

COMMUNICATIONS

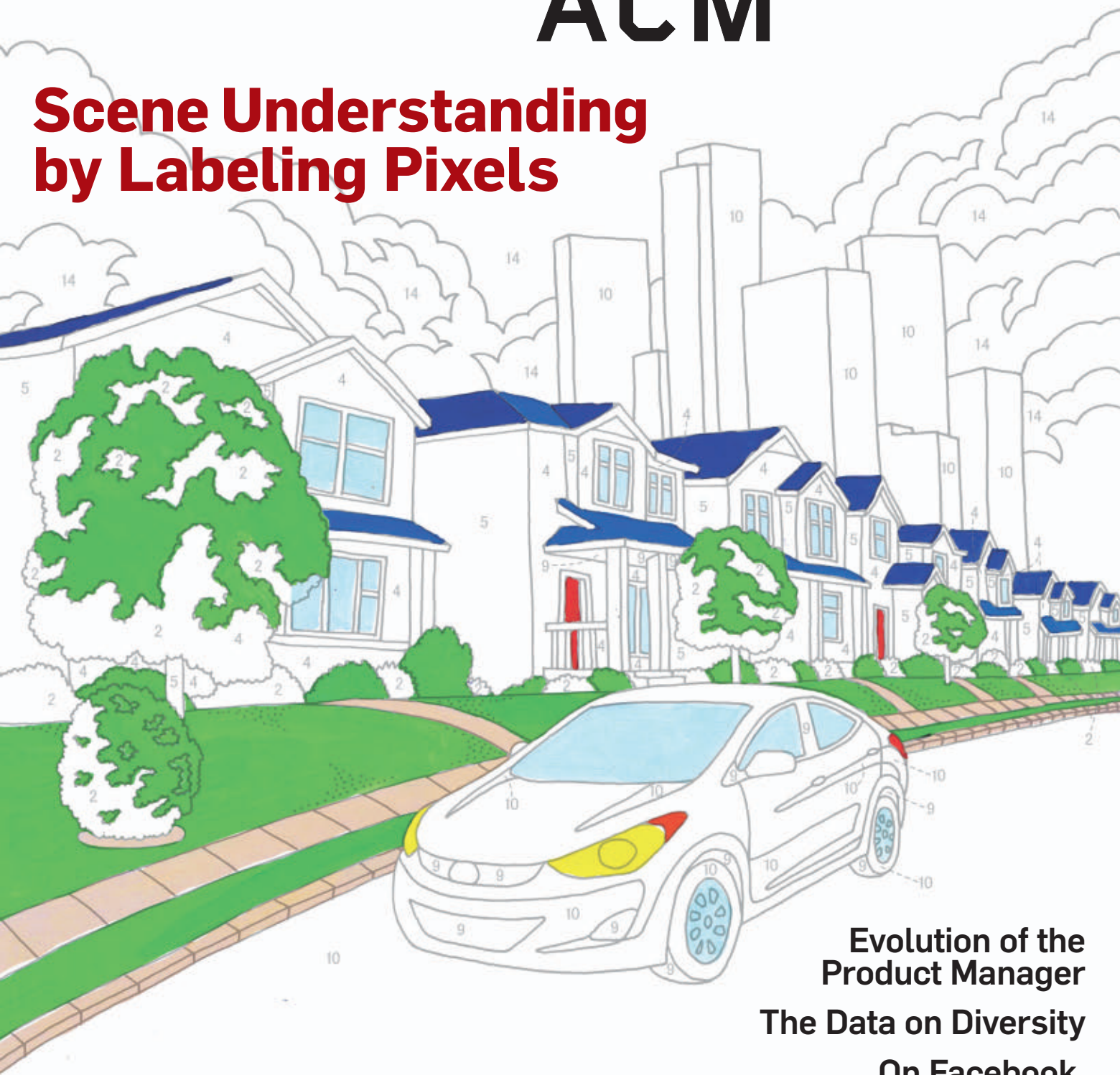
CACM.ACM.ORG

OF THE

ACM

11/2014 VOL.57 NO.11

Scene Understanding by Labeling Pixels



**Evolution of the
Product Manager**

The Data on Diversity

**On Facebook,
Most Ties Are Weak**

**Keeping Online
Reviews Honest**



2015

ACM INTERNATIONAL CONFERENCE
ON INTERACTIVE EXPERIENCES FOR
TELEVISION AND ONLINE VIDEO

3-5 JUNE, 2015
BRUSSELS, BELGIUM

Course and Workshop
proposals by
15 November 2014

Paper Submissions by
12 January 2015

Work in Progress, Demos,
DC, & Industrial
Submissions by
2 March 2015

Welcoming Submissions on

Content Production

Systems & Infrastructures

Devices & Interaction Techniques

Experience Design & Evaluation

Media Studies

Data Science & Recommendations

Business Models & Marketing

Innovative Concepts & Media Art



Association for
Computing Machinery



iMinds

TVX2015.COM

INFO@TVX2015.COM



ACM Books



MORGAN & CLAYPOOL
PUBLISHERS

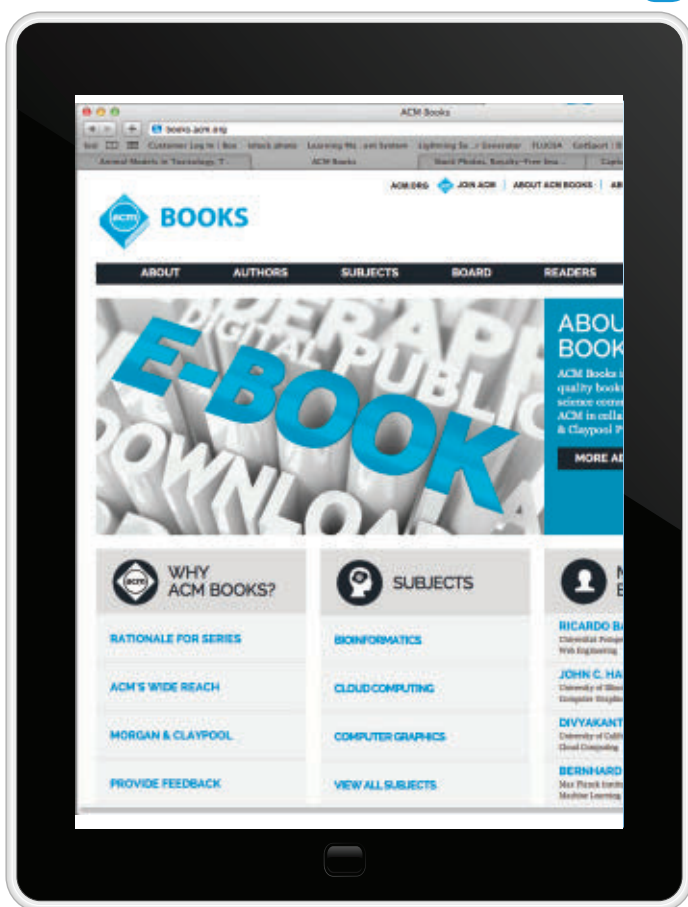
Publish your next book in the ACM Digital Library

ACM Books is a new series of advanced level books for the computer science community, published by ACM in collaboration with Morgan & Claypool Publishers.

I'm pleased that ACM Books is directed by a volunteer organization headed by a dynamic, informed, energetic, visionary Editor-in-Chief (Tamer Özsu), working closely with a forward-looking publisher (Morgan and Claypool).

—Richard Snodgrass, University of Arizona

books.acm.org **ACM Books**



- ◆ will include books from across the entire spectrum of computer science subject matter and will appeal to computing practitioners, researchers, educators, and students.
- ◆ will publish graduate level texts; research monographs/overviews of established and emerging fields; practitioner-level professional books; and books devoted to the history and social impact of computing.
- ◆ will be quickly and attractively published as ebooks and print volumes at affordable prices, and widely distributed in both print and digital formats through booksellers and to libraries and individual ACM members via the ACM Digital Library platform.
- ◆ is led by EIC M. Tamer Özsu, University of Waterloo, and a distinguished editorial board representing most areas of CS.

Proposals and inquiries welcome!

Contact: **M. Tamer Özsu**, Editor in Chief
booksubmissions@acm.org



Association for
Computing Machinery

Advancing Computing as a Science & Profession

Departments

- 5 **From the President**
Dealing with the Deep, Long-Term Challenges Facing ACM (Part I)
By Alexander L. Wolf
-
- 7 **Cerf's Up**
Heidelberg Laureate Forum II
By Vinton G. Cerf
-
- 9 **Letters to the Editor**
Accountability Is No Excuse for Surveillance
-
- 10 **BLOG@CACM**
Teach the Teachers, and Contribute to Humanity
Mark Guzdial suggests computer science education needs to be aimed at adults before teaching it to children; Lawrence M. Fisher finds the volunteer spirit behind ACM Special Interest Groups.
-
- 41 **Calendar**
-
- 106 **Careers**

Last Byte

- 120 **Upstart Puzzles**
Proving without Teaching/ Teaching without Proving
By Dennis Shasha

News



- 13 **Researchers Simplify Parallel Programming**
Parallel computing has become increasingly important as chipmakers put more and more processor cores on individual chips.
By Gary Anthes
-
- 16 **Computing What Fits**
New apps and pods improve the virtual and actual retail experiences by ensuring clothing and cosmetics look good on you before you buy them.
By Esther Shein
-
- 20 **Keeping Online Reviews Honest**
As online reviews become increasingly important to e-commerce and social media sites, computer scientists work on ways to ensure their authenticity.
By Logan Kugler

Viewpoints

- 24 **Economic and Business Dimensions**
The Dark Side of the Sharing Economy ... and How to Lighten It
Improving the sharing economy will require addressing myriad problems.
By Arvind Malhotra and Marshall Van Alstyne
-
- 28 **Legally Speaking**
Updates on the Intellectual Property Front
Recapping influential U.S. Supreme Court decisions rendered earlier this year.
By Pamela Samuelson
-
- 31 **Computing Ethics**
Big Data's End Run Around Procedural Privacy Protections
Recognizing the inherent limitations of consent and anonymity.
By Solon Barocas and Helen Nissenbaum
-
- 34 **Education**
A Future for Computing Education Research
Seeking to address the most important issues facing the computer education research community.
By Steve Cooper, Shuchi Grover, Mark Guzdial, and Beth Simon
-
- 37 **Privacy and Security**
Summing Up
Considering recent privacy- and security-related events through this column.
By Susan Landau
-
- 40 **Viewpoint**
An Embarrassment of Riches
A critical review of open innovation systems.
By Mark Klein and Gregorio Convertino
-
- 43 **Viewpoint**
Designing User Incentives for Cybersecurity
How to encourage better user security practices and behavior.
By Terrence August, Robert August, and Hyoduk Shin



Association for Computing Machinery
Advancing Computing as a Science & Profession

Practice



- 48 **Evolution of the Product Manager**
Better education needed to develop the discipline.
By Ellen Chisa
-
- 53 **JavaScript and the Netflix User Interface**
Conditional dependency resolution.
By Alex Liu
-
- 60 **A Decade of Progress in Parallel Programming Productivity**
Looking at the design and benefits of X10.
By John T. Richards, Jonathan Brezin, Calvin B. Swart, and Christine A. Halverson

Q Articles' development led by **acmqueue**
queue.acm.org

Contributed Articles



- 68 **Scene Understanding by Labeling Pixels**
Pixels labeled with a scene's semantics and geometry let computers describe what they see.
By Stephen Gould and Xuming He
-
- 78 **On Facebook, Most Ties Are Weak**
Strong ties connect individuals in the same community; weak ties connect individuals in different communities.
By Pasquale De Meo, Emilio Ferrara, Giacomo Fiumara, and Alessandro Provetti



About the Cover:
Creating a computer that can automatically understand a visual scene has been a perennial problem in the AI and computer vision fields. This month's cover story (p. 68) presents a promising approach that calls for annotating every pixel in a scene with a label that indicates what and where it is. Cover illustration by Andrew Joyce.

Review Articles



- 86 **The Data on Diversity**
It's not just about being fair.
By Beryl Nelson

Research Highlights

- 97 **Technical Perspective**
The Intricate Dance of Fabric and Light
By Szymon Rusinkiewicz
-
- 98 **Building Volumetric Appearance Models of Fabric Using Micro CT Imaging**
By Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala



ACM, the world's largest educational and scientific computing society, delivers resources that advance computing as a science and profession. ACM provides the computing field's premier Digital Library and serves its members and the computing profession with leading-edge publications, conferences, and career resources.

Executive Director and CEO

John White

Deputy Executive Director and COO

Patricia Ryan

Director, Office of Information Systems

Wayne Graves

Director, Office of Financial Services

Darren Ramdin

Director, Office of SIG Services

Donna Cappel

Director, Office of Publications

Bernard Rous

Director, Office of Group Publishing

Scott E. Delman

ACM COUNCIL

President

Alexander L. Wolf

Vice-President

Vicki L. Hanson

Secretary/Treasurer

Erik Altman

Past President

Vinton G. Cerf

Chair, SGB Board

Patrick Manning

Co-Chairs, Publications Board

Jack Davidson and Joseph Konstan

Members-at-Large

Eric Allman; Ricardo Baeza-Yates;

Cherri Pancake; Radia Perlman;

Mary Lou Soffa; Eugene Spafford;

Per Stenström

SGB Council Representatives

Paul Beame; Barbara Boucher Owens;

Andrew Sears

BOARD CHAIRS

Education Board

Mehran Sahami and Jane Chu Prey

Practitioners Board

George Neville-Neil

REGIONAL COUNCIL CHAIRS

ACM Europe Council

Fabrizio Gagliardi

ACM India Council

Srinivas Padmanabhuni

ACM China Council

Jianguang Sun

PUBLICATIONS BOARD

Co-Chairs

Jack Davidson; Joseph Konstan

Board Members

Ronald F. Boisvert; Marie-Paule Cani;

Nikil Dutt; Roch Guerrin; Carol Hutchins;

Patrick Madden; Catherine McGeoch;

M. Tamer Ozsu; Mary Lou Soffa

ACM U.S. Public Policy Office

Renee Dopplick, Acting Director

1828 L Street, N.W., Suite 800

Washington, DC 20036 USA

T (202) 659-9711; F (202) 667-1066

Computer Science Teachers Association

Lissa Clayborn, Acting Executive Director

COMMUNICATIONS OF THE ACM

Trusted insights for computing's leading professionals.

Communications of the ACM is the leading monthly print and online magazine for the computing and information technology fields. *Communications* is recognized as the most trusted and knowledgeable source of industry information for today's computing professional. *Communications* brings its readership in-depth coverage of emerging areas of computer science, new trends in information technology, and practical applications. Industry leaders use *Communications* as a platform to present and debate various technology implications, public policies, engineering challenges, and market trends. The prestige and unmatched reputation that *Communications of the ACM* enjoys today is built upon a 50-year commitment to high-quality editorial content and a steadfast dedication to advancing the arts, sciences, and applications of information technology.

STAFF

DIRECTOR OF GROUP PUBLISHING

Scott E. Delman

publisher@cacm.acm.org

Executive Editor

Diane Crawford

Managing Editor

Thomas E. Lambert

Senior Editor

Andrew Rosenbloom

Senior Editor/News

Larry Fisher

Web Editor

David Roman

Editorial Assistant

Zarina Strakhan

Rights and Permissions

Deborah Cotton

Art Director

Andrij Borys

Associate Art Director

Margaret Gray

Assistant Art Director

Mia Angelica Balaquiot

Designer

Iwona Usakiewicz

Production Manager

Lynn D'Addesio

Director of Media Sales

Jennifer Ruzicka

Public Relations Coordinator

Virginia Gold

Publications Assistant

Juliet Chance

Columnists

David Anderson; Phillip G. Armour;

Michael Cusumano; Peter J. Denning;

Mark Guzdial; Thomas Haigh;

Leah Hoffmann; Mari Sako;

Pamela Samuelson; Marshall Van Alstyne

CONTACT POINTS

Copyright permission

permissions@cacm.acm.org

Calendar items

calendar@cacm.acm.org

Change of address

acmhlp@cacm.acm.org

Letters to the Editor

letters@cacm.acm.org

WEBSITE

<http://cacm.acm.org>

AUTHOR GUIDELINES

<http://cacm.acm.org/guidelines>

ACM ADVERTISING DEPARTMENT

2 Penn Plaza, Suite 701, New York, NY

10121-0701

T (212) 626-0686

F (212) 869-0481

Director of Media Sales

Jennifer Ruzicka

jen.ruzicka@hq.acm.org

Media Kit acmmidasales@acm.org

Association for Computing Machinery (ACM)

2 Penn Plaza, Suite 701

New York, NY 10121-0701 USA

T (212) 869-7440; F (212) 869-0481

EDITORIAL BOARD

EDITOR-IN-CHIEF

Moshe Y. Vardi

eic@cacm.acm.org

NEWS

Co-Chairs

William Pullleyblank and Marc Snir

Board Members

Mei Kobayashi; Kurt Mehthorn;

Michael Mitzenmacher; Rajeev Rastogi

VIEWPOINTS

Co-Chairs

Tim Finin; Susanne E. Hambrusch;

John Leslie King

Board Members

William Aspray; Stefan Bechtold;

Michael L. Best; Judith Bishop;

Stuart I. Feldman; Peter Freeman;

Mark Guzdial; Rachelle Hollander;

Richard Ladner; Susan Landau;

Carl Landwehr; Carlos Jose Pereira de Lucena;

Beng Chin Ooi; Loren Terveen;

Marshall Van Alstyne; Jeannette Wing



PRACTICE

Co-Chairs

Stephen Bourne

Board Members

Eric Allman; Charles Beeler; Bryan Cantrill;

Terry Coatta; Stuart Feldman; Benjamin Fried;

Pat Hanrahan; Tom Limoncelli;

Kate Matsudaira; Marshall Kirk McKusick;

Erik Meijer; George Neville-Neil;

Theo Schlossnagle; Jim Waldo

The Practice section of the CACM

Editorial Board also serves as

the Editorial Board of *Queue*.

