

## Jonathan Weese

1605 Sansom Street  
Apartment 215  
Philadelphia, PA 19103  
cell: 410 913 2781  
jonny@cs.jhu.edu

### Education

**Johns Hopkins University**, Baltimore, MD  
Ph.D. in Computer Science, expected July 2014  
M.S.E. in Computer Science, May 2011

**The University of Chicago**, Chicago, IL  
A.B. in Mathematics, June 2008, GPA 3.21/4.0  
Dean's List, 2006–2008

### Experience

**Center for Language and Speech Processing**, Baltimore, MD  
July 2008 – Present  
Research assistant.

**Human Language Technology Center of Excellence**, Baltimore, MD  
June – August 2013  
Researcher, SCALE 2013 Workshop.

**Google, Inc.**, Mountain View, CA  
June – August 2012  
Software engineering intern, Google Translate.

**Johns Hopkins University**, Baltimore, MD  
January – May 2012  
Teaching assistant for 600.468 *Machine Translation*.

**Human Language Technology Center of Excellence**, Baltimore, MD  
June – August 2011  
Researcher, SCALE 2011 Workshop.

**Johns Hopkins University**, Baltimore, MD  
September – December 2009  
Teaching assistant for 600.465 *Natural Language Processing*.

### Projects

**Thrax** (principal author), a grammar extractor that builds context-free grammar translation models from aligned parallel corpora. It uses Hadoop for efficient model extraction even from very large corpora.

**Joshua** (contributor), an open-source machine translation decoder developed at Johns Hopkins. I contributed a framework for visualizing the decoder output to aid in debugging translation models. I also adapted the decoder to perform synchronous parsing of sentence pairs.

**cdec** (contributor), another machine translation decoder, developed mostly at the University of Maryland. I helped in feature function implementation.

## Publications

### 2013

*Joshua 5.0: Sparser, Better, Faster, Server.* M. Post, J. Ganitkevitch, L. Orland, J. Weese, Y. Cao and C. Callison-Burch. *Proc. WMT*.

*PARMA: a Predicate Argument Aligner.* T. Wolfe, B. Van Durme, M. Dredze, N. Andrews, C. Beller, C. Callison-Burch, J. DeYoung, J. Snyder, J. Weese, T. Xu and X. Yao. *Proc. ACL* (short papers).

*UMBC\_EBIQUITY-CORE: Semantic Textual Similarity Systems.* L. Han, A.L. Kashyap, T. Finin, J. Weese and J. Mayfield. *Proc. \*SEM*.

### 2012

*Joshua 4.0: Packing, PRO, and Paraphrases.* J. Ganitkevitch, Y. Cao, J. Weese, M. Post and C. Callison-Burch. *Proc. WMT*.

*Using Categorical Grammar to Label Translation Rules.* J. Weese, C. Callison-Burch and A. Lopez. *Proc. WMT*.

*Processing Informal, Romanized Pakistani Text Messages.* A. Irvine, J. Weese and C. Callison-Burch. *NAACL Workshop on Language in Social Media*.

### 2011

*Joshua 3.0: Syntax-based Machine Translation with the Thrax Grammar Extractor.* J. Weese, J. Ganitkevitch, C. Callison-Burch, M. Post and A. Lopez. *Proc. WMT*.

### 2010

*Joshua 2.0: A Toolkit for Parsing-based Machine Translation with Syntax, Semirings, Discriminative Training and Other Goodies.* Z. Li, C. Callison-Burch, C. Dyer, J. Ganitkevitch, A. Irvine, L. Schwartz, W.N.G. Thornton, Z. Wang, J. Weese and O.F. Zaidan. *Proc. WMT*.

*cdec: A Decoder, Alignment, and Learning Framework for Finite-State and Context-Free Translation Models.* C. Dyer, A. Lopez, J. Ganitkevitch, J. Weese, F. Ture, P. Blunsom, H. Setiawan, V. Eidelman and P. Resnik. *Proc. ACL* (system demonstrations).

*Visualizing Data Structures in Parsing-based Machine Translation.* J. Weese and C. Callison-Burch. *Prague Bulletin of Mathematical Linguistics*.

*The Machine Translation Toolpack for LoonyBin: Automated Management of Experimental Machine Translation HyperWorkflows.* J.H. Clark, J. Weese, B.G. Ahn, A. Zollmann, Q. Gao, K. Heafield and A. Lavie. *Prague Bulletin of Mathematical Linguistics*.

### 2009

*Demonstration of Joshua: an Open Source Toolkit for Parsing-based Machine Translation.* Z. Li, C. Callison-Burch, C. Dyer, J. Ganitkevitch, S. Khudanpur, L. Schwartz, W.N.G. Thornton, J. Weese and O.F. Zaidan. *Proc. ACL* (system demonstrations).

*Joshua: an Open Source Toolkit for Parsing-based Machine Translation.* Z. Li, C. Callison-Burch, C. Dyer, J. Ganitkevitch, S. Khudanpur, L. Schwartz, W.N.G. Thornton, J. Weese and O.F. Zaidan. *Proc. WMT*.

*Joshua: an Open Source Toolkit for Parsing-based Machine Translation.* Z. Li, C. Callison-Burch, C. Dyer, J. Ganitkevitch, S. Khudanpur, L. Schwartz, W.N.G. Thornton, J. Weese and O.F. Zaidan. *Prague Bulletin of Mathematical Linguistics*.

**Skills**

Programming languages: Java, C, Haskell, Perl  
Distributed frameworks: Hadoop  
Natural languages: English, Mandarin Chinese, French

**Other Interests**

Drawing, board games, piano, cycling