11-719 Computational Models of Discourse Analysis

Instructor
Dr. Carolyn P. Rosé (cprose@cs.cmu.edu)

Office hours: Students are encouraged to request meetings with the instructor as needed

Units: 12 (PhD/Master’s)

Readings and On-Line Discussions: All required readings will be posted as pdfs to blackboard. The following book is recommended but not required for the course:


Students are expected to do the readings and post a response to discussion questions on-line to the weekly blogs in Blackboard by 8pm the night prior to each class meeting.

Prerequisites: students should be reasonably strong programmers and have taken or audited at least one machine learning course

Class Meets: Monday and Wednesday, 10:30am-11:50am, Gates Hillman Center 4102

Course Description

Discourse analysis is the area of linguistics that focuses on the structure of language above the clause level. It is interesting both in the complexity of structures that operate at that level and in the insights it offers about how personality, relationships, and community identification are revealed through patterns of language use. A resurgence of interest in topics related to modeling language at the discourse level is in evidence at recent language technologies conferences. This course is designed to help students get up to speed with foundational linguistic work in the area of discourse analysis, and to use these concepts to challenge the state-of-the-art in language technologies for problems that have a strong connection with those concepts, such as dialogue act tagging, sentiment analysis, and bias detection.

This is meant to be a hands on and intensely interactive course with a heavy computational modeling component. The course is structured around 3 week units, each of which have a hands on assignment.

Course Procedures and Grading Criteria

Most assignments will involve hands on computational work. You are welcome to use any tools you know and have access to. One way of approaching assignments with computational components is to program plugins for the LightSIDE text mining tool bench (http://ankara.lti.cs.cmu.edu/side/download.html). Plugins with either be novel feature extractors, classification algorithms, or meta-classifiers.

Grades will be assigned as follows:
20% for each of four assignments  
10% for class participation  
10% final project write up (Due by May 11)

**Project/Assignments**

Each of the 4 assignments contributes directly to the course project, which will be a team project. The project should involve 3 (or at most 4) students each.

1. Preliminary: Pick one of the 4 data sets by January 25  
   a. *Example Data Sets*: Street Gangs Forum, Coursera MOOC discussions, Medical Support Forum, Wikipedia Talk Pages, GitHub Issue Corpus
2. Assignment 1 Part A: Model interpretation, due before class on February 1  
3. Assignment 1 Part B: Analysis by hand, revisit model interpretation, due before class on February 10  
4. Assignment 2: Preliminary analysis of project data set, due before class on February 29  
5. Assignment 3: Project Proposal with preliminary results, due before class on March 28  
6. Assignment 4: Final analysis, schedule meeting between team and instructor for the week of May 2  
7. Final paper due May 11

**Semester Reading Schedule and Deadlines**

**UNIT 1: Foundations: Analysis of Conversational Strategy (Jan 11, 2016)**

**Week 1 Lecture 1 (Jan 11): Course Intro**

**Week 1 Lecture 2 (Jan 13): Defining Social Goods and Achievement of Social Goals**


**Week 2 Lectures 2 (Jan 20): Modeling Achievement of Social Goals**

- Caroline Langlet Chloe Clavel (2015). Improving social relationships in face-to-face human-agent interactions: when the agent wants to know user’s likes and dislikes. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Beijing, China

**Week 3 Lecture 1 (Jan 25): Simple Graphical Models of Conversational Structure**

- Review LDA if you are not familiar with it:  

• DEADLINE: Project dataset Sign-Up (using Google Form sent by email)

Week 3 Lecture 2 (Jan 27): Studying Discussion Practices in Social Media


UNIT 2: Power, Politeness and Intimacy (Feb 1, 2016)

Week 4 Lecture 1 (Feb 1): Power, Gender and Style in Blogs


• DEADLINE: Assignment 1 Part A

Week 4 Lecture 2 (Feb 3): Theoretical Framing of Politeness Theory


Week 5 Lecture 1 (Feb 8): Modeling Politeness


Week 5 Lecture 2 (Feb 10 – LP Guest Lecture): Emotion and Intimacy

Reading TBA

• DEADLINE: Assignment 1 Part B

Week 6 Lecture 1 (Feb 15): Modeling Intimacy and Self-Disclosure

Week 6 Lecture 2 (Feb 17): Power Differences at Work and in the Courtroom


UNIT 3: Perspective and Meaning Making (Feb 22)

Week 7 Lecture 1 (Feb 22)


Week 7 Lecture 2 (Feb 24): Accommodation and Transactivity: From Theory to Computation


Week 8 Lecture 1 (Feb 29): Modeling Argumentation and Perspective


- **DEADLINE: Assignment 2**

Week 8 Lecture 2 (Mar 2): Metaphor and Meaning Making


SPRING BREAK!!!

Week 9 Lecture 1 (Mar 14): Similarity and Difference


**Week 9 Lecture 2 (Mar 16): Metaphor and Conversation**


**Week 10: Project work week, full proposal due following Monday**

**UNIT 4: Legitimacy and Identity: Interactions “on the Street” (Mar 25, 2015)**

**Week 11 Lecture 1 (Mar 28): Sociolinguistic Perspective on Style: What Linguistic Resources Buy You**


- DEADLINE: Proposal and preliminary results

**Week 11 Lecture 2 (Mar 30) Negotiating Legitimacy in Interaction**


**Week 12 Lecture 1 (Apr 4) Computational Analysis of Gang Language**


Week 12 Lecture 2 (Apr 6) Impersonation and Multiple Identities


UNIT 5: Social Positioning through Narrative (April 11, 2015)

Week 13 Lecture 1 (Apr 11) Positioning events and characters

- Page, R. (2012). *Stories and Social Media: Identities and Interaction*, Routledge (Chapter 4)


Week 13 Lecture 2 (Apr 13) Positioning the Author among Colleagues


Week 14 Lecture 2 (Apr 18): Analysis of Point of View


Week 14 Lecture 2 (Apr 20) Positioning the Narrator


Week 15 Lecture 1 Reflecting situational factors through personal expression

- Miaomiao Wen, Zeyu Zheng, Hyeju Jang, Guang Xiang, and Carolyn Rosé. Extracting Events with Informal Temporal References in Personal Histories in Online Communities. ACL’13, 2013


Week 15 Lecture 2: Group work day, work on final project
DEADLINE: Assignment 4. Each team must set up an in person meeting with the instructor during the week of May 2 to report on in progress results for the final project and get feedback. Final Paper due on May 11, 2016