

Keith D. Stevens

805-433-2295; 18375 Ventura Blvd #574 Tarzana, CA 91356

fozziethebeat@gmail.com

<http://fozziethebeat.github.com>

Education

University of California, Los Angeles

Los Angeles, CA

- *PhD in Computer Science, Focus: Word Sense Induction* GPA:3.847; In Progress
- *Masters in Computer Science, Focus: Computational Linguistics* GPA:3.847; Fall 2011
- *Bachelors in Computer Science* GPA:3.588; Fall 2007
 - Key Courses: Machine Learning, Learning Theory, Language and Thought, Connectionist Systems, Advanced Operating Systems, Static Analysis, Parallel Languages
 - Courses Taught: Modern Compilers, Operating System Principles, Data Structures and Algorithms, Software Construction, Ethics and Engineering

Experience

Lawrence Livermore National Lab

Livermore, CA

Lawrence Livermore Scholar

June 2011 – Present

- Evaluated accuracy and effectiveness of many Word Sense Induction models with Ensemble Clustering
- Developed a new Ensemble Clustering package for Java
- Expanded a new Word Sense Induction package for Java
- Extended evaluations for Word Sense Induction models with new semantic measures and baselines
- Evaluated three Topic Models using coherence metrics to better define which use cases suited each algorithm

Lawrence Livermore National Lab

Livermore, CA

Summer Research Intern

June 2010 – September 2010

- Developed a new API for directly modifying the WordNet lexical ontology as the C-Cat Package
- Applied distributional semantics to customizing lexical ontologies
- Developed a framework for building Word Sense Induction Models

Google Inc.

Kirkland, WA

Software Engineering Intern

June 2009 – September 2009

- Enhanced a document selection system with the addition of semantic analysis that represented the change of semantic content in the document.
- Enhanced crawling framework for the document selection system, improving savings by up to 50% usage of external resources
- Restructured document selection system to be more scalable and abide by internal policies

Google Inc.

Kirkland, WA

Software Engineering Intern

January 2008 – September 2009

- Built a highly scalable, stateless server for responding to a wide variety, and high volume, of requests
- Developed several plugins for the server allowing uniform access to a wide variety of internal data sources
- Improved response query formulation time with a plugin for the above server

Skills and Interests

Topics of Interest: Distributional Semantics, Word Sense Disambiguation, Word Sense Induction, Ontologies, Topic Models, Machine Learning, Clustering, Evaluation Metrics, Statistical Inference, Dimensionality Reduction

Programming Languages: Scala, Java, Bash, R, Python, C/C++, LaTeX

Open Source Projects

The C-Cat Project: An open source project centered around customizing lexical ontologies based on distributional semantics mined from large scale corpora. This includes a new API for directly modify WordNet and several Word Sense Disambiguation implementations. This package uses the Hadoop and HBase frameworks for distributed processing and storage.

The S-Space Project: An open source framework for using, building, and evaluating word space algorithms. It provides several reference implementations and a wide variety of utility classes in addition to several highly optimized clustering algorithms. This also includes a framework for Word Sense Induction models.