Syllabus for LIN 650

Computational Morphology – Fall 2018

August 18, 2018

Instructor	Course
Jeff Heinz	LIN 650
jeffrey.heinz@stonybrook.edu	Computational Morphology
N237 SBS	MW 14:30-15:50
L160 IACS	N117 SBS
Office Hours	
TTh 2:30-4:00pm and by appointment	

https://jeffreyheinz.net/classes/18F/

What is this course? Morphology is the study of words, word formation, and word analysis. This course is about *computational* morphology, which is the study of these topics from a computational perspective.

One of the most influential approaches today, and the one that drives much of the software on laptops and smartphones, is based on finite-state technology. Much of the course is dedicated to understanding how this technology works, to understanding how it relates to linguistic theories of morphology, and to becoming a proficient user of this technology for morphological analysis.

Course Objectives. The goal of this course is to become familiar with both the *theory* and the *practice* of computational methods of morphological analysis and recognition. In particular we study weighted finite-state transducers and the *composition* operation, which generalizes many morphological processes including concatenation, truncation, root-and-pattern morphology and others. We pay special attention to irregular morphology and exceptions. Students will learn to use the grammar libraries **Thrax** and **Pynini** that are part of the **openFST** toolkit (http://www.openfst.org/twiki/bin/view/FST/WebHome).

In addition to these core topics, depending on the students' interests, additional topics covered include higher-level logical and DATR-like descriptions of morphology and the machine learning of morphology.

Course Materials

- Brian Roark and Richard Sproat. *Computational Approaches to Morphology and Syntax*. Oxford University Press, Oxford, 2007.
- http://www.openfst.org/twiki/bin/view/GRM/Pynini
- http://www.openfst.org/twiki/bin/view/GRM/Thrax

Grading. There are three components to final grades: Assignments (30%), Presentations (20%) and a Research Project (50%).

Assignments. Students enrolled for n credits are required to complete $\lfloor \frac{A \times n}{3} \rfloor$ written and/or coding assignments during the semester where A is the total number of assignments (I am aiming for A=9).

Presentation. Students enrolled for n credits are required to lead the class $n - 1$ times during the semester. Students leading the class may lead discussion of a research article or project in computational morphology, teach the class some software development for morphological analysis related to openFST, or present something else related to their own research. While classes may be led lecture style, it will probably be more fun and a better learning experience if class leaders encourage discussion and interaction.

Research Project. Students enrolled for 3 credits are required to complete a research project. Projects can be theoretical or applied.

Students completing theoretical projects will turn in a written paper, like a short journal article. Theoretical projects should propose new or improved methods for conducting computational morphology.

Students completing applied projects will turn in one or more files that constitute a program implementing a morphological recognizer/analyzer for some non-trivial aspect of the morphology of a language. The program will be written with the tools we learn to use in class. Students completing applied projects will also turn in a 1-2 page report which evaluates the implementation: What is satisfactory and what is unsatisfactory about it?

Projects are due by 5pm, Friday December 14, 2018.

University Policies and Services

Student Accessibility Support Center (SASC) Statement: If you have a physical, psychological, medical or learning disability that may impact your course work, please contact the Student Accessibility Support Center (SASC), ECC (Educational Communications Center) Building, room 128, (631) 632-6748. They will determine with you what accommodations, if any, are necessary and appropriate. All information and documentation is confidential.

Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and the staff at the Student Accessibility Support Center (SASC). For procedures and information go to the following website: `http://www.stonybrook.edu/ehs/fire/disabilities`

Academic Integrity Statement: Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental

Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at `http://www.stonybrook.edu/commcms/academic_integrity/index.html`

Critical Incident Management Statement: Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Judicial Affairs any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures.