Language Technology: Research and Development

Introduction

Sara Stymne

Uppsala University
Department of Linguistics and Philology
sara.stymne@lingfil.uu.se

Teaching Team

- Course coordinator, examiner and lectures:
  - Sara Stymne
- Seminars
  - Beáta Megyesi
  - Paola Merlo
  - Sara Stymne
- Assistant (popular science abstracts)
  - Samuel Douglas
- Alumni guest lecturers

Course Content

Theory	**Practice**
Philosophy of science	Survey a research field
Research methods in LT	Plan and implement a project
Scientific writing	Write and review scientific papers

- Lectures covering theory (large group)
- Seminars devoted to practice (small group)
- Individual projects on a common theme (small group)

Research Themes

- Digging the Past: Digital Philology and the Analysis of Historical Sources [Bea]
- Beyond the Benchmarks: Linguistically-oriented analysis and generalisations in Neural Networks [Paola]
- Cross-lingual NLP [Sara]

Research Themes

- Digging the Past: Digital Philology and the Analysis of Historical Sources [Bea]
  - Campus seminars
- Beyond the Benchmarks: Linguistically-oriented analysis and generalisations in Neural Networks [Paola]
  - Zoom seminars
- Cross-lingual NLP [Sara]
  - Campus seminars

Course Structure

1. Background part:
   - Philosophy of science and research methods [lectures]
   - Survey of the state of the art in research theme [seminars]
   - Planning an R&D project [lecture, seminar]
2. Project part:
   - Implementing an R&D project [seminars]
   - Writing a scientific paper [lecture, seminars]
   - Reviewing scientific papers [lecture]

Reading List

- Textbooks:
  - Okasha, S. (2002) *Philosophy of Science: A Very Short Introduction*. Oxford University Press.
  - Zobel , J. (2004) *Writing for Computer Science*. Second Edition. Springer.
- Papers:
  - Available online from the course home page https://cl.lingfil.uu.se/kurs/rd21/

Assignments and Examination

1. Take home exam on philosophy of science (15%) [written]
2. Research paper presentation and discussion (15%) [oral]
3. Project proposal (15%) [written, oral]
4. Term paper (40%) [written, oral]
5. Review of term papers (15%) [written]

Assignments and Examination

1. Take home exam on philosophy of science (15%) [written]
2. Research paper presentation and discussion (15%) [oral]
3. Project proposal (15%) [written, oral]
4. Term paper (40%) [written, oral]
5. Review of term papers (15%) [written]

- Pass (G) = all assignments passed
- Distinction (VG) = at least 50% of 1, 3–5 with distinction

Deadlines

Choose your preferred topic	September 3, 13.00
Hand in take home exam	September 16
Project proposal	October 8
Present project proposal	October 13
First version of term paper	December 13
Peer review of (other) term papers	December 22
Final seminar	January 13
Final term paper	January 14

Deadlines

Choose your preferred topic	September 3, 13.00
Hand in take home exam	September 16
Project proposal	October 8
Present project proposal	October 13
First version of term paper	December 13
Peer review of (other) term papers	December 22
Final seminar	January 13
Final term paper	January 14

Backup deadlines available on course web page, but important to try to respect original deadlines! (This course is a prerequisite for the master thesis course.)

Seminars

- All seminars are obligatory!
- Group seminars:
  - Research papers
  - Project proposal (presentations with slides)
  - Progress reports (and ethics)
- Final seminar in full group
  - Full day "mini workshop"
  - First-year master students also invited
  - Social event (if possible)

Seminars

- All seminars are obligatory!
- Group seminars:
  - Research papers
  - Project proposal (presentations with slides)
  - Progress reports (and ethics)
- Final seminar in full group
  - Full day "mini workshop"
  - First-year master students also invited
  - Social event (if possible)
- If you miss a seminar, there will be a compensation task

