

Weakly-Supervised Semantic Processing (Seminar)

Thursday 11:15, IMS

Kyle Richardson

kyle@ims.uni-stuttgart.de

http://www.ims.uni-stuttgart.de/institut/mitarbeiter/kyle/wssp_course.html (course materials)

Course Description: This seminar looks at recent advances in machine learning for modeling natural language meaning, or computational semantics. In particular, the focus is on *weakly-supervised* learning methods for semantic parsing (i.e. translating text to formal meaning representations) and various applications thereof (e.g. automated question-answering, natural language problem solving, program induction).

Prerequisites: Familiarity with basic machine learning and linguistic semantics. A knowledge of basic programming concepts is also helpful.

Requirements:

- **Weekly Readings:** Students will be required to keep up with weekly readings drawn from work in the field. Supplementary readings will also be provided to give additional background or context.
- **Reading Summaries:** In addition to the weekly readings, students will also be required to submit summaries of weekly readings to demonstrate an understanding of the work being covered.
- **Presentations:** Students will be required to give presentations of individual papers and lead discussions.
- **Term Paper:** Students will be required to write an in depth term paper on a particular topic (e.g. the paper being presented) in the course. Alternatively, students can implement a particular model (e.g. a semantic parser or generator) from the class if such an idea is approved and discussed with the instructor.

Grading: Reading Summaries: 15%, Presentation: 35%, Term Paper: 50%.

Deadlines:

- **Weekly Reading Summaries:** Midnight the evening before each class in PDF format. Send to kyle@ims.uni-stuttgart.de, with subject line “READING SUMMARY: class date”.
- **Presentation Topic:** Start of lecture 5 (12/5/2016).
- **Term Paper/Project:** July 7th, 2016 by the start of class.

Reading Summaries: Should be a 250 word summary of 5/9 of the week's readings, which can include the **main** papers and **secondary** papers marked with a * (see schedule below). Only one paper per week can be summarized, and at least 3 must be from the **main lecture** period.

The summaries should explain the main points of the paper, as well as any misunderstanding you might have had. In particular, each summary should *explicitly* cover the following questions:

- **Motivation:** What problem is the paper trying to solve? What is the overall objective?
- **Method:** Which methods are used to solve this problem? How does this relate to previous work (based on the paper's description)?
- **Evaluation:** How are these methods evaluated? What are the results?
- **Problems:** Are there any shortcomings of this work? Parts you didn't understand?

Presentations: Student will give a conference-style presentation of an individual paper, explaining in detail the main points of the paper, the methods being used, and the results. Depending on the size of the course, students will either be responsible for one presentation or will work in teams. They will be expected to answer questions from the class and instructor.

- **Deadline:** Student should choose a set of papers/presentation date by the beginning of lecture 6 (12/5/2016).
- **Preparation:** Slides should be prepared a week in advance and sent to instructor for initial approval.

Term Paper or Project: This term paper will be a review of 2 or more papers that focus on a particular idea or problem covered in the class. The paper should be 5-6 pages long, and should cover in detail the methods and models used in the papers. *More details later.*

Main Lectures

| | |
|-----------------------|--|
| Lecture 1 (7/4/2016) | Course Overview |
| <i>no readings</i> | |
| Lecture 2 (14/4/2016) | Big Picture: Mixing Statistical and Compositional Semantics |
| main: | [Liang and Potts, 2015] (required reading summary) |
| secondary: | [Domingos, 2012], [Mooney, 2007] |
| Lecture 3 (21/4/2016) | Learning Rewrite Rules and Parsing |
| main: | [Kate et al., 2005] |
| secondary: | [Kate et al., 2004]*, [Crouch and King, 2006] |
| Lecture 4 (28/4/2016) | Semantic Parsing and Machine Translation |
| main: | [Wong and Mooney, 2006] |
| secondary: | [Wong and Mooney, 2007]*, [Andreas et al., 2013]*, |
| (5/5/2016) | No class, holiday |
| Lecture 5 (12/5/2016) | Structure Prediction and Classification (Presentation topics due) |
| main: | [Zettlemoyer and Collins, 2012] |
| secondary: | [Kwiatkowski et al., 2010]*, [Zettlemoyer and Collins, 2007]* |

