

Project Summary: Student Research Workshop in Computational Linguistics, at the ACL 2005 Conference

The Association for Computational Linguistics (ACL) is the primary international organization in the field of natural language processing and language engineering. The Association's annual conference is the *major* international conference in the field.

This is a proposal for funding to subsidize travel, conference and housing expenses of students selected to participate in the Association of Computational Linguistics (ACL) Student Research Workshop to be held on June 27th, 2005 in Ann Arbor, Michigan. In its current format, the workshop consists of paper presentations by student authors and will take place during the main conference. Each presentation is followed by a discussion panel which consists of leading researchers in the field. The workshop is organized and run by students.

The workshop, which has existed in some form since 1991, has gained an increasing popularity over the years. This year, the workshop received a record number of submissions — 70 papers authored by graduate and undergraduate students, a clear evidence of the enthusiastic reception of the event in the community.

Intellectual Merit: The Student Research Workshop provides a valuable opportunity for the next generation of computational linguists to enter the CL community. It allows the best students in the field to take their first important step towards becoming professional computational linguists by receiving critical feedback on their work from experts outside their advisory team, and making contact with other students and key researchers in their field. This has provided all the students involved with invaluable opportunities for professional growth and interaction with senior researchers on the organizing committee of the main conference.

Broad Impact: The student research workshop contributes to the maintenance and development of a skilled and diverse computational natural language workforce, helping to produce a pool of researchers with the scientific and engineering knowledge for Natural Language Engineering. The workshop encourages a spirit of collaborative research, and builds a supportive environment for a new generation of computational linguists.

1 Project description

This is a proposal for funding to subsidize travel, conference and housing expenses of students selected to participate in the Association of Computational Linguistics (ACL) Student Research Workshop to be held on June 27th, 2005 in Ann Arbor, Michigan.

The Student Research Workshop provides a valuable opportunity for the next generation of computational linguists to enter the CL community. It allows the best students in the field to take their first important step towards becoming professional computational linguists by receiving critical feedback on their work from experts outside their advisory team, and making contact with other students and key researchers in their field. Another benefit of past ACL student sessions is that they have been organized and run by students. This has provided all the students involved with invaluable opportunities for professional growth and interaction with senior researchers on the organizing committee of the main conference.

This year, the workshop attracted 70 submissions from 19 different countries. The sharp increase in the number of submissions from previous years convinced us that the current workshop format is indeed a stimulating environment for students to present their work and be exposed to outside perspectives at a critical time in their research.

For the first time, the 2005 ACL Student Research Workshop will include presentations by both beginning and advanced graduate students. The submissions are organized in two tracks. The emphasis of the first track is original and unpublished research, that is still in progress. This track is aimed at students who have settled on their thesis direction but who still have significant research left to do. The emphasis is on presenting significant research towards a PhD, with the opportunity of getting feedback from the community before submitting the thesis. This framework will further facilitate interaction between students at various stages of their academic careers and improve the quality of the presented papers.

To sustain the outreach tradition of the student workshop, the proposal requests funding for student participants. Students are less likely than faculty to be able to pay the expenses associated with conference attendance. This is especially true for participants who will benefit the most from the workshop (i.e., students that are not members of established research groups.)

2 Background

The student workshop at the Association for Computational Linguistics annual meeting has existed in some form since 1991. Statistics on the number of accepted papers (ACC) and submitted papers (SUB) at past ACL student sessions are shown below:

year	# accepted	# submitted	format
1991	14	unknown	4 sessions, 2 at a time, 20 minutes per student
1992	20	48	6 sessions, 2 at a time, 20 minutes per student
1993	11	30	1 session, 18 minutes per student
1994	10	41	2 sessions in parallel, 20 minutes per student
1995	19	48	poster session, 2 hours and 30 minutes
1996	14	32	poster session, 1 hour and 45 minutes
1997	10	42	two sessions, 20 minutes per student
1998	12	46	four lunch-time sessions, 20 minutes per student
1999	10	30	4 paper sessions, 2 at a time, 25 minutes per student
2000	10	36	4 paper sessions, 2 at a time, 30 minutes per student
2001	12	50	3 parallel paper sessions, 30 minutes per student
2002	13	29	3 parallel paper sessions, 30 minutes per student
2003	12	31	3 parallel paper sessions, 30 minutes per student
2004	10	40	1 session, 30 minutes per student

3 Goals

The specific goals of 2005's workshop build upon the goals of past student sessions. During the workshop, the students selected to participate will be given the opportunity to present research in progress and receive feedback from members of the computational linguistics research community. This would provide a valuable opportunity for students to not only gain exposure within the research community but also to solicit comments on their research from a broader audience than would otherwise be possible. Students at all levels working in any area of computational linguistics will be encouraged to participate.

Goals of the workshop are:

- To help doctoral students with their research by providing advice on participant's current research and direct feedback from peers, faculty, and other researchers
- To facilitate networking events that will allow doctoral students to meet students and faculty from other universities, and industry representatives
- To provide advice on research methodology and career paths in industry and academia
- To encourage a supportive community of scholars and a spirit of collaborative research

4 Organizing Committee Information

The Student Research Workshop Student Co-chairs are Chris Callison-Burch (Edinburgh University) and Stephen Wan (Macquarie University).

Additionally, the committee is supported by a faculty advisor, Dr. Regina Barzilay, from the Massachusetts Institute of Technology. Her role is to help students in developing the structure of the workshop and to serve as a liaison between workshop chairs and the ACL organizing committee. The ACL Executive Committee is monitoring this effort.