CONTRIBUTED ARTICLES

Co-Chairs

Al Aho and Andrew Chien

Board Members

William Aiello; Robert Austin; Elisa Bertino;

Gilles Brassard; Kim Bruce; Alan Bundy;

Peter Buneman; Peter Druschel;

Carlo Ghezzi; Carl Gutwin; Gal A. Kaminka;

James Larus; Igor Markov; Gail C. Murphy;

Shree Nayar; Bernhard Nebel;

Lionel M. Ni; Kenton O'Hara;

Sriram Rajamani; Marie-Christine Rousset;

Avi Rubin; Krishan Sabnani;

Ron Shamir; Yoav Shoham; Larry Snyder;

Michael Vitale; Wolfgang Wahlster;

Hannes Werthner; Reinhard Wilhelm

RESEARCH HIGHLIGHTS

Co-Chairs

Azer Bestavros and Gregory Morrisett

Board Members

Martin Abadi; Amir El Abbadi; Sanjeev Arora;

Dan Boneh; Andrei Broder; Stuart K. Card;

Jeff Chase; Jon Crowcroft; Matt Dwyer;

Alon Halevy; Maurice Herlihy; Norm Jouppi;

Andrew B. Kahng; Xavier Leroy; Kobbi Nissim;

Mendel Rosenblum; David Salesin;

Steve Seitz; Guy Steele, Jr.; David Wagner;

Margaret H. Wright

WEB

Chair

James Landay

Board Members

Marti Hearst; Jason I. Hong;

Jeff Johnson; Wendy E. MacKay

ACM Copyright Notice

Copyright © 2014 by Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or fee. Request permission to publish from permissions.acm.org or fax (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page or screen display, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center; www.copyright.com.

Subscriptions

An annual subscription cost is included in ACM member dues of \$99 (\$40 of which is allocated to a subscription to *Communications*); for students, cost is included in \$42 dues (\$20 of which is allocated to a *Communications* subscription). A nonmember annual subscription is \$100.

ACM Media Advertising Policy

Communications of the ACM and other ACM Media publications accept advertising in both print and electronic formats. All advertising in ACM Media publications is at the discretion of ACM and is intended to provide financial support for the various activities and services for ACM members. Current Advertising Rates can be found by visiting <http://www.acm-media.org> or by contacting ACM Media Sales at (212) 626-0686.

Single Copies

Single copies of *Communications of the ACM* are available for purchase. Please contact acmhlp@acm.org.

COMMUNICATIONS OF THE ACM

(ISSN 0001-0782) is published monthly by ACM Media, 2 Penn Plaza, Suite 701, New York, NY 10121-0701. Periodicals postage paid at New York, NY 10001, and other mailing offices.

POSTMASTER

Please send address changes to *Communications of the ACM*
2 Penn Plaza, Suite 701
New York, NY 10121-0701 USA

Printed in the U.S.A.



Association for
Computing Machinery



Dealing with the Deep, Long-Term Challenges Facing ACM (Part I)

IN AN EDITORIAL last month, ACM CEO John White reported on the outcomes of the November 2013 ACM Strategic Planning Retreat. The retreat generated a number of important new ideas in the areas of membership, conferences, publications, community, and practitioners. While they are likely to have a significant impact on ACM's activities going forward, I view them as defining only a relatively short-term agenda not fully addressing some deep issues facing ACM, issues we still need to understand and deal with.

The challenges and opportunities of open access served as the original motivation for holding the retreat. Despite setting the modern standard for a liberal copyright policy in the digital age (over a decade ago), opening more content under the discretion of ACM authors and SIGs, fully embracing Green/Gold/Hybrid OA publishing, and complying with government mandates, there is a sense among a portion of our community that we have still not done enough. A sense that if an ACM publication sits behind any sort of paywall—regardless of it also being *freely available* via an author's site, an author's institutional site, a SIG site, or even a conference site—we are somehow failing to meet our commitment to nurture the free flow of information.

This position informed much of the retreat discussion. It led to our decision to encourage SIGs to open conference proceedings around the event and until its next occurrence, and it led to our decision to begin work on understanding the “article of the future,” in the broadest sense, along with the digital collection that might serve as host.

But what it failed to do is lead us into a discussion of whether there can (even should be) a model for ACM in which publication revenue plays little or no role.

ACM has a pretty straightforward

business model. There are three major revenue streams: membership dues, conference registration fees, and publication subscription fees. Each has related sets of expenses. Membership runs at a loss—mainly because we subsidize students and members from developing countries. Conferences typically run at a surplus (not always—some conferences have lost considerable sums), but that surplus is retained by the SIGs and invested directly into serving and subsidizing their respective technical communities however they see fit. That leaves publications.

Right now, publications overall (but not universally) generate a surplus. That surplus is used to underwrite the membership loss associated with subsidies for students and developing countries, and to significantly supplement the operating funds flowing back to the SIGs.

It is also used to underwrite the programs our volunteers and members want, build, and run—but for which there is no revenue. This set of programs (sometimes referred to as the “good works” of ACM) includes: ACM-W—our 20-year effort to support and see more women succeed in computing; the Committee to Diversify Computing and the Tapia conference—activities focused on broadening participation in computing; CSTA—the organization for primary and secondary school computing teachers; the Education Policy Committee that advocates for policy changes to see real computer science taught in secondary schools; significant support for the Computing Research Association; USACM—the volunteer committee focused on six major areas of technology policy; ACM Europe/China/India—our efforts to increase our relevance in these regions by giving a special voice, visibility, and autonomy to our members there; the Education Board and its five-decade effort to develop and

maintain international curriculum standards for computer science; Computer Science Education Week and Code.org—opening computing to millions of people around the world; and the Practitioner Board and its development of *Queue* and the Practice section of *Communications*.

These programs exist not because some abstract corporate entity (“the ACM”) built them, but because members *wanted* them, and in many cases *demand*ed they exist. And they thrive because the community is committed to funding the enormous work needed for each.

What we did not do at the retreat is look at a future ACM model in which there is no publication revenue and ask ourselves: “How do we the community do what we the community wants done having only membership dues and conference registration fees to support us?” But when segments of the community ask that ACM stop generating revenue from publications, we need to address this question. Of course, addressing the question will mean taking a fundamental look at membership structures and a fundamental look at how ACM conferences should be run. We need to face the fact that reducing or eliminating publication revenue is a choice that would have profound consequences for the *entire* ACM community.

As president, I am committed to engaging these hard questions; as a first step I will be establishing a Presidential Task Force to take a deep look specifically at future, long-term models for ACM. I also seek your personal input on the difficult choices we may need to make. We cannot simplistically engage in single-issue debates. I ask your help in finding a way forward that supports the broad interests and needs of all our members. 

Alexander L. Wolf, ACM PRESIDENT

Copyright held by author.

Inviting Young Scientists



Association for
Computing Machinery

Meet Great Minds in Computer Science and Mathematics

As one of the founding organizations of the Heidelberg Laureate Forum <http://www.heidelberg-laureate-forum.org/>, ACM invites young computer science and mathematics researchers to meet some of the preeminent scientists in their field. These may be the very pioneering researchers who sparked your passion for research in computer science and/or mathematics.

These laureates include recipients of the ACM A.M. Turing Award, the Abel Prize, and the Fields Medal.

The Heidelberg Laureate Forum is **August 23–28, 2015** in Heidelberg, Germany.

This week-long event features presentations, workshops, panel discussions, and social events focusing on scientific inspiration and exchange among laureates and young scientists.

Who can participate?

New and recent Ph.Ds, doctoral candidates, other graduate students pursuing research, and undergraduate students with solid research experience and a commitment to computing research

How to apply:

Online: <https://application.heidelberg-laureate-forum.org/>
Materials to complete applications are listed on the site.

What is the schedule?

Application deadline—**February 28, 2015**.

We reserve the right to close the application website early depending on the volume

Successful applicants will be notified by **April 15, 2015**.

More information available on Heidelberg social media





DOI:10.1145/2674716

Vinton G. Cerf

Heidelberg Laureate Forum II

The second Heidelberg Laureate Forum (HLF) has just ended. Taking place annually in this historic city, this conference was even better than the first, if that is possible.

The primary mover, Klaus Tschira (a founder of the German software firm SAP), has again generously supported this weeklong event through his personal foundation. ACM (Turing Award), the International Mathematics Union (Fields Medal and Nevanlinna Prize) and the Norwegian Academy of Science and Letters (Abel Prize) are additional sponsors and organize the selection of the 200 young researchers who join the laureates for the week. There were plenty of opportunities for formal lectures and informal small group gatherings to link the 100 mathematics and 100 computer science researchers to each other and the laureates in attendance. The HLF staff provided superb planning, operational and logistical support throughout this event, adapting in real time to suggestions for improvement.

It would be difficult to overstate the energy levels reached at this event. If we had been arranged in a circle, we might have outdone the Large Hadron Collider at CERN! As usual, I learned many new things from both the mathematicians and the computer scientists. While I have only limited space in this column, I would like to share some of the interesting surprises that came my way.

Ivan Sutherland was in attendance again this year and while discussing computer architecture using asynchronous circuits, he told me about a very interesting idea that has implications for improving software safety. We all know a register of N bits has 2^N values, $0-2^N-1$. Right? Wrong! Turns

out, as Ivan grinningly reported, there is another value, “empty.” That is, if nothing has been stored in the register, the hardware should produce an exception to any attempt to make use of the (non)value in that register. Flagging such an event should reduce the incidence of using the contents of an uninitiated register that could lead to any of a variety of bad effects. This begs the question, “When does the ‘empty’ flag get set?” It is obvious that it must be reset when something is stored in the register. One might set the bit when a “new” register is created in a push-down stack, for instance. Or when memory is returned to the “free pool” in a garbage-collecting system. When a register is popped off the push-down stack it should be treated as empty of any value. This idea reinforces the more general sense that safety and security of programs might be improved by creating a partnership in the form of hardware-reinforced security or safety measures.

Leslie Lamport, recipient of the 2013 Turing award, gave one of the lectures. He spoke eloquently on the topic of proofs (not just of program correctness but of any proofs). He made a strong and vivid case for changing from the 17th-century style of narrative proofs to a 21st-century style in which it is clear at each step what precedents are being invoked for each step in the proof. The use of a terse symbolic representation versus a narrative made the basis for each assertion far more apparent.

John Hopcroft gave an example-rich talk illustrating how different his

lectures on computer science are today compared to only 15 years ago. Today, he is far more focused on reasoning and computing in high-dimensionality data rather than operating systems, programming languages, and data structures. What completely surprised me was the reasoning that leads to the conclusion that the unit sphere in extremely high-dimensionality spaces has vanishing volume. There were surprises like this laced throughout an entertaining but very thought-provoking examination of the modern computer science curriculum as he sees it.

A visit to the European Molecular Biology Laboratory (EMBL) on the outskirts of Heidelberg proved to be very stimulating. Surrounded by energetic and thoughtful mathematics and computer science graduates, post-docs, and junior faculty, we heard a remarkable lecture about what is known about the complex operation of cells and the particularly interesting way in which microtubules are used to guide the transport of complex biochemicals within the cell. Computational biology was “in the air” at EMBL and the HLF. If anything stood out for me, it was the increasing role of computing and mathematics in diverse subjects including astrophysics, biochemistry, and mathematical topology among so many other modern interests.

The Heidelberg Laureate Forum is one of the best new activities in which ACM is engaged and I predict it will become a strong and new tradition among the many ways in which ACM contributes to and reinforces the importance of computing in our society. 

Vinton G. Cerf is VP and Chief Internet Evangelist at Google. He served as ACM president from 2012–2014.

Copyright held by author

**Previous
A.M. Turing Award
Recipients**

1966 A.J. Perlis
1967 Maurice Wilkes
1968 R.W. Hamming
1969 Marvin Minsky
1970 J.H. Wilkinson
1971 John McCarthy
1972 E.W. Dijkstra
1973 Charles Bachman
1974 Donald Knuth
1975 Allen Newell
1975 Herbert Simon
1976 Michael Rabin
1976 Dana Scott
1977 John Backus
1978 Robert Floyd
1979 Kenneth Iverson
1980 C.A.R Hoare
1981 Edgar Codd
1982 Stephen Cook
1983 Ken Thompson
1983 Dennis Ritchie
1984 Niklaus Wirth
1985 Richard Karp
1986 John Hopcroft
1986 Robert Tarjan
1987 John Cocke
1988 Ivan Sutherland
1989 William Kahan
1990 Fernando Corbató
1991 Robin Milner
1992 Butler Lampson
1993 Juris Hartmanis
1993 Richard Stearns
1994 Edward Feigenbaum
1994 Raj Reddy
1995 Manuel Blum
1996 Amir Pnueli
1997 Douglas Engelbart
1998 James Gray
1999 Frederick Brooks
2000 Andrew Yao
2001 Ole-Johan Dahl
2001 Kristen Nygaard
2002 Leonard Adleman
2002 Ronald Rivest
2002 Adi Shamir
2003 Alan Kay
2004 Vinton Cerf
2004 Robert Kahn
2005 Peter Naur
2006 Frances E. Allen
2007 Edmund Clarke
2007 E. Allen Emerson
2007 Joseph Sifakis
2008 Barbara Liskov
2009 Charles P. Thacker
2010 Leslie G. Valiant
2011 Judea Pearl
2012 Shafi Goldwasser
2012 Silvio Micali
2013 Leslie Lamport

ACM A.M. TURING AWARD NOMINATIONS SOLICITED

Nominations are invited for the 2014 ACM A.M. Turing Award. This, ACM's oldest and most prestigious award, is presented for contributions of a technical nature to the computing community. Although the long-term influences of the nominee's work are taken into consideration, there should be a particular outstanding and trendsetting technical achievement that constitutes the principal claim to the award. The recipient presents an address at an ACM event that will be published in an ACM journal.

Nominations should include:

- 1) A curriculum vitae, listing publications, patents, honors, and other awards.
- 2) A letter from the principal nominator, which describes the work of the nominee, and draws particular attention to the contribution which is seen as meriting the award.
- 3) Supporting letters from at least three endorsers. The letters should not all be from colleagues or co-workers who are closely associated with the nominee, and preferably should come from individuals at more than one organization. Successful Turing Award nominations should include substantive letters of support from prominent individuals broadly representative of the candidate's field or related field impacted by the candidate's contribution.

**For additional information on ACM's award program
please visit: www.acm.org/awards/**

**Additional information on the past recipients
of the A.M. Turing Award is available on:
<http://amturing.acm.org/byyear.cfm>.**

**Nominations should be sent electronically
by November 30, 2014 to:
[Barbara Liskov, MIT CSAIL c/o mcguinness@acm.org](mailto:Barbara.Liskov@MIT.CSAIL)**



Association for
Computing Machinery

Accountability Is No Excuse for Surveillance

BRITISH PHILOSOPHER AND social theorist Jeremy Bentham would have wholeheartedly endorsed many of the accountability mechanisms Stefan Bechtold and Adrian Perrig outlined in their Viewpoint “Accountability in Future Internet Architectures” (Sept. 2014). It reminded me of Bentham’s Panopticon (late 18th century), a prison where the prisoners would be motivated to behave in a more civilized manner by being made to think they were always under surveillance. Likewise, Bechtold and Perrig took the view that network users being tracked and made accountable for their actions would improve the Internet.

I am certain the majority of governments today would endorse this architecture, in which it would be possible to trace all Internet Protocol communication packets from source to destination and guarantee everyone is using the network responsibly. Indeed, many governments already pursue such a goal.

On the other hand, I am concerned the pervasive monitoring already present in today’s global Internet without these technical aids might not be in society’s best interests. I have been working with U.S. State Dept. sponsorship aiding a user group of journalists and democracy advocates in African countries, many with authoritarian tendencies. I am developing anonymization tools and training participants to use them. In many of the countries, accountability for accessing information considered innocuous in the West has dire consequences. Many of those lacking human rights protections found in Western democracies indeed use the technology produced in the West.

I dislike the idea of making people accountable for the information they consume, which would be a by-product of the ideas Bechtold and Perrig proposed.

Richard R. Brooks, Clemson, SC

Authors’ Response:

Like Brooks, we strongly support privacy and anonymity for users. However, we also

strongly disagree with an interpretation of our Viewpoint that says we envision a future Internet architecture that tracks users. Our aim was (and is) more discerning. As we pointed out, it is sometimes possible to achieve both privacy and accountability, whereby users maintain their privacy and become accountable only if they violate some policy as by, say, perpetrating an attack. Moreover, anonymity can be achieved through an overlay network, even if the underlying network is accountable. We also highlighted the research challenges involved in balancing accountability, privacy, anonymity, political freedom, and other values. Brooks seems to have missed this core point.

Stefan Bechtold and Adrian Perrig,
Zurich, Switzerland

To Counter Cybercrime, Try Security Diligence

Cormac Herley’s article “Security, Cybercrime, and Scale” (Sept. 2014) focused on logical analysis of narrowly defined financial cybercrimes gainfully performed by untrusted remote perpetrators, not by embezzlers. The objective Herley specified in this logical model is improved security to reduce the risk of rational financially motivated untrusted perpetrators able to carry out all possible scaled attacks.

Having interviewed more than 200 cybercrime perpetrators over the past 40 years, I suggest reality is quite different. First, perpetrators possess only partial knowledge. They also cause errors that change their objectives, take less financial assets than are available, do not necessarily consider cost, perform copy-

cat attacks, and act under many other personal irrational conditions and circumstances that were always present in all of the cases I studied.

Here is my threat model: Alice knows she cannot be sufficiently secure from attacks by Mallory and thus seeks to avoid negligence after Mallory (inevitably) attacks, successfully or not.

Herley correctly noted the limitations of successful risk reduction, but a different objective and strategy are more desirable for my model. The objective I advocate is security diligence, rather than risk reduction. It is a safer, more easily obtained and measured objective for the enterprise, more likely meets insurance requirements, and reduces a broader range of risk reduction that includes possibly reducing the risk of negligence on the part of the victim enterprise and the stakeholders within it. I have found (in practice) this is often more important than financial loss.

The diligence strategy is to implement security controls by engaging in benchmark studies, using standards, compliance, contracts, audits, good practices, available products, cost control, experts’ opinions, and experimentation. The tough high-cost decisions are made by management fiat, not necessarily by risk reduction.

Donn B. Parker, Los Altos, CA

Correction

The Milestones section “Computer Science Awards, Appointments” (Sept. 2014) reported Jack Dongarra of the University of Tennessee, Knoxville, as the recipient of the ACM-IEEE Computer Society Ken Kennedy Award. Allow us to clarify. Dongarra was the 2013 recipient. Charles E. Leiserson of the Massachusetts Institute of Technology was recently named the recipient of the award for 2014 (see page 14).

Communications welcomes your opinion. To submit a Letter to the Editor, please limit yourself to 500 words or less, and send to letters@cacm.acm.org.

© 2014 ACM 0001-0782/14/11 \$15.00

I dislike the idea of making people accountable for the information they consume.

The *Communications* Web site, <http://cacm.acm.org>, features more than a dozen bloggers in the BLOG@CACM community. In each issue of *Communications*, we'll publish selected posts or excerpts.



Follow us on Twitter at <http://twitter.com/blogCACM>

DOI:10.1145/2667107

<http://cacm.acm.org/blogs/blog-cacm>

Teach the Teachers, and Contribute to Humanity

Mark Guzdial suggests computer science education needs to be aimed at adults before teaching it to children; Lawrence M. Fisher finds the volunteer spirit behind ACM Special Interest Groups.



Mark Guzdial
**If You Want to Teach
Kids CS, First
Teach the Adults**
<http://bit.ly/1wlGgUV>
July 14, 2014

Recently, ACM past president Vint Cerf testified to the U.S. Senate on research needs, and he addressed the need for more CS education:

As a recent president of the Association for Computing Machinery and a member of the Google staff, I have been a strong proponent of the proposition that computer science should be a required part of the K–12 curriculum. Every student should have some exposure to the concept of programming, not only because it promotes logical thinking but also because it is important for everyone to understand and appreciate the potential weaknesses in all software-controlled systems. Computer science should be treated on a par with biology, chemistry, physics, and mathematics in K–12 and undergraduate curricula,

not simply as an elective that bears no STEM credit.

—Vint Cerf, testimony to U.S. Senate Committee on Commerce, Science and Transportation, July 17, 2014

How do we achieve the goal Vint identifies, to *require* CS in primary and secondary school curricula in the U.S.?

We can agree we need schools and teachers to achieve this goal. MOOCs have not yet shown success with early subjects, and completion rates for MOOCs are lowest for students with little prior background. We really do not know how to use online learning to successfully engage the majority of secondary school students in CS education, let alone the greater challenges of reaching *all* students in high school and in primary school.

The challenges of getting more schools and more CS teachers are highlighted in a recent working paper by the Exploring CS (ECS) group at the University of California at Los Angeles (UCLA), *The Revolving Door: CS for All*

and the Challenge of Teacher Retention (<http://bit.ly/1rMVdm7>):

This means that fully 45 of the 81 teachers who have participated in the ECS program have experienced a teaching “disruption” which has ended their participation in the ECS teacher community for a year or longer.

What is worse is that losing a teacher often means losing the whole school—no replacement CS teacher can be found quickly enough to avoid losing a year or more. What is the problem? Why are so few teachers being retained as high school CS teachers in Los Angeles? Simply requiring or mandating CS in schools does not solve this problem of getting enough teachers.

To be clear, it is not a problem of more professional development. We have far more capacity for CS teacher professional development (<http://bit.ly/1lIgsmN>) than we have teachers interested in learning CS. We have to make teaching CS a valued and inviting career.

The ECS working paper cites Linda Darling-Hammond, who explains the real challenge of having enough teachers is getting support for those teachers:

*“The main problem is an exodus of new teachers from the profession, with more than 30% leaving within five years, and higher rates of turnover in lower-income schools. An additional problem is the flight of teachers from less-affluent schools to more-affluent schools. This is strongly tied to **working conditions**—including administrative support and strong colleagues as well as tangible teaching conditions and*

salaries. Research also finds that teachers leave the profession much faster if they have less preparation before they enter and less mentoring support when they arrive.”

In Lijun Ni’s research on CS teacher identity (<http://bit.ly/1qyJROS>), we heard similar stories from math teachers who did not want to teach computer science—it was the lack of community and support:

“I’m a better math teacher, just because I’ve had so much support. Whenever I have problems, I can talk with the people that I work with, most of who[m] have taught for many years in math. ...Every day, I’m eating lunch with math teachers. With computer science, I’ve got nobody to talk to.”

If we want to teach kids CS, we first have to teach adults CS. We need more teachers who know computer science. We need more administrators and principals who value CS and build support for it within their schools, including mentoring support for new teachers. Yes, we also need legislators to define education policy to value CS so teaching a science class does not take priority over teaching a CS class.

I propose a specific, concrete step to improve CS education and bring us closer to Vint Cerf’s goal: Require computer science as part of the general education requirements for all undergraduate majors in all universities and colleges. Undergraduate students would take courses in social sciences, humanities, science, mathematics, and computer science. If computer science is so important that all school children need it, then it is obviously even more important for all students who are choosing career paths that require higher education. Implementing this requirement is far easier than implementing a requirement in K–12, since most colleges and universities already have a CS department with teachers who know how to teach CS.

The positive impact of this step is that it makes long-term change. All new teachers will already have had a course in computer science, making professional development in how to *teach* computer science much easier. Administrators, principals, and parents with undergraduate degrees will have had a CS course that helps them understand the value and importance of computer science for children.

The ECS working paper ends with an important admonition. The desire to improve computing education in schools is going to lead to short-term quick fixes. An example quick fix is asking IT professionals to teach CS classes so that *something* gets offered in schools. That kind of quick fix does not do anything to grow support for teaching CS in schools so teachers are recruited, mentored, supported, and retained. That kind of quick fix does not lead to *change* in how we teach CS, so we broaden participation in computing and grow a more diverse CS workforce.

We need to plan for teaching computer science for the long term. Think in terms of decades (<http://bit.ly/1kF6O1u>). We have to create a system for growing CS in schools. To do that, we need to grow teachers and schools, which means we first need to think about teaching the *adults* about CS.

Comments

I agree adults should be taught for them to value and pass CS to the next generation.

—E. A. Ichu

As a 1989 college CS graduate I can attest to its value, though how I arrived at becoming interested in CS was not through a mandatory requirement from the small-town public school I attended. My interest began more as a curiosity of computers and how they worked. The truth is I spent many hours playing games on my Commodore 64, which led to eventually writing very simple programs for the C64. What I found as a curiosity, most kids today take for granted. If parents, teachers, and schools want to promote CS, they should recognize there are numerous channels in which kids can be invited to explore what CS is about. Introductory CS classes should focus on fun and multiple disciplines of CS as a way for the student to explore and develop their own curiosity.

—Paul Parker



Lawrence M. Fisher
Contributing to
Humanity

<http://bit.ly/1qcqGYI>

August 11, 2014

This is my first time attending SIGGRAPH, and only the first day of the conference, and I am already struck by the spirit of teamwork, volunteerism,

and contributing to humanity underlying much that I have seen.

That spirit is as pervasive as the youngsters bedecked in brightly colored shirts patrolling the halls, ballrooms, and meeting rooms of the Vancouver Convention Centre, who have donated their time to provide information and directions to those of us new to the venue. It was notable in each of this year’s recipients of ACM SIGGRAPH Awards, but none more than Scott Lang, an educator now at the Bergen County Academies, whose association with SIGGRAPH began as a volunteer many years ago. It was blatantly obvious in the keynote remarks of Elliot Koteck, co-founder of Not Impossible Labs, a firm whose founding tenet was “Technology for the Sake of Humanity,” and which brings together “people with technical savvy” in multiple disciplines to develop open source assistive technologies to help people who are paralyzed to communicate, those missing limbs to walk and feed themselves, the silent to speak.

Conference chair Dave Shreiner observed that the themes exemplified by the Special Interest Group on Graphics and Interactive Techniques include how technology can be used to enhance people’s lives; compassion, in using technology to help make people’s lives better, and “bridging a collection of communities” across computer and other scientific disciplines to cross-pollinate or innovate useful technologies.

ACM SIGGRAPH president Jeff Jortner discussed the value of volunteering in an organization like SIGGRAPH, including the opportunity to network with mentors and peers, develop leadership skills, interact with a wide range of disciplines, make friends from all over the world, and contribute to the larger community.

Quite an unexpected perspective, for a group best known for its development of complex computer graphics and special effects for blockbuster motion pictures.

ACM president Alex Wolf noted SIGGRAPH is just one of ACM’s 36 Special Interest Groups. ACM, he said, is actively “building and nurturing communities of excellence in computing.” That seems an apt description for SIGGRAPH.

Mark Guzdial is a professor at the Georgia Institute of Technology. Lawrence M. Fisher is Senior Editor/News for ACM Magazines.

© 2014 ACM 0001-0782/14/10 \$15.00

29th
IEEE



www.ipdps.org

25-29 May 2015
Hyderabad, India

International Parallel & Distributed Processing Symposium

IPDPS 2015 IN HYDERABAD, INDIA

The Hyderabad International Convention Centre will host an event that offers the full IPDPS program of workshops, contributed papers, and keynote speakers as well as representatives from industry and opportunities for students to hear from and interact with senior researchers attending the conference. The Hyderabad airport (RGIA) has direct flights from all international hubs and direct flights to all major cities in India. Hyderabad offers a variety of tourist attractions for attendees and their families and a chance to explore other parts of India. Visit the IPDPS Website regularly to see new program developments, including announcement of accepted papers and keynote speakers, and to get tips on travel to the premier conference in parallel and distributed computing.

General Co-Chairs

Susamma Barua, California State University, Fullerton, USA
R. Govindarajan, Indian Institute of Science, Bangalore, India

Program Chair

Srinivas Aluru, Georgia Institute of Technology, USA

Workshops Chair

Ümit V. Çatalyürek, Ohio State University, USA

Workshops Vice Chair

Bora Uçar, CNRS and ENS Lyon, France

PhD Forum Chair

Luc Bougé, ENS Cachan, France

Industry Chairs

Kalyana Chadalavada, NCSA, USA
Jigar Halani, Wipro, India

The IPDPS PhD Forum event will continue with the traditional poster presentations by students working toward a PhD and will again offer an enhanced student program to provide coaching in scientific writing and presentation skills and to give students an opportunity to hear from and interact with senior researchers and industry partners. IPDPS extends a special invitation to companies both in and outside India to consider becoming IPDPS 2015 Industry Partners, sharing in the benefits of associating with an international community of top researchers and practitioners in fields related to parallel processing and distributed computing.



IPDPS 2015 WORKSHOPS

Held on the first and last day of the conference, workshops are a major part of the IPDPS week-long compendium of events and provide the IPDPS community with an opportunity to explore special topics and/or present work that is more preliminary or cutting-edge than the more mature research presented in the main symposium. Proceedings of the workshops are published by the IEEE Digital Library and are distributed at the conference. All of the workshops are linked from the workshops page on the IPDPS Website. Please note that each workshop has its own requirements and schedule for submissions and that the **submission deadline for most workshops is after the main conference author notification date of December 12th.**

HCW	Heterogeneity in Computing Workshop
RAW	Reconfigurable Architectures Workshop
HIPS	High-Level Parallel Programming Models & Supportive Environments
NIDISC	Nature Inspired Distributed Computing
HiCOMB	High Performance Computational Biology
APDCM	Advances in Parallel and Distributed Computing Models
HPPAC	High-Performance, Power-Aware Computing
HPBC	High Performance Big Data and Cloud Computing
ASHES	Accelerators and Hybrid Exascale Systems
PLC	Programming Models, Languages and Compilers for Manycore and Heterogeneous Architectures
EduPar	NSF/TCPP Workshop on Parallel and Distributed Computing Education
GABB	Graph Algorithms Building Blocks
PDSEC	Parallel and Distributed Scientific and Engineering Computing
DPDNS	Dependable Parallel, Distributed and Network-Centric Systems
LSPP	Large-Scale Parallel Processing
PCO	Parallel Computing and Optimization
Par Learning	Parallel and Distributed Computing for Machine Learning and Inference Problems
HPDIC	High Performance Data Intensive Computing
JSSPP	Job Scheduling Strategies for Parallel Processing
HiPeR	High-Performance Runtime
HPDAV	High Performance Data Analysis and Visualization
iWAPT	International Workshop on Automatic Performance Tuning

Sponsored By:

IEEE  computer society
Technical Committee on Parallel Processing

In Cooperation With:



ACM SIGARCH

IEEE Computer Society Technical Committee on Computer Architecture

IEEE Computer Society Technical Committee on Distributed Processing

Researchers Simplify Parallel Programming

Parallel computing has become increasingly important as chipmakers put more and more processor cores on individual chips.

USING STATISTICAL MODELS, computer scientists have shown that certain kinds of parallel computation are not as difficult as previously thought.

Researchers from the Massachusetts Institute of Technology (MIT), Microsoft Research, and the Israel Institute of Technology (Technion) showed that for a large class of non-blocking parallel programs, lock-free algorithms will perform fully as well as the more complex wait-free algorithms.

“What we have shown is that you really don’t need to design these wait-free algorithms,” said Nir Shavit, a computer science professor at MIT. “Most methods that are lock-free are also wait-free.”

The two types of algorithms are part of a class of techniques used in shared-memory systems that are non-blocking, meaning concurrent threads in a program are not halted indefinitely by mutual exclusion. “Lock-free” now generally refers to algorithms that guarantee some concurrent operation will always make progress; that is, that there is guaranteed progress somewhere in the system, regardless of scheduling. The more rigorous “wait-free” algorithms go further, guaranteeing every thread makes some progress.



Technicians at Argonne National Laboratory work on MIRA, the fifth-fastest supercomputer in the world (as of July 2014).

While wait-free seems better on theoretical grounds, it is designed to handle worst-case conditions that very rarely occur, the researchers found. Using a technique called Markov chain lifting, they modeled the average behavior of multiple threads across an entire system over time to simulate the threads' "expected asymptotic behavior." The result showed that for many problems, the two kinds of algorithm will perform about the same. (Nevertheless, wait-free algorithms will still be the right choice for real-time systems where processes have hard deadlines, the researchers said.)

Maurice Herlihy, a professor of computer science at Brown University who was not involved in the work, said the kind of analysis the researchers applied may have broader implications elsewhere in computer science. "I think this kind of thing could be applied to a lot of situations, where theory tends to focus on worst case, rather than expected case," he said.

The World Goes Parallel

Parallel computation has become increasingly important in recent years, moving from the realm of scientific su-

percomputing into corporate database servers and into the cellphones and personal computers of users. Limits on energy consumption, heat dissipation, and chip technology are preventing the speeds of individual processors from increasing at the rate we have seen over the past decades. To increase overall performance, chipmakers are putting more and more processor cores on a single chip. Yet using those processors' efficiency requires the often-difficult task of rendering a program to run in multiple concurrent threads.

In some cases, such as in graphics processing, that is relatively easy, as each processor can be allocated its own separate portion of data. However, many programs contain instructions that need to share data, and programmers often "lock" those during program execution so data in memory is protected while it is being worked on.

Programmers have employed such locks for many years, but they charge a penalty. The operations inside a single lock, which might be in multiple tens of instructions, then can become a bottleneck to the overall execution of the system. One goal in these shared-

memory systems is to reduce the sequential bottlenecks caused by locks, Shavit said. They can be minimized by use of fine-grained locking, which locks up fewer instructions at a time. Yet the use of lock-free algorithms opens up the possibility of parallelism at the lowest possible level of granularity, at the level of individual machine instructions, he said.

Another problem with locks is that they have "terrible semantics," Shavit said. Programmers may set a lock in one place in a program but release it at some number of other, unpredictable places, making it hard to read and debug the code.

