Syllabus (handleiding) Natural Language Models and Interfaces 2012

Lecturer: Jelle Zuidema, email: none (email teaching assistants instead)
TA1: Frank Smit (F.Smit@uva.nl)
TA2: Maarten van de Velden (maarten.vdvelden@gmail.com)

Lectures - part 1
1. Intro Language, NLP, Shannon, ngrams
2. Generative models: from ngrams to PCFGs, for modelling, disambiguation, transduction
(probably no class on 13/4: VIA Congres)
3. Smoothing
4. Noisy-channel model, spelling correction, deciphering - *sample exam 1 distributed*
5. Dynamic programming: Earley, CYK, Viterbi

Computer Labs - part 1
(*You may work in groups of 2 on the programming assignments. There's a new assignment every lab session; show the assistants that you have completed the assignment at the end of the session or start of the next; you must choose a different partner in lab 5-9 than you had in lab 1-4; you hand in individual papers at the end of part 1 and 2*).
1. Word frequency distributions, Zipf, ngram frequencies (3pts) - *sign up for presentations*
2. Language modelling: bigram and trigram models without unknown words (3pts)
3. Language modelling: bigram and trigram models with unknown words (3pts)
4. Language modelling: selecting optimal path through word graphs
5. Deciphering - *change programming partner*

Evaluation - part 1
Paper on word graphs (due 23/4, 12h00) (3pts if you submit a complete paper on time; contents are only graded after part 2)
Exam 1: Theory (27/4) (max. 25 pts)

Lectures - part 2
6. Syntax & Compositional semantics: lambda calculus
7. Wide-coverage semantic parsing
8. Lexical semantics
9. Discourse
10. Capita selecta - *sample exam 2 distributed*

Computer & practical labs - part 2
6. Deciphering - *ctd*
7. Lambda calculus exercises (3pts)
8. Boxer (3pts)
9. C&C parser (3pts)
10. Extra computer lab
11. Extra tutorial session

Evaluation - part 2
Paper on word graphs and deciphering (due 25/5; max. 24 pts)
Exam 2: Theory (1/6; max 25 pts)

Student presentations (max 5pts): *In groups of maximum 7 students you present once during the course your solutions to the programming assignments, and post model code (except for the 'deciphering' assignment) to the Blackboard site. Presentations are brief (15 minutes maximum) and scheduled at the start of lecture 2 (ngram frequencies), 4 (language modelling with unknown words), 6 (word graphs), 8 (deciphering) and 10 (boxer / C&C parser). Sign up by following the link on Blackboard.*

Papers: The programming assignments are evaluated in the final paper, handed in at the end of part 2. A preliminary version of this paper is handed in, anonymously, at the end of part 1 and commented on by a fellow students and further discussed in the Tutoraat meeting at the beginning of part 2.

Students that have already passed Natural Language Interfaces may skip exam 2, the practical sessions of part 2, but do hand in the papers at

the end of part 1 and 2 and participate in the student presentations.

Students that have already passed Language Models may skip exam 1, the practical sessions of part 1, and may replace the papers by an individual paper on computational semantics. Talk to the lecturer after class to discuss a good topic and explain your circumstances in an appendix to the final paper you hand in on 29/5.

Missed classes and assignments: If you miss a class or an assignment, don't email me. If you missed some of the points awarded that day, and you have a good reason (illness, funeral etc), explain your circumstances in an appendix to the final paper you hand in on 25/5.

Timetable

Week		Tuesday	Wednesday		Friday
14 2/4 6/4			11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111		Goede Vrijdag Feestdag
15 9/4 13/4			11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111		11-13: Hoorcollege ScP G4.15 13-15: Computerpracticum ScP G0.23-G0.25
16 16/4 20/4		11-13: Hoorcollege ScP D1.112 13-15: Computerpracticum ScP F2.04	11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111		11-13: Hoorcollege ScP G4.15 13-15: Computerpracticum ScP G0.23-G0.25
17 23/4 27/4					13-15: Deeltoets ScP C0.110

18		Dag van de Arbeid Feestdag		
19 7/5 11/5			11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111	11-13: Hoorcollege ScP G4.15 13-15: Werkcollege ScP A1.04
20 14/5 18/5		13-15: Hoorcollege ScP D1.112 15-17: Computerpracticum ScP D1.111	11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111	Dag na Hemelvaartsdag Feestdag
21 21/5 25/5			11-13: Hoorcollege ScP D1.116 13-15: Computerpracticum ScP D1.111	11-13: Hoorcollege ScP G4.15 13-15: Werkcollege ScP A1.04
22 28/5 1/6				✉ 13-15: Deeltoets ScP C1.110