

Final Project

L445/L515

Autumn 2008

Who this is for

If you are taking this course as L515, this is a mandatory project. If you are taking this as L445, you can elect to do a project, for the purposes of obtaining extra credit, namely one homework's worth of credit (a significant 6%).

1 Project overview

Your project is to delve more deeply into a topic we have not covered in class. That is, your task is to identify an area of computational linguistics where interesting work is being done; describe such work; and then identify a part of that area where more work needs to be done. Most of your project will consist of summarizing what has already been done in the literature, but I want you also to be thinking of what could be done in the future.

Getting started I want you to find a topic in the following manner:

- Go to the Association for Computational Linguistics (ACL) anthology (<http://aclweb.org/anthology-new/>), which contains many papers in the field from the last 30 years.
- Browse paper titles from any of the workshops, conferences, or journals from this decade (i.e., click only on years starting with '0').
 - The workshops might be easiest to start with, since they give clearer categorizations as to what the paper is about.
 - Feel free to ask others (me, more advanced grad students, etc.) for suggestions.
- Select a paper or two which sounds interesting (even if you don't understand the title completely).
- Read the paper, understanding that you won't understand most of it.

- Find a part of it which: a) you at least partially understand, b) is interesting to you, and c) you see how it relates to some component of the course or some topic of study in linguistics you're interested in
- Follow up on that topic by looking at the paper's references, other papers by the same author(s), searching for that topic, etc.

2 Project details

Your project will basically consist of a paper approximately 10 pages in length, single-spaced.

The paper should roughly be organized as follows:

1. The general issue (e.g., word sense disambiguation (WSD)), including a discussion of why this issue is important. Include pointers to relevant literature.
2. The specific issue (e.g., WSD in Arabic), including a lengthy discussion showing that you understand how previous approaches worked, why those approaches were used, what needs to be improved, etc.
3. An open question in that specific issue and your brief idea for handling that question.

You may not be able to state precisely how to tackle the problem at this point, but in general terms at least state what could/should be tried to improve the state of the art.

3 Timeline

- **Wednesday, November 5, 2008** (9:30am): topic selected, including preliminary references
 - You will turn in a one-paragraph description (no more, no less), with accompanying references.
 - At any time between now and then, you are encouraged to come talk to me about your topic.
- **Wednesday, November 19, 2008** (9:30am): outline of paper due
 - This should give clear indication that you understand the essentials of the topic and how to form them into a paper. (If I had to guess, I'd say this would take you 1-2 pages.)
- **Wednesday, December 17, 2008** (5:00pm): final paper (approx. 10 pages) due