While lock-free algorithms solve some problems, they in turn can be tricky to code and debug. Although the move to parallel computing is leading to increased interest in these algorithms, most programmers need not be concerned with this latest analysis of lock-free versus wait-free, said Marc Snir, a computer scientist at Argonne National Laboratory and a professor at the University of Illinois at Urbana-Champaign. "It's a concern for people who work at the bottom of the software stack, say

Milestones

Computer Science Awards, Appointments

LEISERSON RECEIVES 2014 KENNEDY AWARD

Charles E. Leiserson of the Massachusetts Institute of Technology (MIT) will receive the 2014 ACM-IEEE Computer Society Ken Kennedy Award on Nov. 1, 2014, at SC14 in New Orleans.

Leiserson is being recognized for his influence on parallel computing systems and their adoption into mainstream use through scholarly research and development. He also was cited for mentoring computer science leaders and students.

He introduced the notion of cache-oblivious algorithms, which can exploit the memory hierarchy near optimally despite having no tuning parameters for cache size or cache-line length, and also developed the Cilk multithreaded programming technology, and led the development of several Cilk-based parallel chess-playing programs, winning numerous prizes in international competition.

Leiserson's research centers on the theory of parallel computing, especially as it relates to engineering reality. His Ph.D. thesis, "Area-Efficient VLSI Computation," won the first ACM Doctoral Dissertation Award, as well as the Fannie and John Hertz Foundation's Doctoral Thesis Award.

Earlier this year, Leiserson, who received his B.S. degree from Yale University and a Ph.D. degree from Carnegie Mellon University, was a co-recipient of the 2014 ACM Paris Kanellakis Theory and Practice Award for contributions to efficient and robust parallel and distributed computing.

MINSKY LAUDED FOR AI ACHIEVEMENTS

Artificial Intelligence (AI) pioneer and 1969 ACM A.M. Turing Award recipient Marvin Minsky was named recipient of the BBVA Foundation Frontiers of Knowledge Award in the information and

communications technologies category, for his role in defining the AI field and in mentoring many of the leading minds in today's AI community.

The award also recognizes contributions to mathematics, cognitive science, robotics, and philosophy by Minsky, who serves as Massachusetts Institute of Technology Media Lab professor emeritus.

Nicholas Negroponte, co-founder and chairman emeritus of the Media Lab, said Minsky's "genius is accompanied by extreme kindness and humor. He listens well, and is oracle-like in his capability to debug an enormously complex situation with a simple, short phrase. Through the 47 years we have known each other, he has taught me to tackle the big problems."

The Frontiers of Knowledge Awards honor achievements in arts, science, and technology. They focus on contributions of

lasting impact for their originality, theoretical significance, and ability to push back the frontiers of the known world.

INTEL'S KAIVOLA HONORED WITH VERIFIED SOFTWARE MILESTONE AWARD

Roope Kaivola, a principal engineer at Intel Corporation, has received a Microsoft Research Verified Software Milestone Award for the Intel Core i7 verification project.

While formal methods were applied within several areas of the Core i7 project, the award was given in recognition for Kaivola's role as intellectual leader of the core execution cluster, as well as his leadership of the verification team.

Said Judith Bishop, principal research director of computer science at Microsoft Research, "We salute Roope Kaivola and his team for some 20 person-years of verification work, one of the most ambitious formal verification efforts in the hardware industry to date."

people writing the runtime of C++ or the parallel framework in C#,” he said.

Snir said the actual results obtained from wait-free algorithms often fall short of theoretical promises for various reasons. “When you implement very low-level coordination algorithms, there are a bunch of low-level issues that come into play, and the use of wait-free algorithms or not is a small part of those considerations.”

Enter Transactions

In recent years, another lock-free control method has attracted considerable attention. Called “transactional memory,” it works a little like database transactions, which are required to maintain the integrity of databases accessed by multiple simultaneous users. In transactional memory, a unit of code that reads and writes to shared memory is treated as a single atomic transaction, but the memory is not locked against access by other threads. If conflicting changes to memory are made by different threads at the same time, the transaction is aborted and the changes reversed. The net result is that parallelism becomes potentially possible at the machine instruction level of granularity.

Hardware-based transactional memory was introduced last year in Intel’s Haswell and IBM’s POWER8 microprocessors. Transactional memory also exists at the software level in, for example, some C, C++, C#, and Java compilers.

Herlihy at Brown said it is too early to tell how well the new transaction-enabled chips will work. He predicts the emergence of a “transactional programming model” that avoids locks by a hybrid of hardware and software transactions. “It’s my hope that the right mix of fast low-level hardware and simpler high-level software will give us something both efficient and well-structured,” he said.

Such systems might be augmented with new programming languages specifically geared for parallel applications. These “functional” languages treat computation as if it were a series of mathematical functions whose output depends only on input and not on the values of other program or system variables. Examples include Haskell and Scala, and Herlihy predicts ideas from those languages will permeate more mainstream languages.

Transactional memory works a little like database transactions, which are required to maintain the integrity of databases accessed by multiple simultaneous users.

Explains Herlihy, “Transactional memory is a little like the ball bearings, where the concurrent parts synchronize. But languages where it’s very tightly controlled which parts of the program can be shared and modified concurrently can go a long way to making large parallel systems easier to understand and maintain.”

Going Forward

Debate and research surrounding the various methods of concurrency control center on the delays that can occur during program execution, and the impact of those delays on system performance. However, said supercomputer user Snir, other criteria—power consumption, reliability, availability, and maintainability—increasingly trump the performance metric. The largest system now at Argonne has 750,000 processor cores, and Snir predicts the arrival of exascale computing in which single applications run in from 100 million to one billion threads. That will require new codes that are hybrids of shared-memory and message-passing architectures, he said. And at those levels, system designers may worry less about which algorithm produces the smallest delays than about which ones are easiest to debug and which ones consume the least power.

Shavit agrees hybrid transactional systems will become the norm, with fast hardware handling smaller, simpler transactions and software larger, more complex ones. He predicts such hybrid systems will find support in the next version of the GNU C compiler within the

next two years. “Once in GCC, they will show up everywhere,” he said. “People will do more and more with transactions and less and less with locking.”

The concepts of lock-free, wait-free, and transactional memory are at present mostly of interest to programmers writing high-performance parallel system software, such as in the operating system or communications software layers. “But this will change as the number of cores on a chip increases,” Shavit said. “What is now the realm of database people and operating system people will become the day-to-day business of other programmers. These programmers will have languages with all these mechanisms built in. They will be able to use transactions, and the compiler will generate code without locks and with sufficiently low granularity.”

Herlihy said there has never been a more exciting time for research in parallel computation. “The economics has changed, the technology has changed, and there are all kinds of exciting new applications like machine learning and graph search—all these things are crying out for better parallelization.” **C**

Further Reading

- Alisarth, A., Censor-Hillel, K., Shavit, N. **Are lock-free concurrent algorithms practically wait-free?** *46th ACM Symposium on Theory of Computing (STOC 2014)*, New York, NY, May 2014 <http://bit.ly/1rUSxTz/>
- Amdahl, G. **Validity of the single processor approach to achieving large-scale computing capabilities.** *Proceedings of the AFIPS Spring Joint Conference*, 1967 <http://bit.ly/1rUSzL7>
- Ferri, C., Wood, S., Moreschet, T., Bahar, R. I. and Herlihy, M. **Embedded-TM: energy and complexity-effective hardware transactional memory for embedded multicore systems.** *Journal of Parallel and Distributed Computing (JPDC)*, Volume 70, Issue 10, October 2010. <http://bit.ly/1pLwZ4g>
- Herlihy, M., Shavit, N. **The Art of Multiprocessor Programming.** Morgan Kaufmann Publishers Inc., San Francisco, March 2008 <http://bit.ly/1sbBYRo>
- Intel Corp. **Introduction to Parallel Programming video lecture series**, 2012. <http://intel.ly/1AuM2pc>

Gary Anthes is a technology writer and editor based in Arlington, VA.

© 2014 ACM 0001-0782/14/10 \$15.00

Computing What Fits

New apps and pods improve the virtual and actual retail experiences by ensuring clothing and cosmetics look good on you before you buy them.

FOR MANY MEN, shopping with their wives is akin to having someone scratch their nails across a chalkboard. For Romney Evans, however, such an outing back in 2004 turned out to be a fateful day. His wife was having “a horrible time finding jeans,” he recalls. “She must have tried on a dozen items ... and I think she was feeling bad about herself. I thought, ‘This is crazy; is it really this difficult?’”

Evans thought there had to be a way to enter some information online and find the right style, so he started researching how other people were attempting to deal with the problem. While in graduate school at Babson College, Evans says his goal became launching a consumer site devoted to helping people find the right fit in jeans. He found a partner in classmate Jessica Arrendondo Murphy, who had a background in apparel and buying, and in 2007 they launched what is now called True Fit. The company’s “fit confidence engine” analyzes reams of data to help consumers understand how well clothes and shoes they see online will actually fit them.

Today, the technology is used by department stores like Macy’s and Nordstrom’s; consumers visiting those companies’ websites can create a True Fit profile and receive recommendations of designers whose clothing will work on their body type. “We’re converting browsers into buyers,” says Evans.

There are numerous apps, walk-in pods, and virtual “fitting rooms” that help consumers determine what clothing will accurately fit them and allow them to virtually try cosmetics, and it is not difficult to figure out why. eMarketer estimates apparel and accessories account for 17% of overall retail ecommerce sales, which are projected to reach almost \$60 billion this year, and to grow to \$86 billion by 2018. Forrester projects that by 2017, 60% of U.S. retail

The goal of these apps is to both help people feel confident about the choices they make, and to reduce the cost of returns for retailers.

sales will be influenced by research on a laptop or mobile device.

Technologies to make such apps and pods possible include augmented reality in the case of makeup apps, according to Charlene Li, founder of research and advisory firm Altimeter Group. A video camera captures an image of your face, and the system superimposes images of makeup products over the first image, as if you were looking in a mirror, she says. Also frequently used is Kinect from Microsoft, a 3D scanning technology that acts like a camera and senses motion, allowing users to interact with it through gestures and spoken commands. Kinect can model objects in 3D based on depth information.

Yet the benefits of makeup apps are being diluted by the experience, maintains James McQuivey, a vice president and principal analyst at Forrester Research. While technology makes the apps appear elegant, companies are dropping the ball in “how they invite you to try the app, how they guide you in making decisions inside the experience, and how they then connect you with tutorials on how to apply that makeup, or even something as simple as help you buy the products you need to complete the look itself.”

Consequently, McQuivey says, consumers are left with an app “that is interesting to try once or twice, but is unlikely to really change the way people shop for and experience makeup.”

Reducing Returns

The goal of these apps is both to help people feel confident about the choices they make, and to reduce the cost of returns for retailers. Clothing is notorious for being one of the most difficult ecommerce categories, and one in four clothing items bought online is returned, mainly because the size the consumer selected was wrong, according to Fits.me founder Heikki Haldre.

London-based virtual fitting room company Fits.me addresses online fit with shape-shifting robotic mannequins “that can contort to thousands of body shapes and sizes,” says Haldre. “The mannequins are used to provide consumers photographic images [of] how a garment looks on a model of their own shape and size.”

Similar to True Fit, consumers can go to a participating retailer’s site, plug in certain body measurements, and Fits.me will offer up actual photos of how a garment will fit their body type on an avatar. At the website of London retailer Thomas Pink, for example, consumers are offered the option to “try it on” by clicking a button and entering a virtual fitting room. They enter various measurements and then are shown the “virtual me.”

“Before Fits.me’s solution, online retailers could only show images of garments on standard models who, for the majority of us, do not represent real consumers at all,” says Haldre.

Fits.me’s proprietary technology was built across several universities in Europe and has been in development for the past five years, Haldre says. The company began commercializing the system in 2013. “We started by collecting 3D body scans from over 20,000



A Fits.me robotic mannequin, which takes a user's measurements to demonstrate how clothing will fit.

people to build robotic mannequins that, as closely as possible, resemble the population.” The shape-shifting robots use about 70 actuators, and “by moving these actuators to different positions, we can control over [one] million ways to manipulate discrete 3D shape[s],” Haldre says.

Solving the Color Problem

The enthusiasm of teens and college students for clothing, shopping, and how they look was the impetus behind White Mirror, a virtual fitting room that began as a senior project for Rice University electrical engineering student Lam Yuk Wong and her friend Cecilia Zhang.

The two students and their friends liked to shop online, finding it convenient “but also troublesome,” Zhang says, because sometimes sizes would be off, or they would wonder whether a certain color would look better on them than another. “It’s related to your skin color or hair—there are many things on your body that you have to be concerned about.”

Wong and Zhang proposed the idea of a more “authentic” system that would help online shoppers visualize how a potential garment would

look on them before making the decision to purchase it. They felt systems that suggest sizes to a customer when they enter body parameters were unreliable “because we tried them, and customers are still not able to see or visualize this garment and whether it fits their skin color or body. There’s no visualization and also, it doesn’t solve the color problem.”

In creating White Mirror, they used Kinect to capture a customer’s body type and create “a ‘virtual you’ in the system,” says Zhang.

They also wanted to make White Mirror operate very quickly, Zhang says. In order to get an accurate simu-

“In our system we combined the physical simulation with machine learning techniques.”

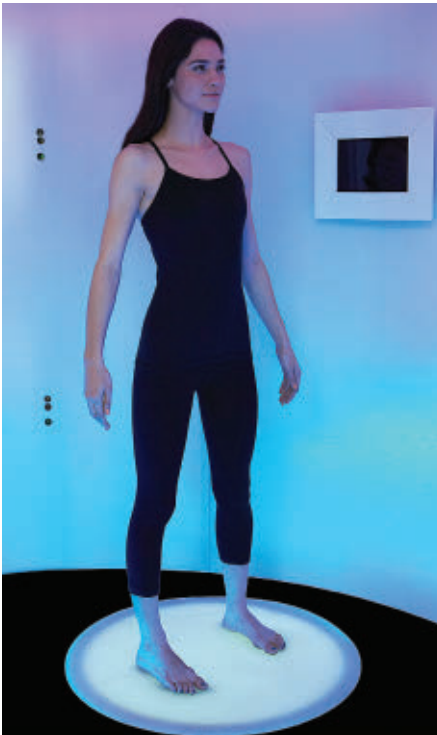
lation of how a garment will look on someone, she says their system utilizes a technique called physical simulation, “which deals with physical particles between the garment and the body in order to see how the garment will form on the body.” However, it takes 40 seconds for physical simulation to work, which is longer than online shoppers will want to wait, Zhang says, so “in our system we combined the physical simulation with machine learning techniques, so our system is able to finish the simulation in 1.2 seconds. Our point is to make a real-time simulation.”

As electrical engineering majors, Zhang and Wong knew little about computer graphics, so they reached out to authors of papers on garment simulation for advice. “We were trying to figure out the algorithm to use to make this all work and make it faster, because we encountered quite a lot of difficulties when implementing this system.” Integrating machine learning techniques allows them to finish the physical simulation behind the scenes, says Zhang, so the system is able to display the customer’s 3D image wearing a selected garment as if in a real fitting room. White Mirror has not yet been tested on shoes.

The system currently requires a second person to hold Kinect to scan a person’s body, but Zhang and Wong are working on refining the system and improving the accuracy of its results. They plan to turn White Mirror into an app, and have received funding to do so from the university, as well as from IBM.

For consumers who hope to avoid having to try on items in a brick-and-mortar store, Bodymetrics has developed walk-in pods that utilize 16 Kinect devices to create a full 3D body scan, says Suran Goonatilake, the company’s CEO and co-founder.

“The pod uses automated ‘body-measurement software’ to generate up to 200 accurate body measurements,” Goonatilake says. The data is securely held on a Microsoft Azure cloud platform and is used to run “a bunch of body analytics.” The pod also generates applications that range from showing the best garments for a body type at a retail store to showing “what you might look like if you were



As many as 200 of a woman's measurements are captured by 16 Kinect devices inside a Bodymetrics walk-in pod, in order to generate a 3D body scan that can be used to match her physical attributes to specific articles of clothing.

to lose 10 pounds at a fitness club," he says.

Altimeter's Li says using Kinect in a walk-in pod is more accurate than an online app. "It's a fixed location," she explains. "It knows exactly where you're standing, so the system has a greater level of accuracy."

However, consumers have to wear tight-fitting garments when entering the pod; "If you wear baggy clothing, the scanner will be getting measurements of the baggy clothes and not your body," Goonatilake explains. He says Bodymetrics has scanned over 10,000 people to date.

Until recently, most Bodymetrics pods were installed on a trial basis, to gain understanding of how users react to the technology, Goonatilake says. The pods now are being rolled out at two "youth-oriented" fashion retail chains in the U.K., and a fitness chain in the U.S. with 50 clubs nationwide, on a subscription basis for \$3,000 per month.

Virtual Makeup

It is not just clothing that users can try on virtually; there are also apps that can help you decide if a particular eye shadow or lipstick works with your coloring. The apps are designed to give

consumers the opportunity to try out products they may not have previously considered without going to a retail store. L'Oreal, for example, last year rolled out the Makeup Genius app, which turns the front-facing camera of an iPad or iPhone into a makeup mirror so consumers can virtually try on a variety of cosmetics.

