

# Sean Gerrish

Princeton University  
Department of Computer Science  
35 Olden Street  
Princeton, NJ 08544

Phone: (650) 861-2108  
Email: [sgerrish@cs.princeton.edu](mailto:sgerrish@cs.princeton.edu)  
Homepage: <http://www.seangerrish.com>

## Academic history

- 2008-present: PhD. Computer Science (in progress), Princeton University (pending 2013). Thesis concentration in machine learning. 3.79 GPA
- 2004: B.S. Honors Mathematics (minor in Physics), University of Michigan. 3.84 GPA
- 1998-2000: Dual enrolled for Mathematics courses, Delta Community College. 4.0 GPA

## Employment and work experience

1. January 2005-August 2008: **Software Engineer, Google, Inc.**
  - November 2007-August 2008: **Senior software engineer and technical lead for an Ads Quality engineering group.** Led a small, data-driven engineering group to improve the quality of certain ads shown on Google.com.
  - **Decision Support.** Worked with a quantitatively focused team of engineers and quantitative analysts to aid internal decision making and build data tools for groups like finance, legal, and Ads.
  - **Ads Quality, Landing Page Quality.** Role included writing software infrastructure to improve the quality of user experience on advertisers' sites.
  - **Conducted over 50 engineering interviews,** served on a hiring committee, and officially mentored two incoming engineers and one statistician.
2. June 2009 - August 2009: **Advanced Technology Research Intern, JSTOR (Ithaka Inc.)** Designed and implemented several search prototypes to provide article-specific themes to users. Demo available at <http://dbrowser.jstor.org> (IE and FireFox). Gave two presentations on results at Ithaka.
3. May 2004-December 2004: **Programmer, University of Michigan Computational Linguistics and Information Retrieval group.** Tasks included simulation of World Wide Web growth based on lexical similarity of documents, data gathering, statistical analysis of network properties, and documentation of projects.
4. June 2004-August 2004: **Assistant in Research, University of Michigan Engineering Research Center.** Tasks included investigation and automation of the cleaning of machining cells; retrofitting machine cells with new software; and design of a DeviceNet inter-node communication protocol.
5. January 2003-August 2003: **REU student, combinatorics research.** Work included study of planar directed networks, polynomials with positive roots, and simplicial complexes. Mentored by Mark Skandera (See publications below).
6. July 2001-April 2003: **Assistant in research, University of Michigan Physics department.** Student technician assembling and testing parts of muon detection endcaps for the ATLAS particle detection project; Work included training two new employees and 6-week assignment at CERN in Geneva, Switzerland.

## Relevant coursework

- Machine learning, data, and applications
  1. Language and Information (**A**)  
Computational linguistics applications, information theory, language models, and text summarization.
  2. Independent Study: Lexical networks. (**A**)  
Directed study on small-world networks, social networks, and the growth of networks with lexical nodes. Mentored by Dragomir Radev.
  3. Analysis and Visualization of Large-Scale Genomic Datasets (**A+**)  
Survey of applications of machine learning to genomic datasets.
  4. Truth in Data (**A-**)  
Seminar on classic and modern work in model selection, causality, and drawing meaningful conclusions from data.
  5. Boosting: Foundations and Algorithms (Not graded). Seminar on boosting.
  6. Machine Learning Summer School, Cambridge University 2009.  
Two weeks of lectures and sessions on theory and applications of machine learning, including MCMC methods, information theory, and Bayesian statistics.
- Selected mathematics courses
  1. Probability (**A, A+, A**)  
Introduction to Probability, measure-based Probability Theory, and Foundations of Probabilistic Modeling (graphical models, inference, MCMC, EM, mixture models, etc.).
  2. Combinatorics (**A+, B+**)  
Combinatorics & Graph theory and Combinatorial Theory, both rigorous (i.e., proof-based) approaches to combinatorics.
  3. Linear Algebra (**A, A+**)  
Introductions to applied linear algebra and honors (abstract) linear algebra.
  4. Markov Chains (**A**)
  5. Calculus I-III and Differential Equations (**A,A,A,A**)

## Teaching, tutoring, and grading experience

- **TA, Interacting with Data** (Princeton COS 424). A senior-level course providing students from various departments with tools for interacting with data, including dimensionality reduction, requisite probability, graphical models, and classification. Responsibilities include hosting an R tutorial session, office hours, and grading.
- **TA, Computers in our world** (Princeton COS 109). An introduction to computer science for non-concentrators.
- **Grader, Honors Introduction to point-set topology** (Umich MATH 590). An honors introduction to (proof-based) topology.
- **Tutor, University of Michigan MathLab.** (August 2003-April 2004) Tutored courses included Honors Calculus I-III; Linear Algebra; Probability; Abstract Algebra; and Differential Equations.

## Publications

### *Refereed journals and conference proceedings*

- Brian Drake, Sean Gerrish, and Mark Skandera. “Two new criteria for comparison in the Bruhat order”, *The Electronic Journal of Combinatorics*. 11 no. 1 (2004).

