Year : 2018/19

8025 - Intelligent Interactive Systems- MSc
32489 - Computational Semantics

Syllabus Information

Academic Course:	2018/19
Academic Center:	802 - Official Postgraduate Programme in Information Technology, Communication and Audiovisual Media
Study:	8025 - Intelligent Interactive Systems- MSc
Subject:	32489 - Computational Semantics
Credits:	5.0
Course:	1
Teaching languages:	Theory: Group 1: Pending
Teachers:	Carina Helga Silberer
Teaching Period:	Quarterly

Presentation

This course provides the basics of how natural language meaning is modeled in Computational Linguistics / Natural Language Processing. We will analyse the relevant semantic phenomena (focusing on word and sentence meaning) and approaches to tackling them, especially data-driven---empirical and Machine Learning---methods. We will begin with traditional approaches to computational semantics, reflect on one particular issue (word meaning) and move to recent approaches in distributional semantics and neural networks. Along the way, we will learn the basic methodology of Machine Learning and strengthen the student's skills in using computational and quantitative tools (in class, we will use Python and Python-based toolkits, such as NLTK).

Associated skills

- Analytical skills (problem solving, data analysis).
- Machine Learning methodology.
- Basic programming (Python, NLTK).

Learning outcomes

The student will acquire:

- a deeper understanding of semantics and how Computational Linguistics can contribute to its study;
- knowledge of the basic methodology of Machine Learning (with---depending on the students' interests---a focus on neural networks
- familiarity with quantitative and computational methods to approach semantic problems.

Prerequisites

- Natural Language Processing / Computational Morphosyntax course **strongly recommended**, but **not required**.
- Previous programming background **not required** (but very welcome).

Contents

1. Traditional computational semantics
    - Practical exercise: Agreeing on semantic categories and inter-annotator agreement.
2. Machine Learning for computational semantics
    - Practical exercise: Machine Learning for Word Sense Disambiguation.
3. "You shall know a word by the company it keeps!" (Firth, 1957): Distributional semantics
    - Practical exercise: Building and analysing a semantic model from text corpora.
4. The neural networks / deep learning revolution.
    - Practical exercise: Training and analysing a neural network semantic model
5. Beyond words and sentences: Reference, multimodality, discourse, dialogue.

Note: The focus on the different topics will be based on the students' interests.

Teaching Methods

The class will be based on lectures, readings, student presentations, and practical exercises. The exercises will build on one another. Some of them will consist of both, a theory-based, analytical task and a programming task, and the students can choose to practically work on both or focus only on one of these tasks. The students will present the outcome of the exercises in the final class of the course, and write an essay summarizing their findings. They will also be expected to present / lead the discussion of one or more of the readings.

Evaluation

- Practical exercises, presentation, essay: 70%
- Presenting / leading the discussion of one of the readings, participation in reading discussions: 30%

Bibliography and information resources

Bibliography:

- Jurafsky, Daniel & Martin, James H. (2009), Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 2nd edition. Prentice Hall. [Especially Chapter 20.]
- Murphy, Gregory L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press. [Especially Chapter 11].

Readings (tentative list):

- Word meaning:

Kilgarriff, Adam. I don't believe in word senses. *Computers and the Humanities* 31.2 (1997): 91-113.

Murphy, Gregory L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press. *Note: Really great book I recommend to everybody. See especially Chapter 11.*

- Symbolic (formal semantics, DRT-based) system for the processing of free English text:

Bos, Johan (2008). Wide-coverage semantic analysis with Boxer. *Proceedings of the 2008 Conference on Semantics in Text Processing*. Association for Computational Linguistics.

Bos, J., & Markert, K. (2005). Recognising textual entailment with logical inference. *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 628-635). Association for Computational Linguistics.

- Distributional semantics, general:

M. Baroni and A. Lenci. 2010. Distributional Memory: A general framework for corpus-based semantics. Computational Linguistics 36(4): 673-721.

