

From: Desiree Cleves <admin@clsp.jhu.edu>
Subject: Summer Research for Undergraduates
Date: February 24, 2009 11:57:21 AM EST

To: chitta@asu.edu, Wayne.Raskind@asu.edu, kchang@eng.auburn.edu, mco@bu.edu, bdh@bradley.edu, mcasey@bradley.edu, jamesp@cs.brandeis.edu, jensen@ee.byu.edu, rt@cs.brown.edu, jhoff@math.brown.edu, Constantine_Dafermos@Brown.edu, bill_warren@Brown.edu, tan@het.brown.edu, gatsonis@stat.brown.edu, clifton@engin.brown.edu, sfkaplan@cs.amherst.edu, waloinaz@amherst.edu, stephen.phillips@asu.edu, robert.nemanich@asu.edu, perez@physics.auburn.edu, nelmsrm@eng.auburn.edu, smith01@auburn.edu, Ben_Kelley@baylor.edu, ralph@bu.edu, rebbi@pthind.bu.edu, fcerrina@bu.edu, slimedog@bradley.edu, jcm@bradley.edu, charney@brandeis.edu, tim@cs.brandeis.edu, aburdene@bucknell.edu, xmeng@bucknell.edu, jlavine@bucknell.edu, schoepf@bucknell.edu, jkat@caltech.edu, flach@caltech.edu, mathieu@caltech.edu, hassibi@caltech.edu, ivakalis@csc.calpoly.edu, amaccarl@calpoly.edu, drawing@calpoly.edu, mmoelter@calpoly.edu, rsmidt@calpoly.edu, shahian@csulb.edu, nagi.elnaga@csun.edu, aghandakly@csuchico.edu, bengiami@csufresno.edu, chrisg@csufresno.edu, dougs@csufresno.edu, stefaan_delcroix@csufresno.edu, mshiva@fullerton.edu, sbarua@fullerton.edu, fmueller@fullerton.edu, zhangd@gaia.ecs.csus.edu, sgs@csun.edu, werner.horn@csun.edu, vadhava@gaia.ecs.csus.edu, Victor.Aitken@sce.carleton.ca, kalyniak@physics.carleton.ca, randall_gess@carleton.ca, jondich@carleton.edu, mark@stat.cmu.edu, ed@ece.cmu.edu, Peter.Lee@cs.cmu.edu, franklin@ernest.phys.cmu.edu, zmo@eecs.case.edu, dxc57@case.edu, dsa5@case.edu, jill.korbin@case.edu, jventura@cbu.edu, pbedross@cbu.edu, jvarrian@cbu.edu, ortmeyer@clarkson.edu, pturner@clarkson.edu, privman@clarkson.edu, lfth@clemson.edu, ddarren@clemson.edu, f.xiong@csuohio.edu, s.misra@csuohio.edu, G.LUPTON@csuohio.edu, M.KAUFMAN@csuohio.edu, kearns@cs.wm.edu, griff@physics.wm.edu, jbreidt@stat.colostate.edu, dieter@lamar.colostate.edu, aam@engr.colostate.edu, whitley@cs.colostate.edu, tavener@math.colostate.edu, sfchang@ee.columbia.edu, wait@cpw.math.columbia.edu, nowick@cs.columbia.edu, paknys@ece.concordia.ca, chair@mathstat.concordia.ca, mfrank@alcor.concordia.edu, mudur@cse.concordia.ca, chair@cs.cornell.edu, jdb2@cornell.edu, director@infosci.cornell.edu, jsb2@cornell.edu, barbasch@math.cornell.edu, tc345@cornell.edu, rockmore@cs.dartmouth.edu, prasad@cs.dartmouth.edu, Erland.Schulson@Dartmouth.edu, lindsay.whaley@Dartmouth.edu, hugo@math.drexel.edu, eileen.abels@ischool.drexel.edu, guceri@drexel.edu, alan@stat.duke.edu, gauthier@phy.duke.edu, jitel@duke.edu, pankaj@cs.duke.edu, lcollins@ee.duke.edu, lesage@polymtl.ca, milton.cone@erau.edu, vss@mathcs.emory.edu, kwanrcke@physics.emory.edu, sharma@fdu.edu, alfredo_tan@fdu.edu, victor.debrunner@eng.fsu.edu, wam@physics.fau.edu, borko@cse.fau.edu, smorgera@fau.edu, horswell@fau.edu, spyros@fau.edu, wds@cs.fit.edu, lakshmik@fit.edu, toswalt@fit.edu, sullivan@fit.edu, banko001@gannon.edu, headley@gannon.edu, frezza001@gannon.edu, afb@cc.gatech.edu, compstat@galaxy.gmu.edu, rehrlich@gmu.edu, hgomaa@gmu.edu, amanitiu@gmu.edu, jchriste@gmu.edu, kfischer@gmu.edu, blakeb@cs.georgetown.edu, schiffd@georgetown.edu, meiyin.chou@physics.gatech.edu, gary.may@ece.gatech.edu, trotter@math