Student Presentations

| | |
|--|--|
| 19/5/2016 | No class |
| 26/5/2016 | No class, holiday |
| 2/6/2016 | Student Presentations 1 |
| <i>Ehsan</i> [Jones et al., 2012b] | <i>Mariia K.</i> [Jones et al., 2012a] |
| 9/6/2016 | Student Presentations 2 |
| <i>Bhavani</i> [Kushman et al., 2014] | |
| 16/6/2016 | Student Presentations 3 |
| <i>Beate</i> [Berant et al., 2013] | |
| 23/6/2016 | Student Presentations 4 |
| <i>Ilmar</i> [Kim and Mooney, 2012] | |
| 30/6/2016 | No class (instructor away) |
| 7/7/2016 | Learning from Denotation (last class, term paper due) |
| <i>Kyle (Last Lecture)</i> | [Liang et al., 2011] |

Papers for Presentations: (partial)

MT related: [Jones et al., 2012a] [Jones et al., 2012b] [Li et al., 2013]

Parsing related: [Börschinger et al., 2011, Kim and Mooney, 2012],

Structure Prediction/ Grounded Learning: [Artzi and Zettlemoyer, 2013], [Kushman et al., 2014], [Kushman and Barzilay, 2013], [Berant et al., 2013], [Liang et al., 2011].

Supplementary Resources and Materials

Tutorial on log-linear models (Michael Collins):

<http://www.cs.columbia.edu/~mccollins/loglinear.pdf>

Tutorial on log-linear models and CRFs (Charles Elkan)

<http://cseweb.ucsd.edu/~elkan/250Bwinter2012/loglinearCRFs.pdf>

Tutorial on Semantic Parsing (Yoav Artzi)

<http://yoavartzi.com/tutorial/>

Lectures on Formal Semantics/Lambda Calculus (Barbara Partee)

http://people.umass.edu/partee/RGGU_2004/RGGU04_formal_semantics.htm

CCG Overview (Mark Steedman and Jason Baldridge)

<http://homepages.inf.ed.ac.uk/steedman/papers/ccg/SteedmanBaldridgeNTSyntax.pdf>

Synchronous Grammars Tutorial (David Chiang)

<http://www3.nd.edu/~dchiang/papers/synchtut.pdf>

Expectation Maximization (EM) Overview (Do, Batzoglu)

<http://www.nature.com/nbt/journal/v26/n8/full/nbt1406.html>

Notes on the Inside-Outside algorithm (Jason Eisner)

<https://www.cs.jhu.edu/~jason/465/iobasics.pdf>

References

- [Andreas et al., 2013] Andreas, J., Vlachos, A., and Clark, S. (2013). Semantic parsing as machine translation. In *in Proceedings of ACL-2013*, pages 47–52. <http://www.anthology.aclweb.org/P/P13/P13-2009.pdf>.
- [Artzi and Zettlemoyer, 2013] Artzi, Y. and Zettlemoyer, L. (2013). Weakly supervised learning of semantic parsers for mapping instructions to actions. *Transactions of the Association for Computational Linguistics*, 1:49–62. <https://tacl2013.cs.columbia.edu/ojs/index.php/tacl/article/view/27>.
- [Berant et al., 2013] Berant, J., Chou, A., Frostig, R., and Liang, P. (2013). Semantic parsing on Freebase from question-answer pairs. In *in Proceedings of EMNLP-2013*, pages 1533–1544.
- [Börschinger et al., 2011] Börschinger, B., Jones, B. K., and Johnson, M. (2011). Reducing grounded learning tasks to grammatical inference. In *Proceedings of EMNLP-2011*, pages 1416–1425. http://homepages.inf.ed.ac.uk/s1051107/borschingerJonesJohnson_emnlp2011.pdf.
- [Crouch and King, 2006] Crouch, D. and King, T. H. (2006). Semantics via f-structure rewriting. In *Proceedings of the LFG06 Conference*, pages 145–165. <http://www2.parc.com/isl/groups/nltp/papers/lfg06crouchking-PREPRINT.pdf>.
- [Domingos, 2012] Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87.
- [Jones et al., 2012a] Jones, B., Andreas, J., Bauer, D., Hermann, K. M., and Knight, K. (2012a). Semantics-based machine translation with hyperedge replacement grammars. In *COLING*, pages 1359–1376. <http://mt-archive.info/Coling-2012-Jones.pdf>.
- [Jones et al., 2012b] Jones, B. K., Johnson, M., and Goldwater, S. (2012b). Semantic parsing with Bayesian Tree Transducers. In *Proceedings of ACL-2012*, pages 488–496. <http://homepages.inf.ed.ac.uk/sgwater/papers/acl12-transducers.pdf>.