5 Posting for the Call of Participation and Event Information

The ACL Student Research Workshop will be held during ACL 2005 (a conference of the Association for Computational Linguistics) on the 27th of June, 2005 in Ann Arbor, Michigan. The call for participation was made available on the ACL2005 conference web site on December 10th. It can be found at URL: <http://www.aclweb.org/acl2005/index.php?studentws>. The web site is a focal point for students to find out how to submit a paper, and later, about the format of the workshop (for students presenting papers or students attending the workshop). It includes not only technical information about submitting and presenting papers, but also information about the structure of the workshop and social activities held in conjunction with the workshop. Information about the format of the presentations will be posted once we have determined how many papers will be accepted. Appendix A contains a text version of the call for papers that was printed from the web site.

We sent the call for papers not only to a wide array of CL-themed mailing lists, but also directly to a large number of relevant departments in universities worldwide.

6 Submission and Paper Selection

Papers were due on February 8th, 2005 and will be selected based on a review process involving both students and established researchers in the area. The review process is a strength of this student workshop. The ratio between student and senior reviewers – 18 to 26 – ensures that every paper

will be reviewed by at least one representative from each group. Thus, students will benefit from expert input from outside their thesis advisory team; additionally, the students on the reviewing committee are gaining valuable professional experience as impractical critics of the work of their peers.

The original list of reviewers has been approved by the ACL Executive Committee. The reviewers are representative of the research areas covered by the organization as well as the geography of the international research community.

The workshop covers the following topics:

- pragmatics, discourse, semantics, syntax and the lexicon
- phonetics, phonology and morphology
- linguistic, mathematical and psychological models of language
- information retrieval, information extraction, question answering
- summarization and paraphrasing
- speech recognition, speech synthesis
- corpus-based language modeling
- multi-lingual processing, machine translation, translation aids
- spoken and written natural language interfaces, dialogue systems
- multi-modal language processing, multimedia systems
- message and narrative understanding systems

Student reviewers include Laura Alonso (U. of Barcelona), Colin Bannard (Edinburgh U.), Phil Blunsom (U. of Melbourne), Marine Carpuat (HKUST), Trevor Cohn (U. of Melbourne), Hal Daume III (U. of South California), Elena Filatova (Columbia U.), Erin Fitzgerald (JHU), Ben Hutchinson (Edinburgh U.), Corrin Lakeland (Otago U.), Maria Liakata (Oxford), Daniel Midgley (University of Western Australia), Ani Nenkova (Columbia U.), Leif Nielsen (King's College), Bo Pang (Cornell U.), Jean-Philippe Prost (Macquarie), Charles Schafer (JHU).

Non-student reviewers include Timothy Baldwin (U. of Melbourne), Bernd Bohnet (U. of Stuttgart), Chris Brew (Ohio State U.), Stephen Clark (Oxford), James Curran (Sydney), Mona Diab (Stanford U.), Pablo Duboue (IBM), Dan Flickinger (Stanford), Pablo Gamallo (Universidade Nova de Lisboa), Graeme Hirst (U. of Toronto), Julia Hockenmaier (U. of Pennsylvania), Frank Keller (Edinburgh U.), Simon King (Edinburgh U.), Alistair Knott (Otago U.), Philipp Koehn (Edinburgh U.), Emiel Krahmer (Tilburg), Mirella Lapata (Edinburgh U.), Alon Lavie (CMU), Daniel Marcu (U. of South California), Cecile Paris (CSIRO), Steve Renals (Edinburgh U.), Philip Resnik (Maryland U.), Advait Siddharthan (Columbia U.), Michael Simard (Xerox), Matthew Stone (Rutgers), Mark Swerts (Tilburg), Leonoor van der Beek (Groningen), Florian Wolf (Cambridge).

We anticipate on accepting between 9 to 12 papers out of 70 submissions depending on the quality of submissions.

7 The Workshop

The Workshop will be organized into a day-long single session on the 27th of June 2005, running in parallel with the main sessions of the conference. At the workshop, each student will be allowed a presentation of 20 minutes in length followed by 15 minutes of discussion by the feedback panel

of researchers. During this time the panel will discuss in depth research issues related to the presentation and make suggestions for future directions. The panelists will be selected to cover areas relevant to the papers accepted to the workshop. The members of the feedback panel will receive copies of the student workshop papers prior to the workshop. All accepted papers will be published in a workshop proceedings.

The workshop will also feature an interactive panel in which established researchers will share their advice on research methodology, and career paths with a new generation of NLP researchers. The discussion will continue in the informal setting during student lunch (sponsored by the ACL).

The organizers will invite all students to attend and participate in the workshop, whether or not they apply to present their work. In previous years, many nonpresenting students said they found it useful to observe their peer's presentation and to participate in the ensuing discussions.

8 Budget

Funding from NSF would contribute toward stipends given to student participants for the cost of travel, registration, and housing during the workshop and conference. Each student will receive \$1,300 to cover expenses of attending the workshop and main conference (\$250 for housing, \$150 for conference registration, \$900 for travel expenses), for a total of up to \$15,600 (for 12 student presenters). In addition, the budget allocates for a best paper prize (\$500) and a percentage of the cost of the workshop proceedings (\$1500). The total amount requested in the proposal is \$17,600.

Panelists will receive no honorarium or other consideration for their efforts other than the gratitude of the research community and an acknowledgment in the workshop co-chairs' report in the proceedings. No additional support is being requested from other funding agencies.

9 Conclusions

As a means of encouraging young and upcoming computational linguists, the NAACL/ANLP Student Workshop seems to be an inexpensive and yet effective method. The more intimate workshop format encourages the student participants to begin building a rapport with established researchers. This nurturing effort should pay dividends by more effectively guiding students in this rapidly changing research field. In addition, by building a supportive environment for these students, it is more likely that down the road, these students will lend a supporting hand to other students who follow.

A Appendix: Call for Papers

Attached is a text version of the call for papers that was printed from the web site for the ANLP/NAACL Student Research Workshop.