantics. SIGLEX is also the umbrella organization for Senseval/Semeval - the evaluation exercise for systems for the semantic analysis of text.

The following ACL members are currently serving as SIGLEX officers:

President: Rada Mihalcea, University of North Texas  
Secretary: Ted Pedersen, University of Minnesota, Duluth

Executive Board:

Francesca Bertagna, Istituto di Linguistica Computazionale (ILC-CNR)  
Stefan Evert, Universitat Stuttgart  
Katrin Erk, Saarland University  
Jimmy Lin, University of Maryland  
Diana McCarthy, University of Sussex  
Claudia Soria, Istituto di Linguistica Computazionale (ILC-CNR)

Board Member with Special Portfolio/Webmaster:  
Ken Litkowski, CL Research

One of the main events for 2007 was the Semeval (former Senseval) exercise for the evaluation of systems for the semantic analysis of text. The evaluation took place in Spring 2007, and the workshop is scheduled to take place on June 23-24 2007, in conjunction with the ACL conference in Prague. The co-chairs are Richard Wicentowski (Swarthmore College), Lluís Marquez (Universitat Politècnica de Catalunya), and Eneko Agirre (University of Basque Country). This has been a highly successful event, with 18 different tasks (selected out of the 25 submitted), more than 100 participating teams, and more than

125 systems. This represents a significant increase compared to the 14 tasks and 55 teams participating in Senseval-3. More information about Semeval is available from the event webpage <http://nlp.cs.swarthmore.edu/semeval/>.

During 2006-2007, SIGLEX has endorsed the following events:

- \* SemEval-2007, 4th International Workshop on Semantic Evaluations, Evaluation period: February 26, 2007 to April 1, 2007, SemEval-2007 Workshop, co-located with ACL-2007, June 23-24, 2007, Prague
- \* TextGraphs-2: Graph-based Methods for Natural Language Processing, HLT/NAACL 2007 Workshop, April 26, 2007, Rochester, New York
- \* A Broader Perspective on Multiword Expressions, ACL 2007 Workshop, June 28, 2007, Prague, Czech Republic (Endorsed by the ACL Special Interest Group on the Lexicon, Submission due March 26, 2007)
- \* 4th ACL-SIGSEM Workshop on Prepositions, ACL 2007 Workshop, June 28, 2007, Prague, Czech Republic (Sponsored by the ACL Special Interest Group on Semantics, Submission due March 26, 2007)
- \* Acquisition and Management of Multilingual Lexicons, Workshop at RANLP-2007, September 30, 2007 (Endorsed by the ACL Special Interest Group on the Lexicon, Submission due June 15, 2007)

=====  
ANNUAL REPORT  
SPECIAL INTEREST GROUP ON MULTIMEDIA LANGUAGE PROCESSING  
(SIGMEDIA)

May 23rd 2007

CHAIR: Elisabeth André (University of Augsburg, Germany, [andre@informatik.uni-augsburg.de](mailto:andre@informatik.uni-augsburg.de))

MAILING ADDRESS: [andre@informatik.uni-augsburg.de](mailto:andre@informatik.uni-augsburg.de)

URL: <http://www.sigmedia.org/>

CONFERENCES AND WORKSHOPS:

Following the tradition of the successful ISCA Tutorial and Research Workshop on Multi-Modal Dialogue in Mobile Environments (IDS02) in 2002, the Tutorial and Research workshop on Affective Dialogue Systems (ADS04) in 2004, SIGMEDIA organized another Tutorial and Research workshop on Perception and Interactive Technologies (PIT 2006) in collaboration with the ACL Special Interest Group SIGDial. The workshop took place at Kloster Irsee, Germany from June 19-21 2006. The organizing committee consisted of: Elisabeth André, Gregory Baratoff (Siemens VDO Automotive AG, Germany), University of Augsburg, Germany, Laila Dybkjaer, University of Southern Denmark, Markus Hennecke (Harman/Becker Automotive Systems, Germany), Wolfgang Minker, Heiko Neumann and Michael Weber all three from University of Ulm, Germany. The proceedings have been published in the Springer LNCS series.