- [Kate et al., 2004] Kate, R. J., Wong, Y. W., Ge, R., and Mooney, R. J. (2004). Learning transformation rules for semantic parsing. <http://www.cs.utexas.edu/users/ai-lab/?kate:tr04>.
- [Kate et al., 2005] Kate, R. J., Wong, Y. W., and Mooney, R. J. (2005). Learning to transform natural to formal languages. In *Proceedings of the National Conference on Artificial Intelligence*, volume 20, page 1062. <http://www.aaai.org/Library/AAAI/2005/aaai05-168.php>.
- [Kim and Mooney, 2012] Kim, J. and Mooney, R. J. (2012). Unsupervised pcfg induction for grounded language learning with highly ambiguous supervision. In *Proceedings of EMNLP-CoNLL-2012*, Jeju Island, Korea. <http://dl.acm.org/citation.cfm?id=2391001>.
- [Kushman et al., 2014] Kushman, N., Artzi, Y., Zettlemoyer, L., and Barzilay, R. (2014). Learning to automatically solve algebra word problems. In *Proceedings of ACL-2014*, pages 271–281. https://people.csail.mit.edu/regina/my_papers/wp.pdf.
- [Kushman and Barzilay, 2013] Kushman, N. and Barzilay, R. (2013). Using semantic unification to generate regular expressions from natural language. In *Proceedings of NAACL-13*. <http://dspace.mit.edu/handle/1721.1/79645>.
- [Kwiatkowski et al., 2010] Kwiatkowski, T., Zettlemoyer, L., Goldwater, S., and Steedman, M. (2010). Inducing probabilistic ccg grammars from logical form with higher-order unification. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 1223–1233. Association for Computational Linguistics. <http://www.aclweb.org/anthology/D/D10/D10-1119.pdf>.
- [Li et al., 2013] Li, P., Liu, Y., and Sun, M. (2013). An extended ghkm algorithm for inducing lambda-scfg. In *AAAI*. <http://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6189>.
- [Liang et al., 2011] Liang, P., Jordan, M. I., and Klein, D. (2011). Learning dependency-based compositional semantics. In *Proceedings of ACL-11*, pages 590–599.
- [Liang and Potts, 2015] Liang, P. and Potts, C. (2015). Bringing machine learning and compositional semantics together. *Annual Reviews of Linguistics*, 1(1):355–376. <http://www.annualreviews.org/doi/abs/10.1146/annurev-linguist-030514-125312>.
- [Mooney, 2007] Mooney, R. J. (2007). Learning for semantic parsing. In *Computational Linguistics and Intelligent Text Processing*, pages 311–324. Springer.
- [Wong and Mooney, 2006] Wong, Y. W. and Mooney, R. J. (2006). Learning for semantic parsing with statistical machine translation. In *Proceedings of HLT-NAACL-2006*, pages 439–446. <http://dl.acm.org/citation.cfm?id=1220891>.
- [Wong and Mooney, 2007] Wong, Y. W. and Mooney, R. J. (2007). Learning synchronous grammars for semantic parsing with lambda calculus. In *Proceedings of ACL-2007*, Prague, Czech Republic. <http://anthology.aclweb.org/P/P07/P07-1121.pdf>.
- [Zettlemoyer and Collins, 2007] Zettlemoyer, L. S. and Collins, M. (2007). Online learning of relaxed ccg grammars for parsing to logical form. In *EMNLP-CoNLL*, pages 678–687. <http://anthology.aclweb.org/D/D07/D07-1.pdf>.
- [Zettlemoyer and Collins, 2012] Zettlemoyer, L. S. and Collins, M. (2012). Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars. *arXiv preprint arXiv:1207.1420*. <http://arxiv.org/abs/1207.1420>.