.gatech.edu, mcdonald@kettering.edu, labay@gonzaga.edu, mckenzie@gonzaga.edu, greera@gonzaga.edu, rgjewell@gcc.edu, WPBirmingham@gcc.edu, meng@stat.harvard.edu, greg@eecs.harvard.edu, chierch@fas.harvard.edu, yau@math.harvard.edu, stubbs@physics.harvard.edu, scarey@wjh.harvard.edu, mike@cs.hmc.edu, abright@hmc.edu, depillis@hmc.edu, townsend@hmc.edu, sphrlb@hofstra.edu, cscgk@hofstra.edu, matsbs@hofstra.edu, Harold.Hastings@hofstra.edu, richard.j.puerzer@hofstra.edu, mohamed.chouikha@gmail.com, blegand@scs.howard.edu, ayakubu@howard.edu, G.Jenkins@howard.edu, hickernell@iit.edu, ms@iit.edu, korel@iit.edu, daaschliman@indianatech.edu, steffen@ipfw.edu, ngp@ipfw.edu, ychen@iupui.edu, sfeng@cs.iupui.edu, arun@iastate.edu, chang@cs.iastate.edu, kkoehler@iastate.edu, grue@ksu.edu, virg@cis.ksu.edu, dzollman@phys.ksu.edu, lpigno@math.ksu.edu, jwneill@ksu.edu, jounyi@lafayette.edu, cs_chair@cs.lafayette.edu, salwachc@lafayette.edu, hoffmang@lafayette.edu, h.mylar@ee.lamar.edu, ljosborne@my.lamar.edu, johnston@ltu.edu, bindschadler@ltu.edu, keith.hansen@lamar.edu, fjb205@lehigh.edu, hfk@lehigh.edu, mjsa@lehigh.edu, aravena@ece.lsu.edu, cox@latech.edu, sawyer@phys.latech.edu, kjeldaas@latech.edu, schroder@coes.latech.edu, bfitzpatrick@lmu.edu, bmarino@lmu.edu, nula@lmu.edu, vcoletta@lmu.edu, bruce.liby@manhattan.edu, jay.friedenberg@manhattan.edu, robert.mauro@manhattan.edu, carol.hurwitz@manhattan.edu, william.hudson@mnsu.edu, steven.case@mnsu.edu, clough@mcs.mu.edu, Benjamin.Brown@Marquette.edu, Edwin.Yaz@Marquette.edu, welg@csail.mit.edu, goemans@math.mit.edu, heim@mit.edu, msur@mit.edu, se_smith@mit.edu, marcia@csail.mit.edu, vkaspi@physics.mcgill.ca, keith.franklin@mcgill.ca, david.plant@mcgill.ca, dudek@cs.mcgill.ca, glyne.piggott@mcgill.ca, chair.math@mcgill.ca, haykin@mcmaster.ca, mohrens@mcmaster.ca, boden@math.mcmaster.ca, venus@physics.mcmaster.ca, allen_r@mercer.edu, peters_rd@mercer.edu, Michael.R.Belote@mercer.edu, leonard_ms@mercer.edu, noonanm@merrimack.edu, michael.bradley@merrimack.edu, whitedp@merrimack.edu, adamsj@merrimack.edu, grotjohn@egr.msu.edu, mutka@cse.msu.edu, prestel@msu.edu, chair@pa.msu.edu, schmitt@msu.edu, ywang@math.msu.edu, fuhrmann@mtu.edu, msgocken@mtu.edu, mneuman@mtu.edu, linda@mtu.edu, Jacques@msoe.edu, durant@msoe.edu, petersen@msoe.edu, hodges@cse.msstate.edu, razzaghi@math.msstate.edu, man@ra.msstate.edu, sbk1@ra.msstate.edu, rscherl@monmouth.edu, jcoyle@monmouth.edu, paxton@cs.montana.edu, bowers@math.montana.edu, smith@physics.montana.edu, block@montana.edu, rmaher@ece.montana.edu, lupton@morgan.edu, ngguerek@jewel.morgan.edu, ali.emdad@morgan.edu, pscottjo@morgan.edu, cjscott@eng.morgan.edu, fwooliver@morgan.edu, pjd@nps.edu, jknorr@nps.edu, luscombe@nps.edu, timothy.n.chang@njit.edu, pmorandi@nmsu.edu, epontell@cs.nmsu.edu, michael.a.baltrush@njit.edu, gibbs@nmsu.edu, shoran@nmsu.edu, daljit.ahluwalia@njit.edu, ysaito@nyit.edu, ajafari@nyit.edu, kolm@cims.nyu.edu, rbcole@stevens-tech.edu, mhw@cs.nyu.edu, jck@ncat.edu, dbdunn@ncat.edu, gvdozier@ncat.edu, mersrc@ncat.edu, loek@math.ncsu.edu, kolbas@eos.ncsu.edu, vouk@ncsu.edu, pantula@stat.ncsu.edu, nagle@ncsu.edu, paesler@ncsu.edu, Dan.Ewert@ndsu.edu, jim.coykendall@ndsu.edu, brian.slator@ndsu.edu, mcowen@neu.edu, s.sridhar@neu.edu, laf@ccs.neu.edu



