Computational Linguistics UK (CLUK '06)

Case for Support - Part 1

Dr Alistair Willis has been a Lecturer at the Computing Dept., The Open University, Milton Keynes, UK, since July 1st, 2003. He was awarded a DPhil from the department of Computer Science at the University of York (June 2000) for his research into natural language processing, in which he developed efficient algorithms for reasoning with underspecified semantic representations. His thesis was selected as the department's entry for the BCS thesis award. His current research interests include methods for aligning (same language) texts across genres. He is also interested in how established techniques from software engineering, especially automated testing, may be applied to natural language systems to improve reliability and maintainability.

Prior to joining the Open University, Dr Willis was employed as a research scientist at Philips Research Laboratories, Redhill (2000-2003). At Philips Research, he worked in the software engineering research group. His research primarily investigated how to automatically generate and execute sets of test cases for reusable components in component based software, with particular interest in integration testing. He also investigated methods of creating "context aware" wearable computing applications.

Dr Willis has been an invited speaker at the 2004 joint COCOSDA and ICCWLRE meeting, and has refereed for the Automated Software Engineering Journal (2004). He currently co-supervises one PhD student on the automatic detection of dangerous ambiguity in requirements documents.

The Computing Department and Centre for Research in Computing

The newly established Centre for Research in Computing (CRC) at the Open University brings together internationally excellent research from the Computing Department, the Knowledge Media Institute and the User Group of the Institute for Education Technology. It undertakes distinctive, leading edge research in software and processes that underpin knowledge management, communication, learning, and interactions among humans and machines. Many of our projects build on the position of the Open University as the world's largest and best distance and e-learning organization, and seek to anticipate how to help learners in a networked world, both at home and at work

Natural Language Engineering is one of the Centre's key themes, and the group has recently expanded to include 2 professors (Professors Donia Scott and Anne De Roeck) as well as 7 further academic staff and about 10 research fellows and PhD students. Main research strands include language generation and document lay-out, corpus profiling and term distribution modelling, terminology based representations for information filtering and for ontology extraction. The group works closely together with the Semantic Web and Ontologies group led by Professor Enrico Motta and has strong links with the Statistics department.

By tradition, CLUK is organised by research students for research students with advice from research staff. This year's student organizers are Dileep Damle and Gaston Burek, both 3rd year students at the Open University who have attended CLUK workshops in the past. Their advisors are Prof. Anne De Roeck, Dr Alistair Willis and Dr Victoria Uren, all active researchers and members of the Natural Language Processing research group (http://mcs.open.ac.uk/NLP/) at the Open University. Dileep Damle's CV is also attached.

Professor De Roeck in particular is a long standing member of CLUK and has attended workshops and advised on their organisation since 1999. Professor Scott was employed at the University of Brighton, when that institution held the event in 2000. Therefore, both have experience of the issues in hosting the conference and are readily available to provide advice and support to the organisers.

Computational Linguistics UK (CLUK '06) March 2006

Case for Support - Part 2

Background

Computational Linguistics UK (CLUK) is an organisation dedicated to the support of Computational Linguistics (CL) research in the UK, in particular:

- to ensure that researchers are able to keep abreast of the many different threads of research in CL
- to establish links across institutions carrying out research in CL, and
- to provide support to PhD students, and to establish them with in the CL community.

The annual CLUK colloquium addresses all these aims, by providing the foremost annual opportunity for UK PhD students in Computational Linguistics and related disciplines to gain valuable feedback on their work in a workshop environment. As such, the colloquium provides essential research training for UK PhD students, and ensures that established researchers can see first hand the directions that future CL will take.

By participation, PhD students benefit by exchanging ideas, receiving feedback on the initial stages of their research and gain experience of writing and presenting papers. This experience is invaluable for the later stages of their PhDs, when they will be taking their work to the international academic community.

The workshop is organised by research students, and submissions are reviewed by a programme committee of research students. In short, the workshop provides experience of all aspects of a full workshop, with academic staff providing only the support as needed by the PhD students.

The annual workshop has proved extremely valuable and popular among the entire CL community in the UK, and is now in its ninth year. The 2006 event is being held at the Open University in Milton Keynes, which has a large CL community, spread across different departments.

The previous colloquiums have been held at the Universities of Sunderland (1998), Essex (1999), Brighton (2000), Sheffield (2001), Leeds (2002), Edinburgh (2003), Birmingham (2004) and Manchester (2005)[1].

Main Programme and Speakers

The main programme will follow the pattern of previous years' colloquiums, covering two days of presentations, a poster session and addresses by two invited speakers.

(a) Presentation and Poster session

The presenting participants are PhD students, typically in the second year of their studies.

We aim to accept 16 full papers and up to 16 posters. An call for abstracts has been sent out using the main mailing list in the field. Authors of accepted abstracts will be invited to present a full paper or poster.

---

[1] See http://nlp.shef.ac.uk/research/cluk

The success of the workshop is dependent upon the attendance of senior members of the academic community to provide feedback to the students in this session. The Open University has many members of staff who are active in the Computational Linguistics community who can provide this function, although our aim with this proposal is to maximise the number of established academics who are able to attend.

(b) Invited Speakers

The keynote speakers for this year's conference are Professor Jon Oberlander of the University of Edinburgh, and Professor Anne De Roeck of the Open University. Both speakers are extremely highly regarded within the Computational Linguistics community and have been chosen specifically to provide expert overviews and opinions on the broad range of subjects we expect to be represented at the main paper and poster presentations.

Dissemination of Results

The submitted papers will be available at the workshop as proceedings. Additionally, the webcam recordings of the conference, poster abstracts and an electronic version of the proceedings will be available via the colloquium website (http://kmi.open.ac.uk/events/cluk06/).

