Andreas Zollmann

www.cs.cmu.edu/~zollmann/publications.html

Am Tatzend 2, D-07745 Jena, GERMANY   **Tel. +1-903-326-3454**   zollmann@cs.cmu.edu

S p e c i a l t i e s

- ☐ *Machine learning, natural language processing (esp. parsing and statistical machine translation), computational logic*

E d u c a t i o n

Carnegie Mellon University   *since 2004*

Pittsburgh, U.S.A.

- ☐ *Ph.D. candidate; Language Technology Institute, School of Computer Science*
  Advisors: Dr. Stephan Vogel and Dr. Alex Waibel
- ☐ *Working mainly on synchronous-grammar based approaches to statistical machine translation*
- ☐ *Teaching assistant in Ph.D. level courses "Machine Learning" and "Inventing the Future of Services"*

University of Amsterdam   *2002 - 2004*

Institute for Logic, Language and Computation

The Netherlands

- ☐ *Enrolled in the Graduate Program in Logic as a Master of Science student*
- ☐ *Degree: Master of Science, grade 9.3 – with honors – (range of grading scheme: 0-10), ECTS grade A*
- ☐ ***Master Thesis** (domain: statistical natural language processing)*
  Title: A Consistent and Efficient Estimator for the Data-Oriented Parsing Model
  Advisors: Dr. Khalil Sima'an and Dr. Detlef Prescher

Technical University of Dresden   *1999 - 2000, 2001 - 2002*

Germany

- ☐ *Study of Computer Science with secondary subject Mathematics*
- ☐ *Degree: **Vordiplom**; grade 1.2 (range 1-5, where 1.0 best); ECTS grade A (ranked 2nd amongst ca. 450 students)*

University of Sheffield   *2000 / 2001*

United Kingdom

- ☐ *Year abroad within the Socrates / Erasmus Exchange Program*
- ☐ *Enrolled in 10 courses in the domains Natural Language Processing, Software Engineering, and Theory of Computer Science; best student in three courses*

R e s e a r c h   V i s i t s

University of Karlsruhe   *Nov 2005 - Dec 2006*

Germany

- ☐ *Stay in the Artificial Intelligence Lab as a fellow of the InterACT exchange program*
- ☐ *Resulted in publications in HLT/NAACL, IWSLT, ICSLP*

Google Inc.   *Aug 2007 - May 2008*

Mountain View, CA, U.S.A.

- ☐ *Intern in the statistical machine translation research group*
- ☐ *Implemented Syntax-Augmented MT model (Zollmann&Venugopal, 2006) there*
- ☐ *Resulted in publication in CoLing-2008*

University of Edinburgh   *Jul 2009 - Aug 2009*

United Kingdom

- ☐ *Developed nonparametric Bayesian model to label phrase pairs, resulting in a probabilistic synchronous context-free grammar for machine translation; work still ongoing*

L a n g u a g e   P r o f i c i e n c y

- ☐ *Germany (native), English (fluent), Dutch (interm.), French (interm.), Russian (basic)*

Additional Skills and Research Activities

- ☐ *Programming skills in Java, C++, Perl, Unix shells, SQL, Visual Basic, Delphi, Prolog, Haskell, Assembler*
- ☐ *System skills in Hadoop (esp. MapReduce), MinorThird, Weka, SVM-light, MySQL, Berkeley DB, CVS, Subversion*
- ☐ *Conference program committee member for: ACL, EMNLP, IJCAI, COLING, NAACL, WMT*
- ☐ *Journal Reviewer for: Journal of Automata, Languages and Combinatorics; Journal of Machine Learning Research*

Awards

- ☐ *30th German Mathematical Competition 1990: **third** (region Jena); **seventh** (state Thuringia)*
- ☐ *33rd German Mathematical Competition 1993: **fifth** (region Jena)*
- ☐ *German Scientific Competition "Jugend forscht" 1997: **winner** (region East Thuringia); **winner** (state Thuringia); **special price** REHA international exhibition; **special price** technology)*
- ☐ *6th National Portuguese Logic Programming Competition (cenPL) 2003: **second***
- ☐ *Fellowship in "Studienstiftung des deutschen Volkes" (German National Academic Foundation)*

