Resume
Ruchira S. Datta

Ruchira.Datta@gmail.com
(510) 761-3949
http://www.ruchiradatta.com

Objective: Part-time, full-time, or contract/consulting work, including systems analysis or modeling, and software architecture, design, or development.

Software Development

  Work Experience
    10/11-Present: Center for Evolution and Cancer, UCSF. Research Specialist. Applying ecology and evolutionary biology to discover how pre-cancerous lesions progress to cancer and how cancer develops resistance to therapy.

    06/07-9/11: Berkeley Phylogenomics Group, UC Berkeley. Research Specialist. Developing novel computational methods in protein informatics, including protein structure prediction and pathway prediction. Directing web application development and scientific computation on new compute cluster under the supervision of Principal Investigator Kimmen Sj√∂lander. Designing and implementing algorithms and protocols using trees, profile HMMs (hidden Markov models), and techniques from statistics and machine learning.
      Designed and implemented the PHOG algorithm for finding orthologs in gene family trees in C++ with STL and Boost, with interfaces to MySQL and PostgreSQL databases, and led the implementation of the PHOG ortholog query engine/web application in Python and Django. Work with MySQL and PostgreSQL also includes architecting the schema and enhancing performance through indexes, clustering, and optimizing queries. Architecting and leading the implementation of version 3.0 of the PhyloFact protein informatics web application using Python and Django. Maintained and modified the previous (currently visible) version of PhyloFacts. Wrote and modified JavaScript, CSS and HTML, in PhyloFacts, PHOG, and SATCHMO-JS. Maintained, modified and refactored Perl scripts. Maintained and modified PHP scripts. Made extensive use of distributed computing with Sun GridEngine and Torque, and wrote custom frameworks for distributing jobs in pipelines, including interacting concurrently with PostgreSQL. Designed schemas and wrote numerous scripts for ETL (extract-transform-load) of third-party bioinformatics data (e.g., UniProt) and computational results (e.g., from HMMER). Have extensive experience with bioinformatics databases such as UniProt, GO, Pfam, PDB,

etc. and computational tools such as BLAST, HMMER, Modeller, etc. Used
Excel and R for statistical analyses. Responsible for the day-to-day
supervision of one or two programmers at any given time.

3/04-9/06: Google, Inc., Mountain View, CA. Software Engineer.
   International Search Quality: Diacritical insensitivity; stemming
        Created and implemented novel algorithm for diacritical
        insensitivity and transliteration in a context of uncertainty as to
        the language of the user.
        Created and implemented novel algorithm for stemming in several
        languages.
   Google Book Search: Web analytics
        Measured, modeled, and analyzed user behavior on a website with a
        complex user interface. Designed and implemented dimensional
        model and infrastructure to support these analyses.
   Designed and implemented numerous programs and pipelines using MapReduce
   in C++. Designed and used many protocol buffers. Wrote complex sawzall
   scripts.

1/04-3/04: Lawrence Berkeley Laboratory, Berkeley, CA.
   Worked on web interface to software for multi-precision computation.

5/01-8/02: Electronics Research Laboratory & Institute for Transportation
   Studies, UC Berkeley. (No relationship with degree research.) Wrote
   Linux kernel module implementing Wireless Token Ring Protocol.
   Designed cross-platform modular interface for link layer protocols.
   Ported Wireless Token Ring Protocol to QNX RTOS to run on automated
   vehicles. Did research on ad-hoc service networks for the MICA
   project.

1/96-5/98: SRC Systems, Inc., Berkeley, CA. Software Engineer. Primarily
   responsible for maintaining and adding functionality to commercial
   software application for thermal analysis of buildings. Implemented
   interactive persistent query of proprietary database, interactive
   graphs, import of foreign proprietary databases; wrote detailed
   proposal for converting proprietary database to standard SQL
   database.

Summer 1995: Hewlett-Packard Labs, Palo Alto, CA. Designed and began
   implementation of a spreadsheet interface for 3D modelling in C++,
   running in X windows, and also using Tcl/Tk and the OSF Motif widget
   library.