Furthermore, SIGMEDIA has started with the preparation of another Tutorial and Research Workshop on PERCEPTION AND INTERACTIVE TECHNOLOGIES to be held in Kloster Irsee from June 16 to June 18, 2008. The organizing committee consists of: Elisabeth André, University of Augsburg, Germany, Laila Dybkjaer, University of Southern Denmark, Wolfgang Minker, Heiko Neumann and Michael Weber, all three from University of Ulm, Germany.

Prof. Dr. Elisabeth André, [andre@informatik.uni-augsburg.de](mailto:andre@informatik.uni-augsburg.de)  
Lehrstuhl für Multimedia-Konzepte und Anwendungen, Institut für Informatik  
Universität Augsburg, Eichleitnerstr. 30, D-86135 Augsburg Germany

Phone: +49 821 598 2341, Fax: +49 821 598 2349  
<http://mm-werkstatt.informatik.uni-augsburg.de>

=====  
ACL\_SIGMOL\_2007.txt  
Report for SIGMOL, Mathematics of Language, June 2007  
Gerald Penn

Our next bi-annual meeting, MoL-10, will take place July 28-30, 2007, and will be hosted by the UCLA Department of Linguistics. Of 23 submissions, 13 were accepted for presentation with 1 withdrawal. There will also be invited talks by Kevin Knight, Larry Moss, Partha Niyogi, Geoffrey Pullum, Marcus Tomalin, and Dag Westerstahl. Proceedings will be prepared after the conclusion of the conference. The possibility of a special journal issue will be discussed at the conference with the attendees.

This year's meeting not only marks MoL's 20th anniversary, but also the 50th anniversary of the publication of Chomsky's *Syntactic Structures*, hence the larger number of invited speakers.

MoL endorsed last year's ACL-COLING workshop on Tree-Adjoining Grammars (TAG+8). We have also redesigned and relocated our web-page.

The next president of MoL, effective at the end of July will be Gerhard Jaeger of the University of Bielefeld. Our election was held last fall.

=====  
ACL SIGNLL - President's Report 2006-2007

Antal van den Bosch

The goals of SIGNLL, ACL's special interest group on natural language learning, are to promote and inform about research on computational modeling of learning in natural language. These are served by (i) the maintenance of an informative and up-to-date website and associated mailing list, and (ii) the organization of annual events (the CoNLL conference and the CoNLL shared task), and support of other related activities.

In 2006-2007 SIGNLL has grown to 484 registered members (as of May 20, 2007). The current elected board (until October 2007) is composed of Antal van den Bosch, president; Hwee Tou Ng, secretary; and Erik Tjong Kim Sang, information officer.

The web-pages, located at URL <http://www.aclweb.org/signll/> and maintained by Erik Tjong Kim Sang, remain an important source of information, complemented by an email list for announcements for SIGNLL-related events. On the web-site, links can be found to relevant associations, networks, research cooperations, research departments, groups, institutes, mailing lists, archives, journals, bulletins, conference reports, online papers (including all papers of all CoNLL proceedings), online courses and slides, bibliographies, software, corpora, companies, meta-information sources etc.

The main events in 2006-2007 were the tenth and the eleventh CoNLL conferences (SIGNLL Conference on Computational Natural Language Learning), the latter organized jointly with SIGDAT under the header of EMNLP-CoNLL 2007, and yet to happen on June 28-30, 2007 in Prague, co-located with ACL-2007.

CoNLL-X

The tenth anniversary of CoNLL, dubbed CoNLL-X for the occasion, was

celebrated in New York, 8-9 June, 2006, co-located with HLT-NAACL (<http://www.cnts.ua.ac.be/conll2006/>). Programme chairs were Lluís Marquez and Dan Klein. The two-day event featured invited talks by Michael Collins and Walter Daelemans. Of 53 full paper submissions, 18 were accepted for presentation (34%). 85 registered participants received a celebration t-shirt showing the "CoNLL World Tour 1997-2006" dates at the backside.

The CoNLL-X shared task focused for the first time on Dependency Parsing. It was organized by Sabine Buchholz, Yuval Krymolowski, Erwin Marsi, and Amit Dubey, and featured training, validation, and test sets for an impressive 13 world languages. The task was shared by 19 participants who together submitted 247 systems.

