NSF Supported Summer Internships

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

APPLICATION DEADLINE – March 19, 2010

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 7 to July 30, 2010.

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

- 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 19, 2010*. For details, contact CLSP, 3400 N. Charles Street, 320 Barton Hall, Baltimore, MD 21218, visit our website - http://www.clsp.jhu.edu, or call 410-516-4237.
Speech Recognition with Segmental Conditional Random Fields

This project will explore an exciting new method for doing speech recognition. Whereas conventional approaches to speech recognition analyze speech in tiny, fixed-length blocks, the proposed Segmental Conditional Random Field (SCRF) approach analyzes it in variable length segments corresponding directly to words. In this approach, we will extract numerous features, each of which measures some aspect of the consistency between the speech segment and the hypothesized word. These features will be combined in a log-linear model, which will allow for the joint training of both acoustic and language modeling features. SCRFs have the potential to make a fundamental impact on the way we do speech recognition, and advances we make in graphical models will have broad relevance to the fields of text and image processing.

Localizing Objects and Actions in Videos with the Help of Accompanying Text

Multimedia content is a growing focus of search and retrieval, personalization, categorization, and information extraction. Video analysis allows us to find both objects and actions in video, but recognition of a large variety of categories is very challenging. Any text accompanying the video, however, can be very good at describing objects and actions at a semantic level, and often outlines the salient information present in the video. Such textual descriptions are often available as closed captions, transcripts or program notes. In this inter-disciplinary project, we will combine natural language processing, computer vision and machine learning to investigate how the semantic information contained in textual sources can be leveraged to improve the detection of objects and complex actions in video. We will parse the text to obtain verb-object dependencies, use lexical knowledge-bases to identify words that describe these objects and actions, use web-wide image databases to get exemplars of the objects and actions, and build models that can detect where in the video the objects and actions are localized.

Synchronous Grammar Induction for Statistical Machine Translation

The last decade of research in Statistical Machine Translation (SMT) has seen rapid progress. The most successful methods have been based on synchronous context free grammars (SCFGs), which encode translational equivalences and license reordering between tokens in the source and target languages. Yet, while closely related language pairs can be translated with a high degree of precision now, the result for distant pairs is far from acceptable. In theory, however, the "right" SCFG is capable of handling most, if not all, structurally divergent language pairs. So we propose to focus on the crucial practical aspects of acquiring such SCFGs from bilingual text. We will take the pragmatic approach of starting with existing algorithms for inducing unlabelled SCFGs (e.g. the popular Hiero model), and then using state-of-the-art hierarchical non-parametric Bayesian methods to iteratively refine the syntactic constituents used in the translation rules of the grammar, hoping to approach, in an unsupervised manner, SCFGs learned from massive quantities of manually "tree-banked" parallel text.