3 Attachments, 435 KB

Save

(Please circulate as appropriate.)

NSF-Supported Summer Internships for Undergraduates

The Center for Language and Speech Processing at the Johns Hopkins University is seeking outstanding members of the current junior class to participate in a summer workshop on language engineering from June 8th to July 31st, 2009. The eight-week workshop provides a stimulating intellectual environment and we hope it will encourage students to eventually pursue graduate study in the field of human language technologies.

The summer workshop provides:

- * An opportunity to explore an exciting new area of research
- * A two-week tutorial on speech and language technology
- * Mentoring by an experienced researcher
- * A \$5,000 stipend and \$2,520 towards per diem expenses
- * Private furnished accommodation for the duration of the workshop
- * Travel expenses to and from the workshop venue

Applications must be received by FRIDAY, MARCH 20, 2009, and should include the name of a faculty nominator who would be willing to serve as a reference if asked. Applications from non-U.S. citizens currently residing outside the United States must be received no later than FRIDAY, MARCH 6, 2009. Apply online here:

<http://www.clsp.jhu.edu/workshops/ws09/app/>

Applicants are evaluated only on relevant skills, employment experience, past academic record, and the strength of letters of recommendation. No limitation is placed on the undergraduate major. Women and minorities are encouraged to apply.



Center for Language and Speech Processing

February 16, 2009

Dear Madam/Sir:

The Center for Language and Speech Processing at Johns Hopkins University is offering a unique summer internship opportunity, which we would like you to bring to the attention of your best students in the current junior class.

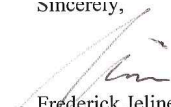
This internship is unique in the sense that the selected students will participate in cutting edge research as full members alongside leading scientists from industry, academia, and the government. The exciting nature of the internship is the exposure of the undergraduate students to the emerging subjects in human language technologies, such as *automatic speech recognition, natural language processing, speaker identification* and *statistical machine translation*.

We are specifically looking to attract new talent into the field and, as such, do not require the students to have prior knowledge of language engineering. Please take a few moments to nominate suitable bright students who may be interested in the internship.

Please post the enclosed colored flyer at a suitable bulletin board in your department. We would also appreciate it if you could keep on file the additional workshop related information enclosed herewith and make it available to students who make inquiries.

If you have any question, please contact us by phone, e-mail, or via the Internet. The contact information is on the enclosed flyer.

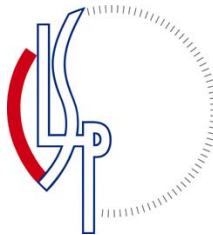
Sincerely,



Frederick Jelinek
Professor and Director

Enclosures

Center for Language and Speech Processing
JOHNS HOPKINS UNIVERSITY
Whiting School of Engineering
CSEB 324, 3400 North Charles Street
Baltimore, MD 21218-2686
(410) 516-4237 / FAX (410) 516-5050
e-mail: clsp@clsp.jhu.edu



NSF Supported Summer Internships

A Language Engineering Workshop for Students and Professionals: Integrating Research and Education

APPLICATION DEADLINE – March 20, 2009

The Center for Language and Speech Processing at the Johns Hopkins University is seeking outstanding members of the current junior class to participate in a summer workshop on language engineering from June 8 to July 31, 2009.