For the first time CoNLL featured a best paper award, selected by a jury appointed by the CoNLL-X programme chairs. The first award went to Rie Kubota Ando for her paper "Applying Alternating Structure Optimization to WSD".

EMNLP-CoNLL 2007

In the fall of 2006, the boards of SIGDAT and SIGNLL decided to propose to ACL a joint EMNLP-CoNLL conference, to be aligned with ACL-2007 in Prague, 28-30 June, 2007. The driving reason for the alignment was that ACL-2007 was the only summer conference in computational linguistics in 2007, and both SIGs intended to align with it. Before, the overlap between EMNLP and CoNLL has caused the two SIGs, by way of liaison David Yarowsky, to develop a careful planning that avoided the co-location of the two events, but now the argument was reversed and forces were joined. Under program chair Jason Eisner an exciting program of 110 presentations has been prepared (<http://www.cs.jhu.edu/EMNLP-CoNLL-2007/>) on the basis of over 400 submissions. 66 papers will be presented as talks, and 44 as posters.

The CoNLL shared task for 2007 is a continuation of the CoNLL-X shared task on Dependency Parsing. A shared task organisation team headed by Joakim Nivre and further composed of Johan Hall, Sandra Kuebler, Ryan McDonald, Jens Nilsson, Sebastian Riedel, and Deniz Yuret organized two subtasks: a Multilingual Track focusing on the development of dependency parsers for 10 languages, and a Domain Adaptation Track aimed at measuring the robustness and adaptability of English dependency parsers for parsing data beyond the genre of the training set. 23 groups participated in the competition. The shared task session at EMNLP-CoNLL will draw the usual attention and suspense.

Conclusion

The SIGNLL board is of the opinion that the SIG remains unique in its focus. We are happy to observe that the CoNLL conference series continues to have a significant impact on the field, partly because of the successful shared tasks, which have been broadly referenced and have contributed benchmark data sets that are commonly used throughout computational linguistics. We keep striving for complementarity with related SIGDAT events such as EMNLP. Creating a joint call for papers for EMNLP-CoNLL 2007 has learned us that CoNLL will remain having a particular focus with unique and complementary elements found nowhere else, and we intend to continue guarding these topics as laid out in the original charter of SIGNLL (<http://ifarm.nl/signll/about/>) as the core elements of CoNLL.

Antal van den Bosch  
Tilburg, The Netherlands  
May 21, 2007

=====

SIGPARSE Annual Report, June 2007

Harry Bunt

The main aim of SIGPARSE is to ensure the continuity of the biennial International Conferences on Parsing Technologies' (IWPT) series.

The last Conference in this series was the one with number 10, which was held in October 2005 in Vancouver, Canada. In the second half of 2006 and the first half of 2007, the organization of the 11th conference has taken shape. IWPT 2007 is a two-day satellite event of ACL 2007 in Prague, with Paola Merlo as program chair, Alon Lavie as logistic arrangements chair, and Harry Bunt as general chair.

IWPT 2007 features Stuart Shieber as invited speaker.

To facilitate its operation and the communication in the parsing community, the SIGPARSE website continues to be maintained at the University of Twente, and a mailing list at CMU in Pittsburgh.

Harry Bunt, May 2007.

=====

SIGMORPHON (Computational Morphology and Phonology)  
2006-2007 Annual Report  
Jason Eisner

MEMBERSHIP  
-----

SIGPHON is ACL's special interest group for computational morphology, phonology, and phonetics. Membership currently stands at 73, up from 55 at the end of July 2006.

(All of these members have explicitly affirmed their membership in the past year or so. Our membership list was previously 240, but this included many inactive members, including many with non-working email addresses. We cleaned it up by explicitly asking people to re-subscribe.)

A new SIGPHON executive committee was elected in 2006, consisting of 3 old and 3 new members.

WORKSHOP  
-----

Our 9th SIG workshop is being held at ACL 2007. (We have customarily held workshops only in even-numbered years, and we did have one in 2006, but there was sufficient interest to hold another this year.)