Shiseido has developed a makeup simulator on an iPad-based device called the "Beauty Tablet," or "B-Tab," which is used by all its beauty consultants in Japan, according to Sho Nagai, a spokesperson for Shiseido Co., Ltd. "The device is equipped with special sensors that recognize and identify different parts on the face," Nagai explains. "The sensors can detect the shape of the eyes and mouth so the virtual makeup stays on the screen, even after blinking or moving."

Improving Recommendations

As True Fit's Evans notes, one of the most frustrating aspects of shopping is that a size 6 in one brand could be very different in another. "If everybody's size is speaking a different language, we're helping translate that."

Powered by Hadoop, an open source framework that enables large-scale processing of datasets, True Fit

gathers data from retailers, consumers, and over 1,000 brands to help the company finesse recommendations based on historical performance and inferences, Evans says.

True Fit assembles the gathered data and uploads the "blueprints of an item" such as fabric details and style attributes (like whether a pair of jeans is high-cut or boot-cut), says Evans. "When a consumer creates a True Fit profile, they tell us about their body—their height, weight and age, and heuristic indicators—are their hips curvy or straight, to understand their perception of themselves." He says the system does not ask for measurements, just that the user identifies the kinds of items they like to wear. For example, a user may say they like a Michael Kors A-line sleeveless dress and that they wear a size 4; "Then we have a really good understanding, because we already have the information from Michael Kors, and so we match that to what consumer is telling us."

Evans likens True Fit to Pandora; in the same way Pandora gathers attributes of songs and looks at the patterns of those attributes to suggest other songs a user will like, True Fit does the same with items of clothing.

The True Fit system gathers sales

feed information from a retailer about individual users, like the number of returns and what a user's successful purchases have been, "and we triangulate that set of items against our database of product information from the brands" to add to a user's profile. "The magic is, we have blueprints of what successfully fit a person, and use machine learning processes to make high-confidence recommendations, which improve over time as consumers buy and return things," says Evans. "If we make a recommendation that doesn't work out, we learn from that and the system fine-tunes that."

Evans will not reveal how many True Fit profiles have been created, but says the total is in the millions.

They Do Not Always Work

For some, such devices and apps may not achieve their desired effect.

"The biggest issue is most of us don't like the way we look," notes Altimeter's Li. "It's one thing when I try something on, [but] I look at it fairly differently on a model than on me."

Forrester's McQuivey predicts Amazon "may do to virtual or magic mirrors what Apple did to phones." He envisions the future of technologies that help us determine appropriate looks and fits will be in a person's bathroom and closet—"not the retail environment, and certainly not the living room"—and will be "much bigger than just trying on virtual shoes or applying virtual cosmetics." The bathroom of the future will have a mirror that will be a huge display, which will show you any particular fashion or cosmetics look you have in mind, he says.

McQuivey also anticipates Amazon or someone else "will have figured out what to show you when you look in the mirror. Perhaps the service can see your calendar and knows that you have a job interview, and so will virtually dress you and style your hair in a variety of looks...which you can then modify interactively."

The service will also care about more than just your look, he adds, and may also be able to add to your mirror image "data about your health, your condition, your muscle tone and body fat, all of which it will sense using cameras and other devices in your bathroom." If you're a few pounds

"Creating a digital relationship with a customer ... could theoretically influence a large share of their consumer spending."

up after a holiday, for example, correspondingly the outfits the mirror will suggest will be the ones it knows you can fit into and are appropriately slimming, McQuivey says.

Technically, all of that is possible today, he says, but businesses have to figure out how to get there. "It will be a mad race to that point because there is so much to be won—or lost—in this race, creating a digital relationship with a customer that could theoretically influence a large share of their consumer spending." **C**

Further Reading

McQuivey, J.

Digital Disruption: Unleashing the Next Wave of Innovation.
<http://bit.ly/1w8RSPR>

Christensen, C.

The Innovator's Dilemma: When New Technologies Cause Great Firms to Fall
<http://bit.ly/1q8Pupq>

Downes, L., Nunes, P.

Big Bang Disruption: Strategy in the Age of Devastating Innovation, 2014.
<http://amzn.to/1rwkxGW>

User mindsets are driving the future of applications.
<http://pwc.to/1welZCj>

Future of enterprise apps: Moving beyond workflows to mindflows.
<http://pwc.to/1BnPxjC>

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., Marrs, A.

Disruptive technologies: Advances that will transform life, business, and the global economy. McKinsey Global Institute, May 2013. <http://bit.ly/1nGHR3R>

Esther Shein is a freelance technology and business writer based in the Boston area.

© 2014 ACM 0001-0782/14/10 \$15.00

ACM Member News

ANDERSON 'WORKING ON THE PLUMBING;' RESEARCHING REAL-TIME EMBEDDED SYSTEMS



"Those of us who do real-time embedded systems research work on the plumbing and it's hard to make

that exciting, but it's absolutely necessary and crucial," says James Anderson, W.R. Kenan Distinguished Professor in the department of computer science at the University of North Carolina, Chapel Hill.

As an undergraduate at Michigan State University, Anderson changed majors from electrical engineering to computer science "because the cleanness of the logic appealed to me." He continued to study computer science and received a master's degree in the discipline from Purdue University and a Ph.D. in computer science from the University of Texas at Austin.

Today, he says, "I love teaching computer science; it gives me the freedom to call my own shots." He relishes working with students and "the thrill of discovering a fact that another human being has never uncovered."

His current research focuses on using multicore processors in real-time embedded safety-critical systems. Dual-, quad-, and eight-core processors are common in general-purpose and high-performance computers, but deploying multicore processors in real-time embedded systems is still rare, because "embedded systems researchers are still figuring out how to use multicore," Anderson says.

Anderson works to resolve the complex resource allocation problems associated with real-time analysis, and validate the timing correctness.

"The real-time community requires formal proof. We focus on pessimistic worst-case scenarios for proof-based reasoning. We can't just come up with a benchmark; things move more slowly for us," he said. "The general theme for me is how to accomplish multicore processing with less infrastructure and at less cost."

—Laura DiDio

Keeping Online Reviews Honest

As online reviews become increasingly important to e-commerce and social media sites, computer scientists work on ways to ensure their authenticity.

AS COMMERCE INCREASINGLY moves online, consumers worldwide are relying more and more on online reviews to help them make decisions about products to buy and businesses to patronize.

According to a recent study by BrightLocal (<http://seind.com/1xzy0Xb>), 88% of U.S. consumers read online reviews “to determine whether a local business is a good business” at least occasionally—39% do so regularly. Also, 72% say positive reviews lead them to trust a business more, while 88% say that in “the right circumstances,” they trust online reviews as much as personal recommendations.

What are “the right circumstances?” In response to that last question in the BrightLocal survey, 30% said they trust online reviews if they believe they are authentic.

The problem of authenticity in online reviews is a long-standing and stubborn one. In one famous incident back in 2004, Amazon’s Canadian site accidentally revealed the true identities of thousands of its previously anonymous U.S. book reviewers. One insight the mistake revealed was that many authors were using fake names in order to give their own books favorable reviews.

The value of a positive review has also led to various review-for-hire operations, in which Web users get paid to write reviews on sites such as Amazon and Yelp. Bing Liu, professor of computer science at the University of Illinois at Chicago and author of *Sentiment Analysis and Opinion Mining*, told *The New York Times* in 2012 that about a third of all Internet reviews were fake.

In their 2013 paper titled “Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud,”

The effects of rampant review fraud can be devastating for both businesses and consumers.

Michael Luca, assistant professor of business administration at the Harvard Business School, and Georgios Zervas, assistant professor of marketing at Boston University School of Management, estimated 16% of Yelp restaurant reviews were fraudulent.

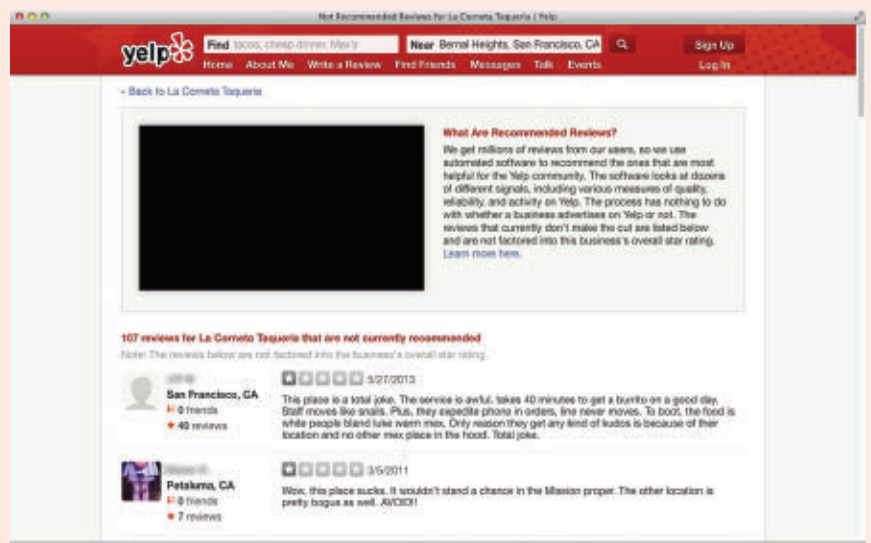
The effects of rampant review fraud can be devastating for both businesses

and consumers. Sometimes fake reviews are posted by business owners either leaving negative reviews for their competitors, or positive reviews for their own businesses. Having to combat false negative reviews can take a lot of time and energy that a business would better apply elsewhere. In addition, many small business owners are not sufficiently computer-savvy to stay on top of such attacks.

The importance of online reviews also leaves a business open to extortion, in which customers threaten to write a negative review if they do not get special treatment—a form of fraudulent review, since it would not reflect how an average customer actually was treated by the business.

For consumers, a fraudulent review can lead to negative outcomes. A traveler researching places to stay on TripAdvisor, a site built on customer reviews of hotels and restaurants, could

Figure 1. Yelp lets users see reviews that do not satisfy its authentication algorithm, but explains why they are in a separate section.



be led by a false positive review to stay somewhere that is not up to their standards, or which does not provide the specific amenities described. A fraudulent review on Amazon could result in the purchase of a shoddy or unwanted product.

How to Spot a Fake Review

Because of the potential for negative experiences—for which consumers may blame the website where they read the erroneous review—efforts to guarantee the honesty and authenticity of online reviews have attracted the attention of researchers.

There are several ways consumers themselves can get a good idea of whether a review is fraudulent or not. On most sites, a potential reviewer must create a user account. That may not (and usually does not) reveal their real name, but it does provide a way of checking to see if the reviewer has reviewed anything else. It certainly is possible for someone to legitimately post just one review—for example, if they join TripAdvisor specifically to rave or complain about someplace they went on a recent trip—but in general, a reviewer with just a single review can be considered a red flag.

Computer algorithms can also help detect fake reviews, and this is an area prime for computer science research. It can be difficult to evaluate the authenticity of a single reviewer, but there are ways to identify fraudulent reviewers operating as a group. Creating a single user account and posting one review of a hotel or restaurant is not likely to be effective at influencing the general opinion of a community, but several reviews all saying the same thing can have a major impact and can subsequently dominate the conversation. To efficiently flood a site that way, a group needs to automate the creation of user accounts. That leads to one way an algorithm can identify such a false source: usernames with more than three numbers at the end strongly suggest accounts created by an automated program. Several reviews by different user names with numbers at the end could indicate a group of review spammers.

Researchers are looking at ways to detect collusion among fake reviewers based on the review contents. Bing Liu

Figure 2. Yelp uses algorithms to ensure reviews in the “Recommended Reviews” section are authentic; this also provides a way to check a reviewer’s activity on the site.



worked with Arjun Mukherjee, a Ph.D. candidate at the University of Illinois at Chicago, and Natalie Gance, a Google software engineer interested in social shopping, on a software algorithm to

detect such activities. In a paper presented at the World Wide Web 2012 conference, they described their use of a behavioral model called GSRank to detect fake reviewer groups. Their

model evaluated relationships among reviewers, groups, and products to detect clusters of people who have all reviewed the same set of products.

In his book, Liu discusses other methods of detecting fraudulent reviews. He cites such clues as consistent word choices across reviews, similarity of content and style among different reviewers, reviewers that fail to maintain a consistent identity (for example, referring to “my husband” in one review and “my wife” in another), multiple reviews from a single user all posted within a short period of time, multiple reviews from a suspected group posted around the same time, and so on. All of these things are clues an algorithm could pick up to suggest the reviews were fraudulent.

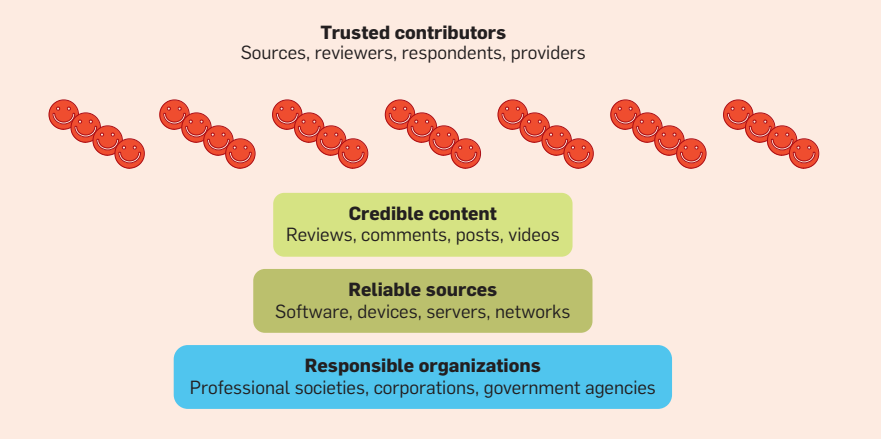
These are the kinds of indicators Yelp’s filtering algorithm looks at, according to Yelp PR specialist Hannah Cheesman. “Our recommendation software takes a large amount of data into account, such as email address, IP address, pattern recognition, etc., to determine which reviews we know enough about to recommend,” says Cheesman. “The software is engineered to weed out possible fakes (several reviews generated from the same IP address), biased reviews (written by a competitor or solicited by a business owner from friends, family, or favored customers), unhelpful rants or raves, and reviews written by users we just don’t know much about (no friends, no picture, few reviews, etc.) and therefore can’t recommend to our community.”

That last point—what the community knows about the reviewer—is a primary concern for Sorin Adam Matei, an associate professor of communications at Purdue University whose work deals with trust, reputation, and authority in social communities. Matei is a cofounder of Kredible.net, a community of researchers interested in ways to define those qualities, as well as co-editor of the forthcoming *Roles, Trust, and Reputation in Social Media Knowledge Markets* (part of the Springer Computational Social Sciences series). For Matei, one key to identifying trustworthy online reviews is the track record of the reviewers. “Trust and credibility in the context of social media depend primarily on functional relationships. What you do is more important than who you are,” he says. Whether human

Figure 3. Amazon provides tools to evaluate the authenticity of reviews, such as links to all a reviewer’s activities, and verification they actually own a product they review.



Figure 4. A framework for analysis of social media communities. From *Roles, Trust, and Reputation in Social Media Knowledge Markets*, courtesy of Ben Shneiderman.



One approach to resolving that doubt is exemplified by Yelp, which separates recommended reviews from those that do not satisfy its screening standards.

or algorithm, following the activity of a reviewer is how you know whether you can trust his or her reviews.

What to Do with Fake Reviews

All these techniques can suggest a review is fraudulent, but by themselves they do not solve the problem of how a site can make certain its visitors can trust the reviews it hosts. One approach to resolving that doubt is exemplified by Yelp, which separates recommended reviews from those that do not satisfy its screening standards. The latter are still available for users to read, but they are marked as “not recommended,” with an explanation of what that means.

Other approaches involve limiting who can write reviews in the first place. Ben Shneiderman, a professor of computer science at the University of Maryland, contributed the chapter “Building Trusted Social Media Communities: A Research Roadmap for Promoting Credible Content” to the book Matei edited. He writes, “Many strategies are being tried to ensure that only trusted contributors participate, such as raising the barriers to entry for contributors by requiring a log-in (no anonymous contributions), identity verification, background check, probation periods, and public performance histories. Greater transparency about who the contributors are and what their past is has the potential to increase trust in their future contributions.”

Transparency is a vital element, in the view of Purdue’s Matei. “Trust and credibility are the product of the transparency of a user’s activities multiplied by the speed and the cost of verifying

them,” he explains. “Individuals that interact online through ratings, reviewing, collaboration, or shopping trust each other on the basis of the traceable activity they leave behind. The easier it is to track them, the more trust and credibility are generated.” He cites eBay as a good example of how that can be implemented: users can rate both buyers and sellers, and their overall rating is visible to all.

Amazon, on the other hand, does not do as good a job, in Matei’s view. “The credibility of its reviewers is mostly a function of productivity,” he says. Amazon does provide links to let users examine a reviewer’s other reviews, and notes whether the reviewer is known to have purchased the item being reviewed (although, of course, if the reviewer purchased it elsewhere, Amazon would not know about it). Some reviewers are in the site’s “Hall of Fame,” meaning other visitors have rated their reviews as helpful, “but the exact algorithm and the numbers behind it are not immediately available,” says Matei.

Matei also points to Wikipedia as a site that, despite its size and popularity, has its credibility questioned because no one really knows who its authors and editors are.