Going for the Real Thing

- The goal is to do real research resulting in real publications
- Guidelines for submission and reviews:
  - Transactions of the Association for Computational Linguistics
    http://www.transacl.org/submission/
- Term papers may be revised and submitted for publication
- Actual submission is not a course requirement

Going for the Real Thing

- The goal is to do real research resulting in real publications
- Guidelines for submission and reviews:
  - Transactions of the Association for Computational Linguistics
    http://www.transacl.org/submission/
- Term papers may be revised and submitted for publication
- Actual submission is not a course requirement

- You are meant to function as a real research group
- Projects are individual, but you should support each other

Publications from Recent Years

Antonia Karamolegkou and Sara Stymne. *Investigation of Transfer Languages for Parsing Latin: Italic Branch vs. Hellenic Branch*. NoDaLiDa 2021.

Harm Lameris and Sara Stymne. *Whit's the Richt Pairt o Speech: PoS tagging for Scots*. VarDial 2021.

Sebastian Reimann and Daniel Dakota. *Examining the Effects of Preprocessing on the Detection of Offensive Language in German Tweets*. KONVENS 2021.

Huiling You, **Xingran Zhu**, and Sara Stymne. *Uppsala NLP at SemEval-2021 Task 2: Multilingual Language Models for Fine-tuning and Feature Extraction in Word-in-Context Disambiguation*. SemEval-2021.

Marsida Toska, Joakim Nivre, Daniel Zeman. *Universal Dependencies for Albanian*. UD workshop 2020.

Arra'Di Nur Rizal and Sara Stymne. *Evaluating Word Embeddings for Indonesian–English Code-Mixed Text Based on Synthetic Data*. Workshop on Computational Approaches to Code Switching 2019.

Learning Outcomes 1

The student should at least be able to do the following, in relation to a scientifically organised language technology project:

- explain the basic principles of scientific work and research methodology in general and in relation to a current project
- make an overview of earlier research and the state of the art within the field that the project treats and identify its most urgent research issues,
- show an ability to identify and formulate research questions in a critical, independent, and creative way

Learning Outcomes 2

The student should at least be able to do the following, in relation to a scientifically organised language technology project:

- plan and carry out research tasks based on sound methodological principles and within given time limits,
- evaluate results and partial results with current validation methods,
- present the purpose of the project and its results in a professional manner, both for scientists and for the general public, orally and in writing, taking the target audience into consideration.

Student Feedback

- 2020 students were very happy with the course: (4.5/5)
- Some comments:
  - Seminars were very useful
  - Philosophy of science part was good
  - I liked the division into research groups with regular seminars
  - The information flow (despite online) and guidance throughout the course worked well
  - It would be good if students could express the strength of their preference for different topics.
  - More peer review during the course, and encourage student to collaborate between seminars

Campus / Zoom

- We will mainly have Campus teaching
  - Lectures will be also on Zoom (mainly)
  - Two online seminar groups, one Zoom group
- Campus activities may be cancelled on short notice

  Check your email+Studium before going to Campus!
- Please follow current regulations!
  - Do not come to Campus if you do not feel well!
  - Try to maintain social distancing

Coming up

- Now: introduction to the topics
- Wish for your preferred topics
  - Rank your preference for the 3 topics, plus indicate if your first choice is a strong preference
  - Indicate your preference for Campus/online seminars (and let us know if you have an approved reason for online teaching)
  - By email to Sara: deadline Friday Spetember 3, 13.00
- Lecture on science, research and NLP: Friday
- Lecture: debates on philosophy of science and NLP: next Tuesday
- First research paper seminars: September 10 or 14

Research Paper Seminars

- Obligatory attendance
- All students are supposed to have read all articles, to bring discussion points, and actively discuss the articles
- Each student is responsible for introducing one article each
  - briefly summarize the paper (2 min)
  - discuss the main points being made
  - bring up difficult to understand parts
  - initiate a discussion by proposing themes to discuss
- Bring the articles to the seminars (on paper or electronically)
- The list of articles and presenters will be available on the web page early next week

Questions?