All 10 talks, 5 of the 6 posters, and the keynote address are related to the workshop's special theme: COMPUTING AND HISTORICAL PHONOLOGY. This special interdisciplinary workshop includes papers not only from computational linguists, but also from a historical linguist, a typologist, a dialectologist, and a geneticist.

The program is at <http://www.let.rug.nl/alfa/Prague/advertisement.html>  
All submissions received 3-4 reviews. The workshop organizers are John Nerbonne, T. Mark Ellison and Grzegorz Kondrak.

ONLINE ACTIVITIES  
-----

sigmorphon.org continues to serve the community by maintaining a

mailing list and online bibliographies at its web site. The website does not fully reflect this new domain name but will be updated this summer.

We plan to advertise the SIG's expanded scope in order to recruit new members to the SIG and increase discussion on its mailing list.

We have corresponded recently with Steven Bird about adding some missing older workshop proceedings to the ACL Anthology. These additional proceedings (with one exception) are already available at sigmorphon.org.

=====

SIGSEM

Johan Bos

SIGSEM elections were held in February 2007 with the following results:

Johan Bos (President)  
Alexander Koller (Secretary)  
Paul Buitelaar (Information Officer)  
Harry Bunt (Information Officer)  
Katrin Erk (Information Officer)  
Claire Gardent (Information Officer)

One of the first activities the new SIGSEM board undertook was changing the SIGSEM webpages into a wiki based site where all SIGSEM officers can easily change and update information. Also, the SIGSEM newsletter was resurrected with a fresh issue sent in May this year to all SIGSEM members, reporting on the SIGSEM working groups and upcoming events. The current SIGSEM membership counts 614 members [May 18, 2007].

The major event organised by SIGSEM in 2007 was IWCS7, the seventh International Workshop on Computational Semantics, which was held from 10 - 12 January 2007 in Tilburg, Netherlands. A detailed conference report was compiled by Jennifer Spender and is published on the SIGSEM webpage.

The events endorsed by SIGSEM in 2006/2007 are:

- [1] OntoLex07  
From Text to Knowledge: The Lexicon/Ontology Interface  
Workshop at ISWC07 (6th International Semantic Web Conference)  
November 11, 2007; Busan, South Korea.
- [2] Joint meeting of ACL-SIGSEM Working Group on the Representation of Multimodal Semantic Information and the ISO TC 37/SC 4 expert group on Semantic Content.  
January 8-9, 2007; Tilburg, Netherlands.
- [3] The 10th Workshop on the Semantics and Pragmatics of Dialogue  
Brandial 2006  
September 11-13, 2006; Potsdam, Germany.
- [4] Fourth ACL-SIGSEM Workshop on Prepositions  
June 28, 2007; Prague, Czech Republic.

[5] ACL-PASCAL Workshop on Textual Entailment and Paraphrasing  
June 28-29, 2007; Prague, Czech Republic.

[6] The Sixteenth Amsterdam Colloquium  
December 17-19, 2007; Amsterdam, Netherlands.

=====

SIG Semitic

Shuly Wintner

The SIG is healthy. We have a SIG meeting this year as an ACL Workshop, which attracted over 30 submissions. We have a working web site and a low-key but active mailing list. The SIG has over 150 members. Last elections were held December 2005, and we plan the next elections for 2008.

Shuly Wintner

=====

SIGWAC report to ACL  
Adam Kilgarriff

23 May 2007

SIGWAC

1) now has an 18-strong committee.

2) will hold its third workshop (WAC3) in Louvain-la-Neuve, Belgium, 15-16 September, (with an invited speaker supported by ACL.) Co-chairs: Cedrick Fairon and Gilles-Maurice de Schryver

3) is organising the first CLEANVAL - shared task and competitive evaluation on turning arbitrary web pages into clean text. This will take place in June 2007 with results discussed at WAC3 in Belgium in September.

4) has developed a proposal for ACL to host a M-linguistic search engineM-T: proposal to follow to the ACL committee separately.