Another approach to promoting trustworthy reviews is illustrated by Reddit, a community where the visibility of content is determined by up-and-down votes from its users. When a user visits a Reddit community, what they see first is based on the opinion of the other members about its relative value. That can be distorted, of course—people tend to vote down comments they disagree with, regardless of their quality—but in general, it represents a consensus about the value of content.

It is vital for e-commerce sites and social media communities for participants to be confident the reviews they read are authentic. As more and more people do their shopping online and rely on strangers’ recommendations, the ability to filter out fraudulent reviews, whether left by a business’s competitors or by non-customers hired for the purpose, becomes of critical importance. Computer science can help with ways to both identify and sequester fraudulent reviews and to promote trustworthy ones, and sites can imple-

ment measures to help visitors track reviewers’ activities and judge which are the most reliable. **C**

Further Reading

Bertino, E. and Matei, Sorin Adam, eds. *Roles, Trust, and Reputation in Social Media Knowledge Markets*, Springer (2014).

Hu, N., Liu, L., and Sambamurthy, V. *Fraud Detection in Online Consumer Reviews*, *Decision Support Systems*, vol. 50, no. 3 (2011).

Liu, B. *Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies)*, Morgan & Claypool (2012).

Liu, B., Mukherjee, A., and Glance, N. *Spotting Fake Reviewer Groups in Consumer Reviews*, *Proceedings of the 2012 World Wide Web Conference*, <http://www2012.org/proceedings/proceedings/p191.pdf>.

Luca, M. and Zervas, G. *Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud*, Harvard Business School NOM Unit Working Paper No. 14-006 (2013).

Logan Kugler is a freelance technology writer based in Clearwater, FL. He has written for over 60 major publications.

© 2014 ACM 0001-0782/14/10 \$15.00



Association for
Computing Machinery

ACM Conference Proceedings Now Available via Print-on-Demand!

You can now order many popular ACM conference proceedings via print-on-demand, through Amazon, Barnes & Noble, Baker & Taylor, Ingram and NACSCORP!

For available titles and ordering info, visit:
librarians.acm.org/pod

Economic and Business Dimensions

The Dark Side of the Sharing Economy ... and How to Lighten It

Improving the sharing economy will require addressing myriad problems.

BECAUSE WE LOVE the sharing economy we want to improve it. But most pundits are telling only half the tale: Naysayers are too bombastic and boosters too unrealistic. Improving the sharing economy means dealing realistically with its dark side. To assure sharing will grow up, we need to avoid market and regulatory failures that allow parts of the market to gain unfair advantage over others. Regulatory arbitrage is not the right answer. Instead, sharing must ultimately create real consumer value.

It is not too early to begin. Sharing is quickly spreading. People already have access to rooms (AirBnB, Roomorama), tools (SnapGoods) cars and bikes (RelayRides, Wheelz), and ad hoc taxi services (Uber, Lyft). These two-sided platforms offer many advantages by unlocking the value

inherent in sharing spare resources with people who want them.⁴ The size of the sharing economy is estimated at \$26 billion.^{1,16} Internet mediaries now match demand and supply in real time on a global scale. The potential macroeconomic gains are colossal, but problems abound.

The Sharing Economy's Dark Side

The Hotel Zone. Transients and tour-

It takes time to balance conflicting needs, but sharing is growing quickly.

ists do not always respect the sensibilities of long-term residents. Conflicts over tenement buildings helped motivate the first U.S. zoning laws⁵ that sharing now circumvents. In addition, short-term rentals create shortages of affordable long-term housing when nightly rates exceed monthly rentals. Passing that tipping point can hurt individuals at lower income levels¹¹ even as it boosts income for homeowners. It takes time to balance conflicting needs, but sharing is growing quickly.

No Soup for You. In a famous episode of "Seinfeld," one of the lead characters was denied soup by a renowned but humorous cook.^a Sharing biases online is as natural as sharing cars and couches. But when

a See the *No Soup for You* episode summary: <http://bit.ly/X5rBCE>.



biases are unfounded and consumers or producers bludgeon each other unfairly in social media, who will intercede? Biases can mislead, ostracize, shill unearned praise, and damn worthy competitors. A recent study found 16% of Yelp reviews are not genuine.¹⁰ By posting unfounded complaints, customers can punish providers for their transgressions. In some cases, providers will have to fight back to save their reputations.^b

Taxing the Taxi. Unlike licensed taxi drivers, private citizens providing ride-share services do not always purchase medallions. They also do not take licensing exams, or necessarily carry commercial insurance. For these reasons and because they are not required to honor all ride re-

quests, a German court banned Uber's basic service throughout the nation.³ Licensed taxi drivers are saddled with greater costs, which hampers their ability to compete with ride sharing. Arguably, ride sharing is growing by circumventing costs and regulations that govern incumbent businesses. Ride sharing can exploit loopholes to avoid rules and taxes. When this occurs, the sharing economy becomes the skimming economy.

Shared Economies or Shared Serfdom. A commentator at *The New York Times* concluded that what he earned by ride sharing barely covered gas and depreciation.^c Peer-to-peer exchange based on sharing ideas or perform-

ing work (such as Mechanical Turk or TaskRabbit) "on the cheap" strips opportunity from the bottom of the pyramid, as jobs move from traditional manufacturing and services to micro-services.¹² Micro-outsourcing that pays for only the task at hand can shed overhead but mortgage the future by covering only marginal costs and leaving nothing for new skills, health care, or retirement. If information goods are an indicator, marginal costs approach zero, so even covering them might not pay much. Jaron Lanier's book *Who Owns the Future?*⁹ calls this issue a matter of dignity: "if you have to sing for your supper for every meal, you're ... one run of bad luck from losing [everything]."^d Going freelance is hollow freedom when the wage for labor is free.

^c See comments on Sundararajan, A. "Trusting the 'Sharing Economy' to Regulate Itself" *Economix Blog New York Times* (Mar. 3, 2014); <http://nyti.ms/1krGSHo>.

^d See <http://bit.ly/1nSYIn9>.

^b One restaurateur fights back against an unjustified Yelp review: <http://bit.ly/1tTJDTq>.

ACM Transactions on Accessible Computing



This quarterly publication is a quarterly journal that publishes refereed articles addressing issues of computing as it impacts the lives of people with disabilities. The journal will be of particular interest to SIGACCESS members and delegates to its affiliated conference (i.e., ASSETS), as well as other international accessibility conferences.

www.acm.org/taccess
www.acm.org/subscribe



Association for
Computing Machinery

History informs us about the willingness of new businesses to accept risk.

Whose Ox Gets Shared? Sharing creates a subtle tug-of-war between the primary producer and secondary sharer. Secondary sharing can become tertiary taking. Consider the extreme case of a Netflix subscriber who pays \$20 per month, rents three DVDs from Netflix, and then rents each for \$1 per night to other individuals. The subscriber makes \$90 ($1 \times 3 \times 30$) each month while paying substantially less for an asset he or she never owns. Similar to the way put and call options share the value of other people's stocks, sharing allows us to go long or short on physical assets. It might be economically efficient, but it can harm the demand for Netflix and the people who produced the movies. Broadcasters sued the startup Aereo for leasing a device that let individuals capture broadcast TV signals to stream shows to devices they (or others) own. The suit reached the U.S. Supreme Court and, on the basis of copyright infringement, the U.S. Supreme Court ruled against Aereo.^e The City of San Francisco likewise shut down a subversive sharing service. Counselors issued a cease and desist order to producers of a mobile app that rewarded drivers for sharing news of their willingness to vacate public parking. It then auctioned off their spaces.^f Insider sharing just became the latest form of insider trading.

The Dark Side of the Moonlighting. Last New Year's Eve, an off-duty driver for the ride-sharing service Uber killed a pedestrian while hunting for fares. Since the driver was a "contractor," the sharing service would not

compensate the victim's family.^g The contract stipulates that the service is a matching platform and "the company does not provide transportation services, and ... has no liability for services ... provided by third parties."^h Who then will bear the costs of such disasters? Jaron Lanier says these new business models enjoy profits while offloading risk to others. When society picks up the tab, these new business models raise concerns. Maybe they are no cause for celebration.

Lightening the Dark Side
Absorb Risks that Benefit the Ecosystem. Shedding risk and never admitting culpability is standard lawyers' advice. In following this advice, Uber did not indemnify its driver yet has reversed itself dramatically expanding coverage. Putting share-rides into taxi-ride perspective, New York City experienced 44 taxi-related fatalities in 2009 alone.ⁱ

History informs us about the willingness of new businesses to accept risk. Banks originally opposed the Fair Credit Reporting Act (FCRA), complaining about the increased liability and responsibility for unauthorized transactions. Banks argued that FCRA would promote fraud, encourage people to be careless with credit cards, and reduce bank willingness to extend credit. Precisely the opposite occurred. The FCRA passed in 1970, with amendment 15 US § 1643 later limiting consumer liability to \$50 for fraudulent credit card use. However, banks ultimately learned that protecting customers is good business: increased credit card use more than offsets growth in fraud-related costs. Most banks now do not even hold consumers responsible for the \$50. If sharing moves in the same direction, sharers and society will be protected.

Invest in Your Customers. As a startup, Airbnb could not gain traction with renters as long as people put up low-quality listings and photos of their

^g Diamicis, C. "Uber driver hits, kills 6 yr old girl. Is 'Not our problem' still an appropriate response?" (Jan. 2, 2014); <http://bit.ly/190hsft>.

^h See <http://bit.ly/1pxnIgK>.

ⁱ New York City Pedestrian Safety Study: Technical Supplement (Aug. 2010). <http://on.nyc.gov/1AqODmU>.

^e Steel, E. "Stung by Supreme Court, Aereo Suspend Service;" *New York Times* (Jun. 25, 2014), A20.

^f Coté, J. "SF Cracks Down on 'Monkey Parking' Mobile App" June 23, 2014; <http://bit.ly/ToFdNA>.

rooms. Improved listing quality doubled revenues.^j That led Airbnb to educate its users in how to improve their listings.^k Author Michael Schrage¹³ points out that investing in customers helps them create more value. When they can create that value, the ecosystem wins. Gary Swart, past CEO of oDesk, had his company partner with SkilledUP to grant access to thousands of online courses because better-trained freelancers charge more for their work and deliver higher quality.

Community Policing and Self-Regulation. Platforms can be better than governments at spotting stalkers, running background checks on sharing service providers, and responding quickly to conflicts among members. Platforms are closer to the action; and they have an incentive to look after their communities. That is how they make money.⁶ Venture capitalist Nick Grossman says peer-sharing systems develop scalable enforcement, like reputation systems, that are more inclusive than licensing regimes. Self-regulation can work (For examples see The American Medical Association and National Association of Realtors.¹⁴) Problems with self-regulation arise from harm to non-members, market power, and network effects that encourage firms to engage in anticompetitive behavior. But with care, these problems can be avoided. Moreover, sharing can also help government regulators. For example, health inspectors are using Yelp ratings to identify restaurants that may be sources of food poisoning.⁸

j See <http://bit.ly/1gmYXsd>.

k See <http://bit.ly/1dvNMm1>.

The problem is not whether to bury or build the sharing economy: it is already on the ground.

Tax Fairly and Don't Promote Arbitrage. The City of Amsterdam has begun to support sharing economies.¹⁵ Hosts renting their homes to others pay income and tourist taxes, and must ensure neighbors stay neighborly. Private individuals do not require liquor licenses or kitchen inspections as do major hotel chains. Regulations that lie between the individual and industrial levels in this fashion can be a wiser way to handle tax arbitrage. Individuals do pay reasonable taxes that support the community but do not pay at industrial levels.

Current laws that gouge one group to benefit another also need reform. If technology permits low-fee providers to substitute for high-fee ones, simplifying and designing laws that do not promote arbitrage is possibly a better answer. When was the last time your airport car rental cost less than you expected? Sticker shock sets in when cities impose head, gas, airport, and other taxes on tourists. Taxi medallions help regulate taxi services, but also create cartels where medallions offer investment-grade returns to their respective cities.⁷ Taxes on private cars used in shared services might make sense because the roads still require maintenance. However, balance is essential. Internet-enabled sharing does not mean “no taxes,” as Amsterdam has shown.

Create a FICO of Reviews and Fair Access of Resources

To a great extent, the viability of shared services hinges on the quality of review systems because people rely on them to decide whether and what to purchase. Authenticating the validity of reviews is critical to prevent abuse. An independent agency might help prevent glowing “sock puppet” reviews or unfair criticisms. Certification might even deflate mutual excess flattery. Credit scores and information have been monitored in a similar fashion for many years by several agencies, including FICO. Also regulators must ensure public access to public information. Sharing news must not be used to make public resources private. The sharing economy requires that complete information and trustworthy reputations be available to all parties.

The problem is not whether to bury or build the sharing economy: it is already on the ground. The gains are too great to pass up because of misdeeds on the part of a few self-serving actors. The larger opportunity is to move forward despite the disruption. In the short run, platform firms should indemnify users and self-regulate the health of their ecosystems. At the same time, consumers should choose sharing platforms based on short and long-term gains as well as individual and community benefits. Learning and appropriate regulation for fair reporting and fraud protection will be central—although it will need a light touch to encourage innovation while still watching for problems. The task is to share the pain and the wealth. If this sharing happens, the wealth will grow and endure. C

References

1. Botsman, R. and Rogers, R. *What's Mine Is Yours: How Collaborative Consumption Is Changing the Way We Live*. Collins, 2011.
2. Byers, J.W., Proserpio, D., and Zervas, G. The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. Boston University School of Management Research Paper (2013).
3. Eddy, M. German court bans Uber service nationwide. *New York Times Bits Blog* (Sept., 2, 2014).
4. Eisenmann, T., Parker, G., and Van Alstyne, M.W. Strategies for two-sided markets. *Harvard Business Review* 84, 10 (2006), 92.
5. Erickson, A. The birth of zoning codes, a history. *The Atlantic* (June 19, 2012); <http://bit.ly/UxVvx2>.
6. Evans, D.S. Governing bad behavior by users of multi-sided platforms. *Berkeley Technology Law Journal* 27, 2 (2012).
7. Keeley, L. NYC taxi medallions wheel in profits rivaling S&P 500. *Bloomberg* (Aug. 6, 2010).
8. Knox, R. How Yelp can help disease detectives track food poisoning. NPR (May 22, 2014); <http://n.pr/1oC9qPp>.
9. Lanier, J. *Who Owns the Future?* Simon & Schuster, 2013.
10. Luca, M. and Zervas, G. Fake it till you make it: Reputation, competition and Yelp review fraud (2013); <http://hbs.me/Sjz92n>.
11. New York Times Editorial Board. The dark side of the sharing economy. (Apr. 30, 2014); <http://nyti.ms/1fB1JXj>.
12. Rifkin, J. The rise of the sharing economy. *Los Angeles Times* (Apr. 6, 2014); <http://lat.ms/1AqMffY>.
13. Schrage, M. Who do you want your customers to become? *Harvard Business Review Press*, 2012.
14. Sundararajan, A. Trusting the 'sharing economy' to regulate itself. *Economix Blog, New York Times* (Mar. 3, 2014); <http://nyti.ms/1krGSHo>.
15. Tam, D. Amsterdam officially approves new Airbnb friendly laws. 2014; <http://cnet.co/1tTJ2kp>.
16. *The Economist*. The rise of the sharing economy. (Mar. 9, 2013); <http://econ.st/1rwIfEx>.

Arvind Malhotra (Arvind_Malhotra@kenan-flagler.unc.edu) is a T.W. Lewis Scholar and Professor of Strategy and Entrepreneurship at The University of North Carolina Kenan-Flagler Business School, Chapel Hill, NC.

Marshall Van Alstyne (mva@bu.edu) is an associate professor in the department of management information systems at Boston University and a research scientist at the MIT Center for Digital Business; Twitter: InfoEcon.

Copyright held by authors.



DOI:10.1145/2668895

Pamela Samuelson

Legally Speaking Updates on the Intellectual Property Front

Recapping influential U.S. Supreme Court decisions rendered earlier this year.

WHEN INTELLECTUAL PROPERTY lawsuits will have important implications for computing professionals, my Legally Speaking columns often assess the merits of the claims, sometimes before appellate courts have completely resolved the issues. This column discusses three significant rulings rendered in June 2014 in cases discussed in previous columns and provides an update on the state of play for each.

Two were U.S. Supreme Court decisions. First, a unanimous U.S. Supreme Court ruled against a computer-implemented invention in the *Alice v. CLS Bank* case. (My November 2013 column predicted this result.)

In a second, the Court ruled that Aereo infringed copyrights in ABC's broadcast television programs because of Aereo's role in transmitting programs to customers to watch on Internet-connected devices. (My July 2014 column said the case was too close to call, and the Court split 6-3 on it.)

Earlier in June, an appellate court ruled that mass-digitizing books to create a full-text searchable database that enhances public access to information without harming the market for the books is fair use. (My March 2013 column predicted this result.)

For those whose thirst for news on



Attorney Mark A. Perry represented CLS Bank in the *Alice v. CLS Bank* case.

the intellectual property front has not been sated by this quick roundup, I offer below a recap and some commentary on the decisions.

Alice v. CLS Bank

Alice Corp. owned a set of patents directed to a computer-implemented means of mitigating losses in settlements of financial transactions so that money would only flow in appropriate directions when both parties to each transaction had completed their obligations.

