

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Shasha, Dennis

eRA COMMONS USER NAME (credential, e.g., agency login): des1.nyu

POSITION TITLE: Julius Silver Professor of Computer Science and Data Science

EDUCATION/TRAINING (*Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.*)

| INSTITUTION AND LOCATION | DEGREE (if applicable) | END DATE MM/YYYY | FIELD OF STUDY |
|-----------------------------------|---------------------------|---------------------|---------------------------------|
| Yale University, New Haven, CT | BS | 1977 | Engineering and Applied Science |
| Syracuse University, Syracuse, NY | MS | 1980 | Computer Science |
| Harvard University, Cambridge, MA | PHD | 1984 | Applied Mathematics |

A. Personal Statement

Dennis Shasha is a professor of computer science at the Courant Institute of Mathematics at New York University where he works with biologists on pattern discovery for network inference; with computational chemists on algorithms for protein design; and on computational reproducibility. Other areas of interest include database tuning as well as tree and graph matching. He has co-authored over seventy journal papers, seventy conference papers, and twenty patents. Because he likes to type, he has written six books of puzzles about a mathematical detective named Dr. Ecco, a biography about great computer scientists, and a book about the future of computing. He has also written four technical books about database tuning, network inference, time series, DNA computing, and statistics. The work in this proposal directly uses the work on graph matching, motif finding, and machine learning.

B. Positions and Honors

Positions and Employment

1984 - 1990 Assistant Professor, Courant Institute, New York University, New York, NY
 1990 - 1995 Associate Professor, Courant Institute, New York University, New York, NY
 1995 - Professor of Computer Science, Courant Institute, New York University, New York, NY
 2018 - Julius Silver Professor of Computer Science and Data Science, Courant Institute, New York University, New York, NY

Other Experience and Professional Memberships

1993 - Co-Editor-in-Chief (with Gottfried Vossen), Information Systems, Publisher: Elsevier North Holland
 2003 - Distinguished Science Advisor, New York Hall of Science
 2014 - Fellow, ACM
 2015 - 2019 International Chair, INRIA

Honors

C. Contribution to Science

1. Understanding causality in transcriptional networks when there are few data points. The goal of this line of work is to find out which genes/proteins are influencing other genes and proteins and how. When there are few data points (e.g. short time series), this problem is very underdetermined. For this

reason, it is important to make use of simple models (linear and perhaps quadratic) and find methods that are robust to noise. Our work in that area includes the following papers:

- a. Krouk G, Lingeman J, Colon AM, Coruzzi G, Shasha D. Gene regulatory networks in plants: learning causality from time and perturbation. *Genome Biol.* 2013 Jun 27;14(6):123. PubMed PMID: [23805876](#); PubMed Central PMCID: [PMC3707030](#).
 - b. Jesse L, Shasha D. *Network Inference in Molecular Biology -- a hands-on framework*. Springer Verlag. 2012;
 - c. Krouk G, Mirowski P, LeCun Y, Shasha DE, Coruzzi GM. Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. *Genome Biol.* 2010;11(12):R123. PubMed PMID: [21182762](#); PubMed Central PMCID: [PMC3046483](#).
 - d. Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN. A gene expression map of the Arabidopsis root. *Science.* 2003 Dec 12;302(5652):1956-60. PubMed PMID: [14671301](#).
2. Querying graphs for topological patterns and finding over-represented patterns in networks. The goal of this work is to locate portions of graphs that reflect known patterns or to describe new patterns. Because most interesting versions of this problem are NP-Complete, performance is a main concern.
- a. Giugno R, Bonnici V, Bombieri N, Pulvirenti A, Ferro A, Shasha D. GRAPES: a software for parallel searching on biological graphs targeting multi-core architectures. *PLoS One.* 2013;8(10):e76911. PubMed PMID: [24167551](#); PubMed Central PMCID: [PMC3805575](#).
 - b. Di Natale R, Ferro A, Giugno R, Mongiovi M, Pulvirenti A, Shasha D. SING: subgraph search in non-homogeneous graphs. *BMC Bioinformatics.* 2010 Feb 19;11:96. PubMed PMID: [20170516](#); PubMed Central PMCID: [PMC2850364](#).
 - c. Ferro A, Giugno R, Mongiovi M, Pulvirenti A, Skripin D, Shasha D. GraphFind: enhancing graph searching by low support data mining techniques. *BMC Bioinformatics.* 2008 Apr 25;9 Suppl 4:S10. PubMed PMID: [18460171](#); PubMed Central PMCID: [PMC2367637](#).
 - d. Ferro A, Giugno R, Pigola G, Pulvirenti A, Skripin D, Bader GD, Shasha D. NetMatch: a Cytoscape plugin for searching biological networks. *Bioinformatics.* 2007 Apr 1;23(7):910-2. PubMed PMID: [17277332](#).

D. Additional Information: Research Support and/or Scholastic Performance

Ongoing Research Support

1934388, NSF Birnbaum (PI) 11/01/19-10/31/23

Comparative developmental dynamics: single-cell analysis of maize meristem trajectories

In this project, the goal is to develop multi-cellular models among genes showing how the evolution of successor cells interacts with gene regulation for the less studied Maize plant.

Role: CPI

1840761, NSF Coruzzi (PI) 04/01/19-03/31/21

Uncovering the molecular mechanisms that integrate nutrient and water dose sensing and impact crop production

In this project, the goal is to develop a computational model that can predict how phenotypes of interest (e.g. grain yield) can be predicted based on gene expression and nutrient/water mixes.

Role: CPI

133962, NSF Coruzzi (PI) 09/01/14-08/31/20

NutriNet: A Network Inspired Approach to Improving Nutrient Use Efficiency (NUE) in Crop Plants

In this project, the goal is to develop computational models to suggest which genes to over-express to

increase nitrogen use efficiency.

Role: CPI

1822156, NSF Birnbaum (PI) 07/01/18-06/30/20

Developing High-Throughput CRISPR/Single-cell RNA-seq Screening in Maize

In this project, the goal is to develop causal models among genes based on CRISPR interference in promoters.

Role: CPI