References

- ☐ *Dr. Stephan Vogel (Carnegie Mellon University):* vogel@cs.cmu.edu
- ☐ *Dr. Noah Smith (Carnegie Mellon University):* nasmith@cs.cmu.edu
- ☐ *Dr. Jay Ponte (Google Inc.):* ponte@google.com

Patents and Selected Publications (full list: `www.cs.cmu.edu/~zollmann/publications.html`)

- ☐ *Design patent no. 297 07 958.1, Germany:* "Foot-controlled pointing device for computers"
  Owners: Gommel, Christoph; Hauser, Gerlind; Zollmann, Andreas; day of registry: 3 July 1997
- ☐ *Andreas Zollmann and Stephan Vogel. 2010.* New Parameterizations and Features for PSCFG-Based Machine Translation. *In* Proc. of the Fourth Workshop on Syntax and Structure in Statistical Translation (SSST) at the 2010 Conference on Computational Linguistics (COLING-2010)*, Beijing, China.*
- ☐ *Ashish Venugopal, Andreas Zollmann, Noah A. Smith, and Stephan Vogel. 2009.* Preference Grammars: Softening Syntactic Constraints to Improve Statistical Machine Translation. *In* Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (NAACL/HLT)*, Boulder, Colorado.*
- ☐ *Ashish Venugopal and Andreas Zollmann. 2009.* Grammar based statistical MT on Hadoop: An end-to-end toolkit for large scale PSCFG based MT. *In* The Prague Bulletin of Mathematical Linguistics No. 91, 2009, pages 67-78.
- ☐ *Ashish Venugopal, Andreas Zollmann, Noah A. Smith, and Stephan Vogel. 2008.* Wider Pipelines: N-Best Alignments and Parses in MT Training. *In* Proc. of the Association of Machine Translation in the Americas (AMTA)*, Waikiki, Hawaii.*
- ☐ *Andreas Zollmann, Ashish Venugopal, Franz Och and Jay Ponte. 2008.* A Systematic Comparison of Phrase-Based, Hierarchical and Syntax-Augmented Statistical MT. *In* Proc. of 22nd International Conference on Computational Linguistics (Coling)*, Manchester, U.K.*
- ☐ *Ashish Venugopal, Andreas Zollmann and Stephan Vogel. 2007. An Efficient Two-Pass Approach to Synchronous-CFG Driven Statistical MT. In* Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (HLT/NAACL)*, Rochester, NY. April 22-27.*
- ☐ *Andreas Zollmann and Ashish Venugopal. 2006. Syntax Augmented Machine Translation via Chart Parsing. In* Proc. of NAACL 2006 - Workshop on statistical machine translation*, New York. June 4-9.*
- ☐ *Andreas Zollmann, Ashish Venugopal and Stephan Vogel. 2006. Bridging the inflection morphology gap for Arabic statistical machine translation. In Short Papers in the Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (HLT/NAACL), New York. June 4-9.*
- ☐ *Roger Hsiao, Ashish Venugopal, Thilo Köhler, Ying Zhang, Paisarn Charoenpornsawat, Andreas Zollmann, Stephan Vogel, Alan W Black, Tanja Schultz and Alex Waibel. 2006. Optimizing Components for Handheld Two-way Speech Translation for an English-Iraqi Arabic System. In Proc. of ICSLP-2006, Pittsburgh.*
- ☐ *Detlef Prescher, Remko Scha, Khalil Sima'an and Andreas Zollmann. 2006. What Are Treebank Grammars? In* BNAIC'06, Namur, Belgium.
- ☐ *Andreas Zollmann and Khalil Sima'an. 2005. A Consistent and Efficient Estimator for Data-Oriented Parsing. In Journal of Automata, Languages and Combinatorics Vol. 10 (2005) Number 2/3, pages 367-388.*
- ☐ *Ashish Venugopal, Andreas Zollmann and Alex Waibel. 2005.* Training and Evaluating Error Minimization Decision Rules for Statistical Machine Translation. *In* Proc. of ACL Workshop on Building and Using Parallel Corpora*, Ann Arbor.*
- ☐ *Detlef Prescher, Remko Scha, Khalil Sima'an and Andreas Zollmann. 2004.* On the Statistical Consistency of DOP Estimators. *In* Proceedings of the 14th Meeting of Computational Linguistics in the Netherlands*, Antwerp, Belgium.*