Alice asserted that a CLS computer

program infringed these patents. CLS challenged the validity of these patents, relying on recent Supreme Court opinions.

The Court of Appeals for the Federal Circuit (CAFC) initially ruled that Alice's computer-implemented claims were valid. But after rehearing the case with all 10 sitting judges, a divided CAFC ruled that the method and the medium claims were unpatentable under Supreme Court precedents.

However, the court split 5-5 on the system claims. Some judges thought

all three types of claims were invalid; some thought the system claims were patentable because they were more technological than the method or medium claims.

The issue before the Supreme Court was whether Alice's claim were eligible subject matter for patenting. In the U.S., only "machines, manufactures, compositions of matter, and processes" can be patented.

Alice asserted that its system claims were for "machines," its method claims were "processes," and its software-stored-on-a-medium claims were "manufactures" within the meaning of the statute.

Supreme Court precedents state that abstract ideas, mental processes, and scientific principles are ineligible for patenting. In 2010 in *Bilski v. Kappos*, for example, the Court ruled that a method of hedging risks of loss due to price fluctuations for commodities ran afoul of the "abstract idea" exclusion from patenting.

CLS analogized Alice's claims to those in *Bilski*, saying that claims are still too abstract even if they call for implementing a risk management strategy on a general-purpose computer. Alice mainly argued that its claims were not abstract because they were detailed.

The Supreme Court held that implementing an abstract idea on a general-purpose computer does not make it patentable and that when the claimed invention is for an abstract idea, it makes no difference whether the claim is stated in system, method, or medium terminology.

Unfortunately, the Court gave very little guidance about what makes a claim too "abstract" to be eligible for patenting. It only reinvoked its prior software-related decisions, such as the 1972 *Gottschalk v. Benson* decision, which rejected patent claims for a method of transforming binary coded decimals into pure binary form.

Even though several briefs urged the Court to say something about the patentability of computer program innovations, the Court did not directly address this question. However, it suggested that software innovations that improved the functioning of computers or of other technologies would likely be eligible for patents. The Court

Supreme Court precedents state that abstract ideas, mental processes, and scientific principles are ineligible for patenting.

pointed to the patent at issue in *Diamond v. Diehr*, which claimed an improved rubber curing process that utilized a computer program.

Some of the worst software patents out there—those that claim carrying out specific functions on computers or on the Internet—are likely to be invalid after *Alice*. This may discourage some patent-trolling activities.

But U.S. courts have yet to develop workable criteria for sorting out which software innovations are patentable and which are not. Because of this, Alice has cast a shadow over the validity of patents on even highly novel and nonobvious innovations embodied in software.

ABC v. Aereo

Aereo developed a system allowing its customers to select broadcast television programs to watch on Internet-

connected devices. ABC and other broadcasters claimed this infringed their public performance rights in the programs.

The public performance right is defined in the statute to include "transmit[ing] or otherwise communicat[ing] a performance ... to the public, by means of any device or process, whether the members of the public capable of receiving the performance ... receive it in the same place or in separate places and at the same time or different times."

ABC likened Aereo's service to cable television retransmissions of broadcast programming, which clearly fall within the public performance right. Cable companies pay for the privilege of retransmitting programs, and these revenues are an important source of income to broadcasters these days. ABC claimed the entire broadcast television enterprise was in jeopardy if Aereo won the lawsuit.

Aereo argued it was merely supplying equipment to enable fair use copying by its customers, akin to the Sony Betamax video recording device the Supreme Court found non-infringing in 1984. Aereo and some amicus curiae briefs expressed concern that adopting ABC's theory of the public performance right would jeopardize many other technologies such as cloud computing services (for example, for backup copying purposes).

Justice Breyer wrote the majority opinion in *Aereo* that reversed an appellate court decision and ruled in ABC's favor. The Court found Aereo's service to be too similar to cable tele-



Supreme Court Justice Breyer announces the majority opinion in the *ABC v. Aereo* case.

vision retransmissions to be legally distinguishable. Cable television had once been outside that right, but in 1976 Congress redefined the public performance right to encompass cable and cable-like system transmissions.

Because Aereo transmitted the same program to multiple people at the same time, the Court thought the performance was “to the public” within the meaning of the statute. The Breyer opinion asserted the holding against Aereo was sufficiently limited in scope that it would not put other technologies, such as cloud computing services, at risk of infringement claims.

Justice Scalia’s dissent asserted that Aereo did not perform the programs at issue because it did not select the programs to be viewed. Aereo merely provided a technology through which users could select programs to watch.

This dissent expressed concern about the implications of the ruling for innovation in the technology sector. Virtually all other copyright cases have rejected direct liability claims against technology developers, focusing instead on whether the developer induced its users to infringe or materially contributed to user infringements. The dissent said that courts should not fill holes in the statute through “make-shift” reasoning that would lead to many other challenges.

Communications readers can only hope that Justice Breyer’s confidence in the limited scope of the ruling is borne out over time. Unfortunately, Justice Scalia may be right that the Court’s interpretation of the public performance right may give rise to more litigation and to chilling effects on the development of beneficial new technologies.

Authors Guild v. HathiTrust

HathiTrust is a consortium of academic and research institutions. It hosts a digital library of 10 million books digitized from libraries around the world. Many are copies of books that Google scanned from research library collections for its Book Search Project. HathiTrust makes public domain books available for viewing and downloads, but restricts usage of in-copyright books.

The Authors Guild sued HathiTrust and several university part-

Unfortunately, the Court gave very little guidance about what makes a claim too “abstract” to be eligible for patenting.

ners for copyright infringement. HathiTrust asserted it made fair uses of the books: first, to enable data-mining so researchers can discover books in physical library collections that are responsive to their search queries, second, to preserve copies of the works for future generations, and third, to make books more accessible to print-disabled persons.

A trial judge ruled in HathiTrust’s favor. The Guild appealed, arguing the digitization of millions of books could not possibly be fair use because it did not result in the creation of a new work of authorship (such as a biography quoting from a writer’s letters), because the books copied were highly creative, because whole works were systematically copied, and because HathiTrust had preempted the development of licensing markets for these uses.

The Second Circuit Court of Appeals rejected the Guild’s argument. It found the creation of a full-text searchable database of books to be “quintessentially transformative” because of the different purpose it served from the originals. Making books more findable and enabling data-mining research are consistent with the purposes of copyright law, which are to “promote the progress of Science and useful Arts.” (The court had similarly positive reactions to the preservation and enhanced access for the print-disabled purposes of HathiTrust, although it sent the case back to the lower court for fact-finding on one issue.)

Because it was necessary to copy whole works to create such a database, this did not weigh much against fair use. The HathiTrust digital library did not supplant demand for the purchase

of books, which is the main interest that copyrights protect. The Guild’s theories of harm were too speculative to matter.

Still pending before the Second Circuit is the Guild’s separate appeal from a fair use ruling in Google’s favor in a case challenging the latter’s scanning of books from research library collections. (Although the Guild sued Google first, some procedural complications resulted in the *HathiTrust* case being appealed first.)

There are two principal differences between the *Google* and *HathiTrust* cases. One is that Google is a for-profit company whereas HathiTrust is a non-profit organization. The other is that Google provides snippets of text in response to search queries whereas HathiTrust serves up only identifying information and page numbers of books containing relevant information.

The commerciality distinction is unlikely to be determinative, but snippets do display some expression from in-copyright works. The Guild must pin its hopes of winning the appeal on the snippet difference. Google will argue that snippets, like the data-mining in *HathiTrust*, makes books more findable and is more likely to enhance than decrease the market for books, especially since Google provides links to libraries and online stores from which the books can be obtained.

Conclusion

Although the game is over for Alice’s patents and Aereo’s service, the implications of the *Alice* and *Aereo* decisions for other software-related patents and other digital technologies are as yet unclear.

Given the *HathiTrust* decision, it seems likely Google will prevail in its fair use defense. It may, however, take another year or so before the Supreme Court decides whether to review the fair use rulings in the *HathiTrust* or *Google* cases. The legality of mass-digitization of in-copyright works is closer to being resolved in the U.S., but we are not quite there yet. 

Pamela Samuelson (pam@law.berkeley.edu) is the Richard M. Sherman Distinguished Professor of Law and Information at the University of California, Berkeley.

Copyright held by author.

Computing Ethics

Big Data's End Run Around Procedural Privacy Protections

Recognizing the inherent limitations of consent and anonymity.

PRIVACY PROTECTIONS FOR the past 40 years have concentrated on two types of procedural mitigations: informed consent and anonymization. Informed consent attempts to turn the collection, handling, and processing of data into matters of individual choice, while anonymization promises to render privacy concerns irrelevant by decoupling data from identifiable subjects. This familiar pairing dates to the historic 1973 report to the Department of Health, Education

& Welfare, *Records, Computers and the Rights of Citizens*, that articulated a set of principles in which informed consent played a pivotal role (what have come to be known as the Fair Information Practice Principles (FIPPs)) and proposed distinct standards for the treatment of statistical records (that is to say, records not identifiable with specific individuals).

In the years since, as threats to privacy have expanded and evolved, researchers have documented serious cracks in the protections afforded by

informed consent and anonymity.^a Nevertheless, many continue to see them as the best and only workable solutions for coping with privacy hazards.^b They do not deny the practical challenges,

^a For concise explanations for the growing lack of confidence in anonymization and informed consent, see Narayanan and Shmatikov⁷ and Solove, D.J.,¹¹ respectively.

^b As evidenced by their continued centrality in the policy documents under discussion in both the U.S. and Europe and the entrenched role that they play in commercial privacy policies and codes of conduct.



but their solution is to try harder—to develop more sophisticated mathematical and statistical techniques and new ways of furnishing notice tuned to the cognitive and motivational contours of users. Although we applaud these important efforts, the problem we see with informed consent and anonymization is not only that they are difficult to achieve; it is that, even if they were achievable, they would be ineffective against the novel threats to privacy posed by big data. The cracks become impassable chasms because, against *these* threats, anonymity and consent are largely irrelevant.¹

Informed Consent

Long-standing operational challenges to informed consent (“notice and choice”) have come to a head with online behavioral advertising. Companies eager to exploit readily available transactional data, data captured through customer tracking, or data explicitly provided by users have crafted notices comprising a mish-mash of practices and purposes. Even before big data entered common parlance, authors of privacy policies faced profound challenges in trying to explain complex flows, assemblages, and uses of data. Anxious to provide easily digestible accounts of information practices, they have confronted something we have called the *transparency paradox*⁸: simplicity and fidelity cannot both be achieved because details necessary to convey properly the impact of the information practices in question would confound even sophisticated users, let alone the rest of us.

Big data extinguishes what little hope remains for the notice and choice regime. Stated simply, upfront notice is not possible because new classes of goods and services reside in future and unanticipated uses.⁴ Two decades ago, experts were already warning that data mining posed insurmountable challenges to the foundations of emerging privacy law;^{5,9,10} the situation now is worse than they had feared. Even if it were possible, as a theoretical matter, to achieve meaningful notice and render informed, rational decisions concerning our *own* privacy, these decisions would nevertheless affect what companies can then infer about *others*, whether or not these others

have consented. The willingness of a few individuals to disclose information about themselves may implicate others who happen to share the more easily observable traits that correlate with the traits disclosed. We call this the *tyranny of the minority* because it is a choice forced upon the majority by a consenting minority.

How might this happen? Consider the familiar trope about “the company you keep.” What your friends say or do (on social networking sites, for example) can affect what others infer about you.^c Information about social ties, however, is unnecessary when drawing such inferences, as we learned from Target’s infamous pregnancy prediction score. To discover the telltale signs of pregnancy, Target looked over the purchase histories of those few customers who also made use of the company’s baby shower registry. Analysts then employed data mining techniques to isolate the distinctive shopping habits of these women and then searched for similar purchases in other customers’ records to identify those who were likely to be pregnant. Target was thus able to induce a rule about the relationship between certain purchases and pregnancy from what must have been a tiny proportion of all its customers.

When analysts can draw rules from the data of a small cohort of consenting individuals that generalize to an entire population, consent loses its practical import. In fact, the value of a particular individual’s withheld con-

c See, for example, Mislove, M. et al.⁶

Long-standing operational challenges to informed consent (“notice and choice”) have come to a head with online behavioral advertising.

sent diminishes the more effectively a company can draw inferences from the set of people that do consent as it approaches a representative sample. Once a dataset reaches this threshold, the company can rely on readily observable data to draw probabilistic inferences about an individual, rather than seeking consent to obtain these details. This possibility also reveals why the increasingly common practice of vacuuming up innocuous bits of data may not be quite so innocent: who knows what inferences might be drawn on the basis of which bits?

Anonymity

Most online outfits make a serious show about anonymity.^d But when they claim they only maintain anonymous records,¹² they rarely mean they have no way to distinguish a specific person—or his browser, computer, network equipment, or phone—from others. Nor do they mean they have no way to recognize him as the same person with whom they have interacted previously, to associate observed behaviors with the record assigned to him, or to tailor their content and services accordingly. They simply mean they do not rely on the limited set of information commonly conceived as “personally identifiable” (for example, name, Social Security number, date of birth, and so forth), while still employing unique *persistent identifiers*. Hence the oxymoronic notion of an “anonymous identifier”²—more accurately labeled a pseudonym. These identifiers are anonymous only insofar as they do not depend on traditional categories of identity while still serving the function of persistent identification.

Such practices may make it more difficult for someone to show up on a person’s doorstep with a folder full of embarrassing, discrediting, or incriminating facts, but they do nothing to limit the ability of these companies to draw upon this information in shaping a growing share of everyday experiences that take place on these companies’ platforms. Stated differently,

d That this type of anonymity bears little resemblance to the rigorous specifications of anonymity developed by computer scientists is not our concern here; ours is a discussion of the *value* of anonymity evinced by these techniques.

while anonymous identifiers can make it more difficult to use information about a specific user outside an organization's universe, they do nothing to alleviate worries individuals might have about their fates within it—the information they are presented, the opportunities they are offered, or the way they are treated in the marketplace.

Whatever protections this arrangement offers are further undermined by the kinds of inferences companies can draw having discovered patterns in large assemblages of diverse datasets. A company that may have been unable to learn about individuals' medical conditions without matching records across datasets using personally identifiable information may be able to infer these conditions from the more easily observable or accessible qualities that happen to correlate with them.¹³ If organizations become sufficiently confident to act on these uncertain inferences, the ability to draw these inferences will pose as serious a threat to privacy as the increasingly well-recognized risk of de-anonymization. But rather than going to the trouble of attempting to re-associate “anonymized” medical files with specific individuals, companies might instead discover patterns that allow them to estimate the likelihood someone has a particular medical condition. That certain institutions could meaningfully affect a person's experiences and prospects in the absence of identifying information or without violating record-keepers' promises of anonymity defies the most basic intuitions about the *value* of anonymity.

We Are Not Saying...

There is no role for consent and anonymity in privacy protection. Consent and anonymity should not bear, and should never have borne, the entire burden of protecting privacy. Recognizing their limits allows us to assess better where and under what conditions they may perform the work for which they are well suited.

Privacy loses the trade-off with big data. This tired argument misunderstands the nature and value of privacy and mistakes means for ends. Weaknesses in existing *procedures* for protecting privacy do not undercut the viability of privacy itself.

Big data extinguishes what little hope remains for the notice and choice regime.

We need to try even harder to achieve fail-safe anonymization and effectively operationalize notice and consent. Though worthy goals, the practices described here bypass not only weak mechanisms but also defeat the ideal.

What to Do?

Mathematicians and computer scientists will continue to worry about re-identification. Policymakers will continue down the rabbit hole of defining personally identifiable information and informed consent. Social scientists and designers will continue to worry about refining notice and choice. In the meantime, miners of big data are making end runs around informed consent and anonymity.

A lesson may be drawn from biomedicine where informed consent and anonymity function against a rich ethical backdrop. They are important but not the only protective mechanisms in play. Patients and research subjects poised to sign consent forms know there are limits to what may be asked of them. Treatment or research protocols that lie outside the norm or involve a higher than normal risk must have passed the tests of justice and beneficence. In other words, clinicians and researchers must already have proven to their expert peers and institutional review boards that the protocols being administered or studied are of such great potential value to the individual subject or to society that the reasonable risks are worthwhile. Consent forms have undergone ethical scrutiny and come at the end of a process in which the values at stake have been thoroughly debated. The individual's signature is not the sole gatekeeper of welfare.

By contrast, informed consent and anonymity have served as the sole gatekeepers of informational privacy.

