Syllabus for LING 575
Unsupervised learning: a case study on unsupervised POS tagging
Winter 2008

Class time & location: T 3:30-5:50pm, MGH 284

Instructor: Fei Xia
Office: Padelford A-210G
Office Hours: Fr: 10:30-11:30am
Phone: (206) 543-9764
Email: fxia at u.washington.edu (include “Ling575” in the subject line).

Website: http://courses.washington.edu/ling575x

Course Description:
The existing work on unsupervised POS tagging can be divided into three categories: the first starts with the forward-backward algorithm and improves performance by using a “filtered” lexicon; the second clusters words and tags the words with the cluster labels; the third takes advantage of bilingual data by projecting POS information from one language to the other. In this course, we will discuss each approach and build systems aiming at improving the state of the art.

Required Textbooks:
None. Reading materials are online.

Some background reading can be found in
“Foundations of statistical natural language processing” (M&S) by Manning and Schütze and/or
Speech and Language Processing (J&M) by Jurafsky and Martin (the Table of Content is available at http://www.cs.colorado.edu/~martin/slp2.html#top)

Prerequisites:
• CS 326 (Data Structures) or equivalent
• Stat 391 (Prob. and Stats for CS) or equivalent
• Programming in Perl, C, C++, Java, or Python
• LING 570
Grading:
- Assignments (45-60%): there will be four assignments plus a final report. All the assignments are due at 2:30pm on Tuesday.
- Presentation (15-25%): each student will present one or two papers and give a final presentation.
- Reading (5-10%): it will be checked but not graded.
- Class participation (10-20%)

Course policies:
- **Course website**: All course information on this web page is tentative and can change at any time. Confirm crucial dates or information with me in person during class.

- **Attendance**: Students are expected to attend all classes: announcements about assigned readings and assignments will be made available at the start of each class. Such announcements may not be made on this web page, so don't rely on information here instead of attending class.

- **Late assignment submission**: There will be a 1% penalty for every hour after the deadline. For instance, suppose the assignment is due at 2:30pm and you turn in the assignment at 2:30am the next morning, you grade would be x * 0.88, where x is the grade you would get if you have turned in before the 1pm deadline. **No assignments will be accepted four (4) days after the due date.**

- **Reading**: You need to read papers before class starts. Write your questions on a piece of paper and bring it to class.

- **Incomplete**: According to UW policy, "incomplete grades may only be awarded if you are doing satisfactory work up until the last two weeks of the quarter." Therefore, it is crucial for you to hand in your homework on time. An “incomplete” grade is given only under extremely unusual circumstances (e.g., health issues, family emergency).

- **Emails**: Use the prefix "ling575: " on all the subject line of your messages. If you do not include the prefix, then the mail might go unanswered. If you don’t receive a reply from me within 48 hours, please send me a reminder.
Tentative schedule (subject to changes):

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Reading</th>
<th>Hw</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/8</td>
<td>Introduction Recap on HMM: by Will</td>
<td>M&amp;S 9.3</td>
<td>Hw1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1/15</td>
<td>EM approach (I): by Fei • Forward-background algorithm</td>
<td>M&amp;S 9.3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1/22</td>
<td>EM approach (II) by Fei • Will EM help?</td>
<td>(Merialdo, 1994) (Elworthy, 1994)</td>
<td>Hw2</td>
<td>Hw1</td>
</tr>
<tr>
<td>4</td>
<td>1/29</td>
<td>EM approach (III): by Alex • Using lexicon (w/ or w/o filtering) • Constraint EM</td>
<td>(Banko and Moore, 2004) (Wang and Schuurnans, 2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2/5</td>
<td>Clustering approach: by Fei</td>
<td>(Schutze, 1995) (Clark, 2000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2/12</td>
<td>Prototype learning: by Ryan</td>
<td>(Haghighi and Klein, 2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2/19</td>
<td>Cross-lingual approach (I): by Ankit</td>
<td>(Yarowsky and Ngai, 2001) (Cucerzan and Yarowsky, 2002)</td>
<td></td>
<td>Hw2</td>
</tr>
<tr>
<td>8</td>
<td>2/26</td>
<td>Bayesian approach</td>
<td>(Blei et al.,)</td>
<td></td>
<td>Hw3</td>
</tr>
</tbody>
</table>
Reading material: (*: papers discussed in class)

- EM approach:

- Clustering approach:
  
  
  
  

- Bayesian approach:
  
  
  
  - (Goldwater and Griffiths, 2007)* A fully Bayesian approach to unsupervised POS tagging. ACL 2007.

- Cross-lingual approach:
  


• Prototype learning: