

Frank Richter:
Introduction to Computational Linguistics

Seminar:	Wednesday 14ct–16 at the Sfs, Hörsaal 0.02		
	Exception: Oct. 26th moves to Friday, Oct. 28th, 16ct–18h		
Regular seminar starts:	Wednesday, October 19th, 2011		
Credits:	3 CP		
Office Hours:	Monday 12.00 – 13.00		
	Office No. at Sfs	Phone #	Email
	2.17	29-78489	fr@sfs.uni-tuebingen.de

- Webpages: www.sfs.uni-tuebingen.de/~fr/teaching/ws11-12/i2c1/
There is a Moodle page with reading material (linked from that page)
- Midterm exam: November 30th
- Final exam: in the last week of classes, on February 1st

Reading Assignments

Please read the assigned reading in advance of the class meeting for which it was assigned. I will presuppose that you have read the material when we discuss it in class.

Grading Policy

Your grade will be based on three components: exam at the end of the compact week with Jochen Saile (20 %), midterm exam (40 %) and final exam (40 %).

General Remark

If you carry your mobile phone with you, please turn it off before class.

Course Objective This introductory course has five major goals:

- (Largely non-technical) introduction to the field of computational linguistics and its history.
- Survey of natural language processing applications.
- In-depth look at machine translation as a means to illustrate the major tasks for natural language processing
- Presentation of tools and resources needed for natural language processing applications.
- To give you credit for your work and to get you one step closer to your degree.

Time Table

- 19.10.: Organizational matters, Introduction
- 28.10.: Overview and History of Computational Linguistics
- 02.11.: Machine Translation I
- 09.11.: Machine Translation II
- 16.11.: Machine Translation III
- 23.11.: Tokenization and Sentence Segmentation I
- 30.11.: Midterm exam
- 07.12.: Tokenization and Sentence Segmentation II
- 14.12.: Regular Expressions, Finite State Automata
- 21.12.: Finite State Transducers
- 11.01.: Finite State Transducers and Replacement Operators
- 18.01.: Finite State Transducers, Morphological Analysis
- 25.01.: Part of Speech Tagging and Course Review
- 01.02.: Final exam

Course Readings

- Arnold, D., Balkan, R., Humphreys, R. L., Meijer, S., and Sadler, L. (eds) 1994. *Machine Translation. An Introductory Guide*. Manchester/Oxford: NCC Balckwell Ltd.
- Dale, R., Moisl, H., and Somers, H. (eds) 2000. *Handbook of Natural Language Processing*. New York, Basel: Marcel Dekker, Inc.
- Garside, R., Leech, G., and A., McEnery (eds) 1997. *Corpus Annotation. Linguistic Information from Computer Text Corpora*. Addison Wesley Longman Ltd.
- Hutchins, John 2003. Commercial systems. The state of the art. In Harold Somers (ed), *Computers and Translation. A Translator's Guide*, 161–174. John Benjamins. Amsterdam.
- Hutchins, W. J. and Somers, H. L. (eds) 1992. *An Introduction to Machine Translation*. San Diego: Academic Press.
- Karttunen, L. 2003. Finite state technology. In Mitkov [Mitkov, 2003], 339–358.
- Leech, G. 1997. Grammatical tagging. In Garside et al. [Garside et al., 1997], 19–33.
- Locke, W. N. and Booth, A. D. (eds) 1955. *Machine Translation of Languages*. Cambridge, Mass.: MIT Press.
- Mitkov, R. (ed) 2003. *The Oxford Handbook of Computational Linguistics*. Oxford University Press.
- Palmer, D. D. 2000. Tokenization and sentence segmentation. In Dale et al. [Dale et al., 2000], 11–36.

Searle, J. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–458.

Slocum, J. (ed) 1988. *Machine Translation Systems*. Cambridge University Press.

Somers, H. 2000. Machine translation. In Dale et al. [Dale et al., 2000], 329–346.

Trost, H. 2003. Morphology. In Mitkov [Mitkov, 2003], 25–47.

Turing, A. M. 1950. Computing machinery and intelligence. *Mind*, 59.