We give two new criteria by which pairs of permutations may be compared in defining the Bruhat order (of type A). One criterion utilizes totally nonnegative polynomials and the other utilizes Schur functions.

- Brian Drake, Sean Gerrish, and Mark Skandera, “Monomial nonnegativity and the Bruhat order”, *The Electronic Journal of Combinatorics*. 11 (2005) (2) research paper R18, 5 pp.

We show that five nonnegativity properties of polynomials coincide when restricted to polynomials of the form  $x_1, \pi(1) \dots x_n, \pi(n) - x_1, \sigma(1) \dots x_n, \sigma(n)$ , where  $\pi$  and  $\sigma$  are permutations in  $S_n$ . In particular, we show that each of these properties may be used to characterize the Bruhat order on  $S_n$ .

- Jonathan Chang, Jordan Boyd-Graber, Sean Gerrish, Chong Wang, and David Blei. “Reading tea leaves: How humans interpret topic models”. *Neural Information Processing Systems*, 2009. **Honorable Mention and NIPS oral**.

... We present new quantitative methods for measuring semantic meaning in inferred topics. We back these measures with large-scale user studies, showing that they capture aspects of the model that are undetected by previous measures of model quality based on held-out likelihood.

- Lada A. Adamic, Xiao Wei, Jiang Yang, Sean Gerrish, Kevin K Nam, Gavin S. Clarkson, “Individual Focus and Knowledge Contribution” *First Monday*. 15 no. 3 (1 March 2010) **Lead Article**

... we measure the relationship between the narrowness of focus and the quality of contribution across a range of both traditional and recent knowledge sharing media, including scholarly articles, patents, Wikipedia, and online question and answer forums. Across all systems, we observe a small but significant positive correlation between focus and quality.

- Sean Gerrish and David Blei. “A language-based approach to measuring scholarly impact.” *International Conference on Machine Learning*, 2010.

We propose using changes in the thematic content of documents over time to measure the importance of individual documents within the collection. We describe a dynamic topic model for both quantifying and qualifying the impact of these documents. We validate the model by analyzing three large corpora of scientific articles. Our measurement of a document’s impact correlates significantly with its number of citations.

### *Workshops, Colloquia, and Student Conferences*

- Sean Gerrish and David Blei. “Modeling Influence in Text Corpora” *NIPS Workshop on Applications for Topic Models: Text and Beyond*. 2009. **Workshop oral**.

- Sean Gerrish and David Blei. “The Ideal Point Topic Model: Predicting Legislative sentiment from text.” *Computational Social Science and the Wisdom of Crowds (NIPS Workshop)*. December 2010. **Oral Presentation**

- Sean Gerrish and David Blei. The Ideal Point Topic Model: Predicting Legislative sentiment from text. New York Academy of Sciences 5th Annual Machine Learning Symposium. November 2010. (Poster)
- Sean Gerrish and David Blei. Modeling Influence in Text Corpora. Applications for Topic Models: Text and Beyond (NIPS workshop). December 2010. **Oral Presentation**
- Sean Gerrish and David Blei. Measuring Article Influence without Citations. Northeast Student Colloquium on Artificial Intelligence (NESCAI), April 2010. **Oral presentation**
- Sean Gerrish and David Blei. Modeling Influence in Text Corpora. New York Academy of Sciences 4th Annual Machine Learning Symposium, November 2009. (Poster)

## Honors and Awards

- 2010: Finalist for the Facebook Fellowship program (applied to the machine learning fellowship)
- 2008,2009: Honorable Mention, NSF Graduate Research Fellowship
- 2007: Nominated for and participated in Google's Edge leadership training
- 2006: Google EMG Award (awarded to outstanding engineering groups)
- 2004: Graduated with high distinction and high honors in mathematics at the University of Michigan
- 2004: Awarded an Outstanding Achievement in Mathematics Award by the University of Michigan Mathematics department  
Awarded to about 10 undergraduates each year for outstanding mathematics coursework
- 2003: Copeland Merit Scholarship for Mathematics coursework
- 2001: William J. Branstom Freshman Prize  
Awarded to students for a 4.0 G.P.A in their first semester.
- 2000: University of Michigan Honors program
- 2000: Society of Automotive Engineers merit scholarship
- 2000: Monsignor Forbes Merit Scholarship
- 2000: Third Place in Level II of the annual Math Olympics hosted by Saginaw Valley State University; offered SVSU full-ride scholarship
- 2000: American Society of Metallurgists merit scholarship
- 2000: Placed in top 100 of 13,800 in 43rd Annual Michigan Mathematics Prize Competition
- 1999: American Society of Metallurgists merit scholarship for 7-day materials science program at Michigan Tech
- 1996: Placed first in region and twenty-first in state in the Michigan Council of Teachers of Mathematics competition
- 1996: Third Place in Level I of the annual Math Olympics hosted by Saginaw Valley State University
- 1996: Fourth Place in regional MathCounts competition