Katrin Erk. Vector space models of word meaning and phrase meaning: a survey. Language and Linguistics Compass 6(10), 635-653, October 2012. *Note: survey article.*

Stephen Clark. 2015. Vector Space Models of Lexical Meaning. *Handbook of Contemporary Semantic Theory — second edition*, edited by Shalom Lappin and Chris Fox. Chapter 16, pp.493-522. Wiley-Blackwell. [PDF] (pre-copy editing). *Note: survey article.*

Alessandro Lenci. 2008. Distributional semantics in linguistic and cognitive research. Italian journal of linguistics, 20 (1), pp. 1-31.

- Multimodal distributional semantics:

Bruni, E., G. Boleda, M. Baroni, N. K. Tran. 2012. Distributional semantics in technicolor. *Proceedings of ACL 2012*, pp. 136-145, Jeju Island, Korea. [**paper**, slides, data]

Silberer, C and Lapata, M. 2013. Learning Grounded Meaning Representations with Autoencoders. *Proceedings of ACL 2013*.

M. Baroni. 2016. Grounding distributional semantics in the visual world. Language and Linguistics Compass 10(1): 3-13. *Note: survey article.*

- Composition in distributional semantics:

M. Baroni. 2013. Composition in distributional semantics. *Language and Linguistics Compass* 7(10): 511-522. *Note: survey article.*

Jeff Mitchell and Mirella Lapata. 2008. Vector-based Models of Semantic Composition. In: ACL. 2008, pp. 236-244.

E. Vecchi, M. Marelli, R. Zamparelli and M. Baroni. 2016. Spicy adjectives and nominal donkeys: Capturing semantic deviance using compositionality in distributional spaces. *Cognitive Science*. To appear.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review*, *122*(3), 485-515. http://doi.org/10.1037/a0039267

Socher, R., Pennington, J., Huang, E.H., Ng, A.Y. and Manning, C.D. 2011. Semi-supervised recursive autoencoders for predicting sentiment distributions. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 151-161).

- Building word vectors with neural networks:

Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, Luke Zettlemoyer. NAACL 2018. Deep contextualized word representations

Jeffrey Pennington, Richard Socher, Christopher Manning. 2014. Glove: Global vectors for word representation Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)

Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781v3.

M. Baroni, G. Dinu and G. Kruszewski. 2014. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. Proceedings of ACL 2014 (52nd Annual Meeting of the Association for Computational Linguistics), East Stroudsburg PA: ACL, 238-247.

Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic Regularities in Continuous Space Word Representations. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746-751). Atlanta, Georgia: Association for Computational Linguistics.

- Current limits of distributional semantics / neural networks:

Bernardi, R., G. Boleda, R. Fernandez, D. Paperno. 2015. Distributional semantics in use. Proceedings of EMNLP 2015 Workshop LSDSem 2015: Linking Models of Lexical, Sentential and Discourse-level Semantics, 95-101. Lisbon, Portugal, September. Association for Computational Linguistics.

Paperno, D., G. Kruszewski, A. Lazaridou, Q. Ngoc, R. Bernardi, S. Pezzelle, M. Baroni, G. Boleda, R. Fernandez. 2016. The LAMBADA dataset: Word prediction requiring a broad discourse context. *Proceedings of ACL 2016* (54th Annual Meeting of the Association for Computational Linguistics), 1525-1534, Berlin, Germany, August. Association for Computational Linguistics.

Boleda, G. and A. Herbelot. 2016. Formal Distributional Semantics: Introduction to the Special Issue. Computational Linguistics 42:4, 619-635.

- General Machine Learning:

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. http://m.mind.oxfordjournals.org/content/LIX/236/433.full.pdf, (if that fails: http://phil415.pbworks.com/f/TuringComputing.pdf)

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78. http://doi.org/10.1145/2347736.2347755

Parloff, R. (2016). The AI Revolution: Why Deep Learning Is Suddenly Changing Your Life. *Fortune* Magazine. *Note: well written, thorough popular science article about deep learning.*

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. http://doi.org/10.1038/nature14539