Summer 1994: Lawrence Berkeley Laboratory, Berkeley, CA. Wrote an OO

program in Symantec C++, using the Think Class Library for the
Macintosh, to interactively generate 2-dimensional Voronoi diagrams.

6/93-11/93: Lockheed Missiles and Science R&D Div., Palo Alto, CA. Wrote
a program in C to reconstruct an object from X-ray images taken in a
circle around the object, using the Feldkamp algorithm.

Summer 1992: Dept. of Materials Science & Mineral Engineering, UC
Berkeley. (No relationship with degree research.) Wrote a program in
C to generate PostScript output showing a rectangular array of
2-dimensional "rocks" in successive stages of cracking. Wrote a
program in C using curses to display PostScript files on the monitor
(using Display PostScript) with some user interaction. Wrote a
program in C using GL on an IBM RISC 6000 running AIX 3.1 to randomly
generate discs of various radii within a rectangular space such that
each disc touches at least one other.

Summer 1991: Woodward-Clyde Consultants, Pasadena, CA. Wrote software in
C to calculate theoretical seismograms due to seismic waves
propagating in a two-dimensional medium, using Kirchhoff integration.

Summer 1990: AT&T Bell Labs, Middletown, NJ. Translated the
specifications of the ASAI communications protocol into a particular
format, and specified particular test cases, in order to generate a
conformance test of a particular PBX switch to the protocol. Began
implementation of the test in C using a PC running UNIX as the
testbed.

Summer 1989: AT&T Bell Labs, Murray Hill, NJ. Developed a software
product in C which input a description of the user's implementation
of the boundary-scan hardware architecture, simulated the
implementation, and generated a test of conformance of the
implementation to the boundary-scan standard. This product was
subsequently sold by AT&T to hardware manufacturers under the trade
name TAPDANCE.

Summer 1988: AT&T Bell Labs, Denver, CO. Developed software in C to
decode several protocols within the OSI layered architecture for a
passive protocol monitor.

Summer 1987: AT&T Information Systems, Denver, CO. Developed software in
C to audit a fifty-field database.

Programming Languages

```
    C++, Python, R, Java, C, ML, Pascal (inc. Delphi), Basic (inc. Visual
    Basic, with Access), Perl, Tcl, PostScript;
    amateur interest in Haskell, Scheme, Prolog

  Window Systems
    Microsoft Windows, X Window, Macintosh

  Operating Systems
    Linux, Windows, Unix workstations, MacOS, DOS
```

Education

Ph.D in mathematics, University of California, Berkeley, 2003.
  Thesis title: "Algebraic Methods in Game Theory"
  Faculty advisor: Prof. Bernd Sturmfels.
M.S. in computer science, University of California, Berkeley, 2002.
  Thesis title: "Using Computer Algebra To Compute Nash Equilibria"
  Faculty advisor: Prof. Richard Fateman.
B.S. in mathematics, California Institute of Technology, 1991.
  Coursework equivalent to physics minor.

Patents

  Inventors: Datta, Ruchira S. Augmenting queries with synonyms from synonyms
    map, U.S. Patent 8255376.

  Inventors: Datta, Ruchira S. Simplifying query terms with transliteration,
    U.S. Patent 7835903.

  Inventors: Datta, Ruchira S., and Lopiano, Fabio. Augmenting queries with
    synonyms selected using language statistics, U.S. Patent 7475063.

Published Patent Application

  Inventors: Lopiano, Fabio, and Datta, Ruchira S. Query language
    determination using query terms and interface language, U.S. Patent
    Application 20070288450.

Teaching Experience

  2004: Lecturer, Dept. of Math, UC Davis
    Winter 2004: Integral Calculus; Short Calculus
  1994-1995: Graduate Student Instructor, Dept. of Math, UC Berkeley
    Fall 1994, Spring & Fall 1995: Linear Algebra & Differential Equations

Spring 1994: Multivariable Calculus

Foreign Languages

   French, Bengali, Russian, Spanish

References

   Upon request.