No limitation is placed on the undergraduate major. Only enthusiasm for research, relevant skills, past academic and employment record, and the strength of letters of recommendation will be considered¹. Students of Biomedical Engineering, Computer Science, Cognitive Science, Electrical Engineering, Linguistics, Mathematics, Physics, Psychology, etc. may apply. *Women and minorities are encouraged to apply.*²

- An opportunity to explore an exciting new area of research.
- A two-week tutorial on speech and language technology.
- Mentoring by an experienced researcher.
- Use of a computer workstation throughout the workshop.
- A \$5,000 stipend and \$2,520 towards per diem expenses.
- Private furnished accommodation for the duration of the workshop.
- Travel expenses to and from the workshop venue.
- Participation in project planning activities.

The eight-week workshop provides a vigorously stimulating and enriching intellectual environment and we hope it will encourage students to eventually pursue graduate study in the field of human language technologies.

Application forms are available via the Internet and will only be accepted electronically. **Applications must be received at CLSP by Friday, March 20, 2009***. For details, contact CLSP, 3400 N. Charles Street, 324 CSEB, Baltimore, MD 21218, visit our website - <http://www.clsp.jhu.edu>, or call 410-516-4237.

***Applications from non-U.S. citizens currently residing outside the United States must be received no later than March 6, 2009.**

¹ Four to eight undergraduate students will be selected for next summer's workshop. It is expected that they will be members of the current junior class. Applicants must be proficient in computer usage, including either C, C++, Perl or Python programming and have exposure to basic probability or statistics. Knowledge of the following will be considered, but is not a prerequisite: Linguistics, Speech Communication, Natural Language Processing, Cognitive Science, Machine Learning, Digital Signal Processing, Signals and Systems, Linear Algebra, Data Structures, Foreign Languages, or MatLab or similar software.

²The Johns Hopkins University does not discriminate on the basis of gender, marital status, pregnancy, race, color, ethnicity, national origin, age, disability, religion, sexual orientation, or veteran status in any student program or activity administered by the University or with regard to admission or employment. Questions regarding Title VI, Title IX and Section 504 should be referred to the Office of Institutional Equity, Garland Hall, Suite 130, Homewood Campus, 410-516-8075; TTY 410-516-6225.

Project Descriptions for the NSF Supported Summer Internships

1. Parsing the Web: Large-Scale Syntactic Processing

Project Leader: **Dr. Stephen Clark** (Cambridge University, UK)

Want to build the fastest linguistically-motivated parser in the world? Want to process billions of words of text, and solve fundamental language processing tasks? Interested in large-scale and distributed computing? If you've answered yes to most or all of these questions then this is the project for you! Check out the parser and try an online demo here: <http://svn.ask.it.usyd.edu.au/trac/candc/wiki>. We plan to adapt this parser to web data and use it to analyze Wikipedia, in the process producing the parser of choice for anyone needing a syntactic analysis of text. Who might want to use such a parser? A good example is any search engine company interested in "Semantic Search", for example Microsoft and Powerset.

2. Low Development Cost, High Quality Speech Recognition for New Languages and Domains

Project Leader: **Dr. Dan Povey** (Microsoft Research, U.S.A.)

This project involves applying newly developed techniques based on factor analysis to speech recognition. These techniques are derived from speaker identification technology, and are based on expressing the models for a particular speaker and speech sound as a linear combination of factors specific to the speaker and the speech sound. The framework allows us to specify the speech models much more compactly. We intend to apply it to recognizing speech from languages where the amount of transcribed speech data is relatively small. Based on very positive initial results, we anticipate that this project will have a big impact on the speech recognition community. Relevant skills include linear algebra, C/C++, Unix shell, and statistical modeling. Already committed team members include researchers working in the USA, UK, Czech Republic and Canada.

3. Unsupervised Acquisition of Lexical Knowledge from N-Grams

Project Leader: **Dr. Dekang Lin** (Google Research, U.S.A.)

The overall performance of machine-learned NLP systems is often ultimately determined by the size of the training data rather than the learning algorithms themselves. The web undoubtedly offers the largest textual data set. Previous researches that use the web as the corpus have mostly relied on search engines to obtain the frequency counts and/or contexts of given phrases. Unfortunately, this is hopelessly inefficient when building large-scale lexical resources.

We propose to build a system for acquiring lexical knowledge from ngram counts of the web data. Since multiple occurrences of the same string are collapsed to a single one, the ngram data is considerably smaller than the original text. Since most lexical learning algorithms only collect data from small windows of text anyway, the ngram data can provide the necessary statistics needed for the learning tasks in a much more compact and efficient fashion. The students participating in the workshop will gain experience in advanced development of search and parallel processing and develop skills for data analysis and exploration.

Center for Language & Speech Processing www.clsp.jhu.edu Johns Hopkins University, Baltimore, MD