Adam Kilgarriff  
Chair

=====

Report on NACLO (North American Computational Linguistics Olympiad)

Lori Levin, co-chair, lsl@cs.cmu.edu  
Thomas Payne, co-chair, tpayne@cs.uoregon.edu  
Dragonir Radev, program chair, radev@umich.edu

Introduction  
-----

NACLO (formerly known as NAMCLO), <http://www.namclo.org>, is an Olympiad style contest for high school students. The first contest was held on March 29, 2007. 195 high school students participated in Pittsburgh, Boston, Ithaca, and "the internet". The problems and solutions and names of the winners are available on the web site. We

are now planning to take the top four winners to the International Linguistics Olympiad (ILO) in St. Petersburg, Russia this summer.

Goals of NACLO  
-----

1. Increase the size and diversity of the pool of future scientists in Linguistics, Computational Linguistics, and Human Language Technologies.

2. Identify talented high school students and help them get the background that they need for higher education in Linguistics, Computational Linguistics, and Human Language Technologies

3. Get the scientific study of language into high school curricula (in cooperation with the LSA's Language in the School Curriculum committee)

4. Identify foundational skills that can be taught before college that can prepare students for coursework in linguistics, computer science, and language technologies.

NAACL Sponsorship of NACLO  
-----

NAACL contributed \$2000 to the 2007 competition. The money was used for prizes (along with additional sponsorship money from Google and books from Cambridge University Press). Prizes were awarded as follows:

National First Place: \$500  
National Second Place: \$350  
National Third Place: \$200

Local First Place (four locations):  
\$150 and the Cambridge Encyclopedia of Language  
Local Second Place (four locations plus one tie):  
\$100 and the Cambridge Encyclopedia of Language  
Local Third Place (three locations):  
\$50 and the Cambridge Encyclopedia of Language

Best solution (eight problems with four ties = 12 recipients):  
\$50 each

Total: \$2900 (\$2000 from NAACL, the rest from Google)

There was an error in not including the NAACL logo on the web site, which has now been corrected. The NAACL logo was, however, on the exam booklet and on large posters in Pittsburgh, Boston, and Ithaca.

In addition, NAACL was verbally acknowledged in Pittsburgh in front of 94 students and several teachers.

History of Linguistics Olympiads  
-----

Linguistics Olympiads have been held since the 1960's in Russia and since the 1980's in Bulgaria. They are now also held in other countries such as the Netherlands and the UK. A Linguistics Challenge was also held several times in Eugene, Oregon. There is now an International Linguistics Olympiad, whose 5th annual contest is July 31 to August 4 in St. Petersburg, Russia ([www.ilolympiad.spb.ru](http://www.ilolympiad.spb.ru)).

This year's ILO is preceded by a one week summer school in Narva-Joesuu, Estonia.

History of NACLO  
-----

About one year ago, Tanya Korelsky (US National Science Foundation) suggested that North America should participate in the ILO and also have its own Olympiad contest. Lori Levin agreed to apply to NSF for a grant in order to hold a planning workshop. The workshop was held in September, 2006 at the Interspeech conference in Pittsburgh, PA. About forty people attended the workshop including representatives from NAACL and ACL, high school teachers, representatives from the ILO and Moscow Linguistics Olympiad, organizers of other contests, and faculty and graduate students from about 5 other universities.

At the workshop, it was determined that Spring would be the best time for the competition. It was decided to ambitiously try for a pilot competition in Spring 2007, with only six months to plan and organize it.

Officers were elected at the workshop:

Co-chairs: Lori Levin and Thomas Payne  
Program chair: Dragomir Radev  
Follow-up chair: Barbara Di Eugenio  
Outreach chair: William Lewis  
Sponsorship chair: James Pustejovsky

Other jobs were created later:

School liaison: Amy Troyani  
School practice program: Rebecca Hwa and Noah Smith  
Administrative assistant: Mary Jo Bensasi

Locations of NACLO  
-----

NACLO was held in Pittsburgh, Boston, and Ithaca. In addition, students could participate in remote locations provided that they were monitored by a teacher. Philadelphia planned to participate, but couldn't due to the tight time line.