This year, in order to make the workshop as accessible as possible to the CL community, we intend to make the speaker presentations fully available via a webcast, with access to the recordings maintained for a minimum of 1 year after the event. This has not been attempted at previous CLUK workshops. However, the benefits will be that:

- students and their mentors will be able to review their own performance at later dates, and
- talks (and, importantly, the question and answer sessions) will be available to those who are unable to attend the event itself.

A real-time broadcast of the proceedings also enables interaction by those who are not able to attend in person.

The venue within the Open University where the presentations will be held has the means and the experience to implement this. The bulk of the funds that we have requested are to cover the cost of the technician who will manage the webcast and set up the website subsequently.

Project Management

By tradition, CLUK is organised by research students for research students with advice from research staff. This year's student organizers are Dileep Damle and Gaston Burek, both 3rd year students at the Open University who have attended CLUK workshops in the past. Their advisors are Prof. Anne De Roeck, Dr Alistair Willis and Dr Victoria Uren, all active researchers and members of the Natural Language Processing research group (http://mcs.open.ac.uk/NLP/) at the Open University. Professor Donia Scott has recently joined the department from the University of Brighton, and is able to provide the organisers with her experiences from the 2000 event, held at Brighton.

In keeping with the traditions of the workshop, abstract refereeing will be carried out primarily by students, with relatively little additional feedback from permanent academic staff. This aims to give PhD students experience with reviewing submissions. A student review committee has been established using PhD students both from the Open University and other institutions across the UK. Abstracts will be reviewed by at least two reviewers from different institutions.

Submission is via the workshop web site, which is maintained by the organisers (http://kmi.open.ac.uk/events/cluk06/).

All administration is being supported by the Open University.

The Open University is strategically located, being easily accessible by many other institutions who are active in Computational Linguistics, in particular, London, Oxford and Cambridge. This will make it easier for other eminent UK based researchers (about 6 have been invited) to attend, who will be able to provide high value feedback.

Relevance to Beneficiaries

CLUK has proved to be a highly beneficial event for research students. The technical benefit comes from feedback from senior researchers and experienced students. This also encourages PhD students to engage with the national community at the earliest possible stage in their studies, while exposure to different areas of their subject gives a greater depth of knowledge of their field. The early opportunity to submit and present work also provides invaluable practice in communication and networking skills, that the students will use throughout their career.

Computational linguistics is an import area of computing research, with the major conferences, ACL and COLING attracting very high quality submissions from international participants. It is also an important area of research in the UK. Because there are many strands of research within CL, events such as this are vital to maintain long term expertise. By providing a forum for current PhD students to discuss their work, researchers in CL can maintain their knowledge of existing and future areas of research.

Many PhD students who have presented at previous CLUK workshops have gone on to present their work at the major international conferences.

The Relationship to Similar Workshops/Schools

The CLUK colloquium is unique in being held in the UK, being specifically and only for presentations by research students and covering the range of sub-disciplines of Computational Linguistics. In comparison with other events such as the workshops associated with the major European and international conferences such as COLING (biennial), and ACL and EACL, which are specific to subfields of CL, CLUK covers a broader spectrum of subfields of CL and will have a higher rate of acceptance.

Research students who present their work at CLUK frequently use the experience as a stepping stone for submission to the major international conferences.

The Numbers of Participants and Target Audience

We are targeting all the universities which have CL research groups that provide research training in the UK. We hope to attract 16 full paper submissions and 16 posters. The event will also be attended by PhD students, some PhD supervisors and OU senior researchers in the field. In total we are aiming at an attendance of about 50 participants.

Justification of Resources

For a workshop of this kind, it is important to keep the registration fee to a minimum, to allow maximum participation by students, and academic staff from other institutions. In particular, because of its unique nature, the Open University has no facilities for cheap student accommodation available, and so a high registration fee may be prohibitive.

We request expenses for the invited speaker at £400 including 2 nights' accommodation, travel from Edinburgh and expenses. This figure allows us to invite a prominent expert in field who can set out the agenda and trends for the future direction of research in this field. One of the keynote speakers (Prof De Roeck) is from the Open University and will not be incurring additional expenses.

The webcam recording will require one experienced technician to be available for the two days of the colloquium, one day in preparation, and a further two days afterwards for preparation and archiving of the recorded material. We also request funds for five days paid for Ben Hawkridge, a member of technical staff with this expertise. For the five day period, this cost is £948 (5 days at £189.44 per day).

In total, we request £1348 in funding for the travel and expenses of the invited speaker and the technical costs of broadcasting the proceedings. These funds will allow us to reduce the registration fee and maximise the number of participants able to attend. Fees for previous two-day events have been up to £70.

Plans for a User-Based Evaluation for the Colloquium

We intend to collect user feedback at the end of the event, collecting information on the students' perception of the value of the workshop. In particular, we will look at:

− skills gained as a result of participating in the workshop, in particular on paper writing and submission, and networking, and the experience of refereeing.
− how far students feel they have broadened their knowledge of the field in current important research topics and key people.
− new insights that the students will use in their research.

The success of the webcast will be measured by monitoring the number of hits on the website during the event and for a period of six months after.

Computational Linguistics UK (CLUK '06) March 2006

Case for Support - Part 3

Time plan:

16th November 2005	First call for astracts
2nd January 2006	Deadline for Abstract Submission
28th January 2006	Notification of Acceptance
18th February 2006	Camera Ready Papers and Posters Copy
25th February 2006	Programme finalised and published
8th-9th March 2006	CLUK 2006 Workshop