Program  
-----

The program chair (Dragomir Radev) is responsible for soliciting and reviewing problems, choosing problems for the competition, and supervising the grading of problems. A call for problems was issued in October 2006. Problems were submitted and reviewed in November and December 2006. The submitted problems were split into practice problems, which were posted on the web, and potential competition problems, which were held secretly in reserve. After NACLO was publicized on the ACL mailing list, additional problems were submitted. Eight problems were chosen for a the competition, some from US faculty and some from foreign faculty with experience in the ILO. The length of the exam was five hours.

Types of problems: We attempted to augment the ILO-style linguistic puzzles with some problems that focus more on computational concepts and applications such as web search and OCR. Pedagogical goals of the problems and effective problem types are ongoing areas of research.

Outreach  
-----

The outreach chair, William Lewis, was responsible for the web site, practice session curriculum that would be presented in schools, and other kinds of publicity. Linguist List joined to help with the web site, with additional funding from NSF. A school practice session was designed by Rebecca Hwa and Noah Smith, including information about linguistics and computer science and practice problem solving.

Several practice sessions were held in Pittsburgh, Boston, and Ithaca. There were also some press releases and some postings in newspaper calendars.

Sponsorship  
-----

The sponsorship chair was James Pustejovsky. Since this was our first pilot year, we did not contact very many potential sponsors. Google contributed \$6000, NAACL contributed \$2000, and Cambridge University Press contributed 12 copies of the Encyclopedia of Language.

Follow up  
-----

The follow up chair (Barbara Di Eugenio) is responsible for getting feedback on the competition and for setting up clubs and mentoring programs so that students know how to find appropriate universities and courses in order to continue studying linguistics and language technologies. The main follow up activities for this year were an evaluation form and the trip for four students to the ILO summer school and competition.

The evaluation form includes a variety of questions including how did you hear about NACLO, would you participate again, how can we make it better, etc. Students were also asked to rate each problem. Results have been tabulated and are available on request.

Diversity  
-----

About half of the top 50 scorers are girls. The national first place winner is a girl.

One of the goals of NACLO is to introduce linguistics and language technologies to populations that are under-represented in the field. However, we did not make any special effort in this direction for the 2007 contest. For example, in Pittsburgh, the contest was publicized through the network of program coordinators for gifted students at high schools. There is some economic diversity in gifted programs, and a bit of racial diversity, but not an overwhelming amount.

The future of NACLO and future financial needs  
-----

The 2007 NACLO recruited 195 participants with less than eight weeks of minimal publicity. We plan to start school outreach sessions in the Fall of 2007 for the 2008 contest. In addition, there will be more cities participating. We are estimating that the 2008 NACLO will have over 1000 participants. There will probably be an internet-based qualifying round.

Our activities for the next year will focus on: establishing a non-profit organization, writing a constitution, refining the lesson plan for high school practice sessions, and refining the types of exam problems to meet pedagogical goals.

The expenses for NACLO include:

1. Operating expenses such as copying, mailing, and room rentals and lunch for the participants.
2. Prizes: We expect the 2008 prizes to be similar to the 2007 prizes.
3. Trip to the ILO: This year's trip to the ILO for four students and two chaperones will cost at least \$17,000. The money is coming from Google and NSF's office of international programs. The costs includes a day of training in New York, 250 euros per student for the summer

school in Estonia, plane tickets, and some hotel and food costs for days that are not covered by the ILO and summer school. The ILO pays for food and lodging during the competition, but since the ILO is growing, the 2008 ILO may not be able to provide food and lodging for all participants. For example, they may be able to cover the costs for one team per country, but several countries send more than one team. NACLO certainly plans to send more than one team in 2008.

4. Personnel: This year there was minimal coverage of faculty time and administrative assistant time through the NSF planning grant. It would be good to work out a way to get more coverage for faculty and administrative salaries. For comparison: chairing NACLO is takes at least as much time as being general chair of a conference; sponsorship and administration also take about the same amount of time as for a conference; program committee takes at least as much time and could take more. For example, the Moscow program committee meets weekly year round and may refine each problem for several months; local arrangements takes more time than for a confernece because of outreach to highschoools and practice sessions. This is ok for one year at a time, but a few devoted people will probably be involved on an ongoing basis and could use some salary coverage.