When consent is given (or not withheld) or the data is anonymized, virtually any information practice becomes permissible. These procedural mitigations have long relieved decision-makers of the burden of rendering judgment on the substantive legitimacy of specific information practices and the ends that such practices serve. It is time to recognize the limits of purely procedural approaches to protecting privacy. It is time to confront the substantive values at stake in these information practices and to decide what choices can and cannot legitimately be placed before us—for our consent. **C**

References

1. Barocas, S. and Nissenbaum, H. Big data's end run around anonymity and consent. In *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. J. Lane, V. Stodden, S. Bender, and H. Nissenbaum, Eds. Cambridge University Press, NY, 2014.
2. Barr, A. Google may ditch 'cookies' as online ad tracker. *USA Today* (Sept. 17, 2013).
3. Cate, F.H. The failure of fair information practice principles. In *Consumer Protection in the Age of the "Information Economy"*. J.K. Winn, Ed. Ashgate, Burlington, VT, 2006, 341–378.
4. Cate, F.H. and Mayer-Schönberger, V. Notice and consent in a world of big data. *International Data Privacy Law* 3, 2 (May 20, 2013), 67–73.
5. Klösgen, W. KDD: Public and private concerns. *IEEE Expert: Intelligent Systems and Their Applications* 10, 2 (Feb. 1995), 55–57.
6. Mislove, M. et al. You are who you know: Inferring user profiles in online social networks. In *WSDM '10 Proceedings of the Third ACM International Conference on Web Search and Data Mining*. ACM, NY, 2010, 251–60; DOI: 10.1145/1718487.1718519.
7. Narayanan, A. and Shmatikov, V. Myths and fallacies of 'personally identifiable information'. *Commun. ACM* 53, 6 (June 2010), 24; DOI: 10.1145/1743546.1743558.
8. Nissenbaum, H. A contextual approach to privacy online. *Daedalus* 140, 4 (Oct. 2011), 32–48; DOI: 10.1162/DAED_a_00113.
9. O'Leary, D.E. Some privacy issues in knowledge discovery: The OECD personal privacy guidelines. *IEEE Expert: Intelligent Systems and Their Applications* 10, 2 (Apr. 1995), 48–59.
10. Piatetsky-Shapiro, G. Knowledge discovery in personal data vs. privacy: A mini-symposium. *IEEE Expert: Intelligent Systems and Their Applications* 10, 2 (Apr. 1995), 46–47.
11. Solove, D.J. Privacy self-management and the consent dilemma. *Harvard Law Review* 126, 7 (May 2013), 1880.
12. Steel, E. and Angwin, J. On the Web's cutting edge, anonymity in name only. *The Wall Street Journal* (Aug. 4, 2010).
13. Walker, J. Data mining to recruit sick people. *The Wall Street Journal* (Dec. 17, 2013).

Solon Barocas (sib237@nyu.edu) is a postdoctoral research associate at the Center for Information Technology Policy at Princeton University.

Helen Nissenbaum (helen.nissenbaum@nyu.edu) is a professor of media, culture, and communication at New York University.

An expanded version of the arguments presented in this Viewpoint appears in Barocas and Nissenbaum.¹ The authors thank Arvind Narayanan and an anonymous reviewer for their helpful feedback and gratefully acknowledge research support from the Intel Science and Technology Center for Social Computing and NSF awards DGE-0966187 and CNS-1355398.

Copyright held by authors.

Education

A Future for Computing Education Research

Seeking to address the most important issues facing the computer education research community.

ONE DECADE AGO, NSF funded two efforts to create a computing education research (CER) community in the U.S. and Europe. Nicknamed “Bootstrappers” and “Scaffolders,”^a these two projects helped create a CER community among existing computing faculty who had an interest in computing education. More than 40 computer science faculty became active in CER. Many publications have resulted, the ICER conference has been created and established, and many of those faculty members now hold leadership positions within the computing education community. What the Bootstrappers and Scaffolders did not focus on was growing the number of Ph.D. students in CER, and trying to establish the legitimacy of CER as a research discipline within computing departments at research-oriented schools. (It is difficult to place CER within schools of education, as computing is not a required component of primary or secondary schools.) With the exploding interest in computing education and technology, with the awareness of the challenges in computing education at the K–12 level,

with NSF’s recent push toward disciplinary-based education research in STEM fields,³ and with a huge growing interest in offering computer science courses via distance, NSF funded another pair of workshops. The purpose of the first summit was to allow the CER community to identify the most important research questions facing the discipline. The second summit focused on identifying and developing strategies to overcome structural impediments to supporting CER at the university level.

This column focuses on the first summit: the most important research questions facing CER. Interested researchers were invited to submit a white paper addressing key research questions likely to be faced by CER over the next five years. The accepted white papers (available from <http://purl.stanford.edu/mn485tg1952>) asked research questions primarily in five areas: broadening participation, computing in K–12 (primary and secondary), computing in STEM education, students and learning issues, and tools. A few papers were written at a meta-level, looking at the direction of CER and its impact, and looking more broadly at challenges facing the teaching of computing at all levels. Here, we summarize across the five research areas while highlighting the two most compelling agendas that currently look to

be in contention in terms of garnering researchers’ attention.

Broadening Participation

A critical issue for CER is “computing for all.” Not unlike other STEM disciplines, computing suffers from underrepresentation of women and minority groups. Understanding why there is such little diversity in computing is a critical area of research in CER, but even more critical is research into *interventions* that make computing available for females, underrepresented minorities, low-income students, students in community colleges, as well as students with disabilities. The focus on increasing access to computing moves a lot of attention in CER into primary and secondary school education. It is the level of schooling where we can reach “all.”

In computing education, we tend to focus on coming to understanding algorithms and information representations. The summit focused at a different level. Computing education researchers study ways to better understand how conflicting cultural values and self-belief that impact motivation, and tackling these through design practices and tools that foster meaningful learning. Because we deal with such a flexible medium that can easily shift modalities, there is no reason why computing tools and

^a There was a third, unfunded, project nicknamed Bracelets, which focused on growing CER within Australasia.

programming environments are not more accessible to those with disabilities. Because computing is typically an elective subject in primary and secondary school, computing education researchers look for any way they can to slide computing into schools, including integrating through humanities (including creative writing, theatre, and media production), through science classes, and through mathematics classes.

The focus in CER on broadening participation makes CER different from other discipline-based education research, like physics education. Our first keynote at the summit was from Nobel laureate Carl Wieman, who argued for rigorous research in physics (and in computing) education. Wieman noted that physics education research has largely stayed away from research in K-12, due to the difficulties working in this space. But the motivation to diversify physics education

is not as strong as it is in CER. Computing is even less diverse than physics. The economic value for computing is even greater than physics, so the lack of access to computing education is an even greater cost to underrepresented groups than is a lack of access to physics education.

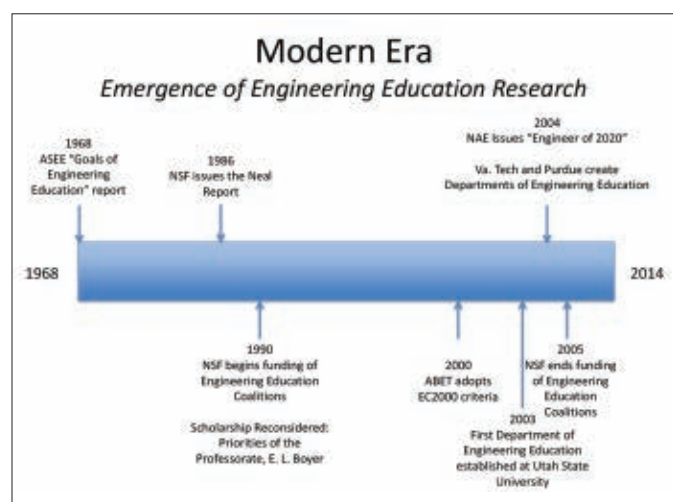
Lance Pérez offered a keynote on the history of engineering education research, and lessons CER could learn. In particular, Pérez noted that engineering education has largely advocated for separate engineering education departments, leading to the conflicting needs to both be strong teachers (of introductory engineering curricula) as well as effective engineering education researchers. He advised computing education researchers to develop a model that avoids the problems that engineering education has faced.

Research in computing education must necessarily incorporate research

into gender studies and other non-majority perspectives. Perspectives and experiences are shaped by the constant navigation of multiple shifting identities of gender, race, class, religion, and ability. Computing education researchers were urged to disaggregate empirical data in CER based on race, ethnicity, gender, and socioeconomic status. Studies at the intersection of race, ethnicity, and gender are needed, to better understand why certain groups choose to enter and persist in computing while others do not. We don't come to understand others' perspectives by counting them, so CER necessitates the inclusion of qualitative research methodologies (such as interviews and case studies), so the richness of personal narratives may inform action.

Fundamental Issues of Computing Knowledge

Computing education is a newcomer



Educational research

- As Carl Wieman said yesterday, we do educational research to make a difference to education – there's no other reason.
- IMO - In Computing learning - there are going to be no bigger issues than the expansion of the student body and characterising and improving the capacity and quality of Computing teaching.
- Your thoughts?

University of Kent

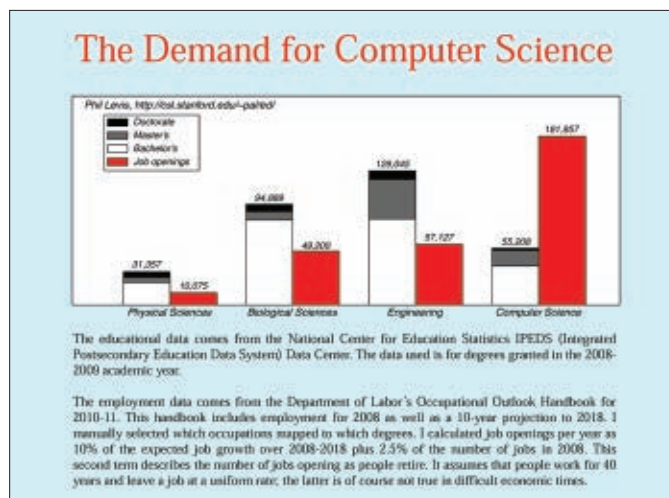
Learning in the in classroom*

Comparing the learning in two identical sections of 1st year college physics. 270 students each.

Control--standard lecture class-- highly experienced Prof with good student ratings.
Experiment-- inexperienced teacher (postdoc) trained to use principles of effective teaching.

- Same learning objectives
- Same class time (3 hours, 1 week)
- Same exam (jointly prepared)- start of next class

*Deslauriers, Schewlew, Wieman, Sci. Mag. May 13, '11



Samples from keynote presentations at the Future Directions in Computing Education Summit Meeting held earlier this year.

to the ranks of discipline-based education research. Mathematics education researchers know a lot about how students come to understand numbers and sets. Physics education researchers have detailed research into the misconceptions that students develop about the physical world from living in it without measuring it. But we know relatively little about the foundations of computing knowledge, and these are more critical issues as we move computing into primary and secondary schools. Foundational questions include “How do students learn to program and what does that development look like?” “What are successful and unsuccessful mental models of challenging computing concepts?” “How do we support successful transfer from beginner programming environments to real-world ones?” and “What are common challenges in conceptual understanding in computing course?” were listed among the many key questions the learning sciences can help address. Even though these topics have been studied extensively since the 1980s, many questions remain unanswered and would benefit from contemporary research in the learning sciences in socio-cultural and situated learning, distributed and embodied cognition, as well as activity, interaction and discourse analyses.²

Researchers are working to define and validate learning progressions for K–12 computing. Some researchers are studying how people “become computer scientists” and to investigate the broader meanings of the computing discipline. The use of design-based research and design-based implementation research is common in the community to support research and assessment in naturalistic settings.

A more critical issue for CER today is how to develop enough teachers to support computing education in primary and secondary schools. We need to support the professional growth of computing teachers, and in parallel, build a robust community of computing teachers. How, for example, will the U.S. meet the CS10K challenge¹ by 2016? A major focus for research in computing education is, “What do we teach

computing teachers?” What do computing teachers need to know that’s *different* from other STEM teachers? That discipline-specific knowledge of teachers is called *Pedagogical Content Knowledge*.⁴ Defining PCK for computing and figuring out how to teach it are major thrusts in CER today. Computing education researchers are developing just-in-time teacher professional development to help teachers communicate and learn best practices, robust assessments, and the latest introductory programming.

Sally Fincher’s keynote focused on the needs of teachers in order to improve computing education. How do we make our computing education research *matter* and influence *practice*? She pointed out that many of the innovations in teaching reading failed because they changed reading in ways that was dissimilar to practice. We need to ensure that we in computing education are focusing on what the students *actually* need to know, or we risk becoming ineffective and useless.

Like other research areas, the CER community is asking itself if “big data” can help with our issues too. Some researchers are calling for data and evidence-driven analyses of student learning as part of adding rigor to drive practice in CER, including the use of big data to develop scalable methods for automated assessments (along the lines of emergent forms of assessment in MOOCs). The hope is to leverage big data to understand better the nature of learning programming and computational problem solving, the problems learners encounter, and the pedagogies and strategies that can be used to address them. Researchers are using evidence-centered design and data mining to study both cognitive and non-cognitive aspects of computational thinking.⁵

Moving Forward with CER

The summit exposed some of the tensions between research groups in CER about what are the most important research areas to be addressing. A large contingent whose research focused on issues of equity argued that all research in CER, whether content pedagogy, culture, and so forth needed to be viewed through an equity lens. Others argued for CER to focus on more foundational

issues, such as learning progressions and integration of computing content and computational thinking into math, science, and non-STEM classes.

Tensions in CER around research methods mirror what is going on in other STEM education research areas. Traditionally, much of CER focused on social science methodologies. New researchers to CER are using big data analytic techniques, and borrowing approaches from other computing disciplines, most notably machine learning. How can these “newer” techniques (for example, looking at click-through rates, intermediate forms of student artifacts, and other forms of “big data”) be used to answer our questions? Do they substitute for our other research approaches, or do we use them in addition to more traditional approaches?

Computing education research is unique among the STEM education research fields. The economic value of knowing computing is greater than any other STEM field, but computing is the least diverse of all the STEM fields. Thus, too many people are losing out on the advantages of computing. We have to figure out how to fix that, while still answering the fundamental questions of our discipline. How do people come to understanding computing, and how can we make it better? **C**

References

1. Cuny, J. Finding 10,000 teachers. *CSTA Voice* 5, 6 (2010), 1–2.
2. Grover, S. and Pea, R. Computational thinking in K–12: A review of the state of the field. *Educational Researcher* 42, 1 (Jan./Feb. 2013), 38–43.
3. National Research Council. *Discipline-based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. The National Academies Press, Washington, D.C., 2012.
4. Shulman, L. Those who understand: Knowledge growth in teaching. *Educational Researcher* 15, 2 (Feb. 1986), 4–31.
5. Wing, J. Computational thinking. *Commun. ACM* 49, 3 (Mar. 2006), 33–35.

Steve Cooper (coopers@cs.stanford.edu) is an associate professor in the computer science department at Stanford University, Stanford, CA.

Shuchi Grover (shuchig@stanford.edu) is a research scientist at SRI International and a visiting lecturer in the computer science department at Stanford University, Stanford, CA.

Mark Guzdial (guzdial@cc.gatech.edu) is a professor in the College of Computing at Georgia Institute of Technology in Atlanta, GA.

Beth Simon (bsimon@eng.ucsd.edu) is the senior associate director of Learning Sciences and Technology the Center for Teaching Development and tenured faculty in the Computer Science and Engineering Department at the University of California, San Diego.

Copyright held by authors.

Privacy and Security

Summing Up

Considering recent privacy- and security-related events through this column.

THE PRIVACY AND SECURITY Viewpoints column had its start six years ago. Though our 25 columns have had 23 different sets of authors, the fact is that an editor puts an imprint on a column simply by selecting potential contributors. No one should hold sway for too long, and so this column is the last under my leadership as section board member for the Privacy and Security column. It is a good time to ask how privacy and security have fared over the last half-dozen years.

The short answer is that for privacy and security, this is not the best of times. If privacy and security were students in a course, right now their grades would be somewhere between an F and an F-minus (for those readers unfamiliar with the U.S. academic system, an “F” for fail is as low as you can go). Consider just the events of the past two years. We have had the Snowden leaks, which show the U.S. National Security Agency (NSA) partnering with the intelligence agencies of the other “Five Eyes” (Australia, Canada, New Zealand, and U.K.) vacuuming up metadata (signals about communications), and, in many instances, heavily collecting content as well. The Snowden leaks also demonstrate a breach of security within the NSA and its contractors. Under legal requirements, the British telecommunications company Vodafone must provide the governments of six unnamed nations with direct access to *all* communications.⁷ U.S. retailer Target ignored the warnings of breach behavior from its security firm, FireEye, resulting in the theft of 40 million



credit and debit cards (more correctly, data sufficient to forge 40 million credit and debit cards).¹³ Meanwhile Facebook has a new product that records ambient sounds on user request—thus not only capturing lots of information about the enrolled user but also capturing conversations of unsuspecting bystanders. Google is strongly pressing users to have an always signed-on experience, particularly on their smartphones.¹¹

The long answer is considerably more nuanced. Yes, the Snowden leaks

demonstrated a level of collection that only the more paranoid among us had anticipated. The continuing revelations, including the amount of collection by the U.K.’s General Communication Headquarters, France’s DGSE, and Chinese and Russian signals intelligence organizations, demonstrated that the U.S. was far from alone in conducting such massive surveillance. Yes, the wide usage of Facebook, the lack of user response, in terms of a drop-off in usage, to changes in the Google privacy policies, the willingness of users

ACM Journal on Computing and Cultural Heritage



JOCCH publishes papers of significant and lasting value in all areas relating to the use of ICT in support of Cultural Heritage, seeking to combine the best of computing science with real attention to any aspect of the cultural heritage sector.



www.acm.org/jocch
www.acm.org/subscribe



Association for
Computing Machinery

to avail themselves of apps that downloaded user location, whether or not such information was needed by the app, demonstrated that users continued to be generally unconcerned about protecting their privacy. But subtle signs indicate these signals are not all there is to the present privacy and security stories.

Consider the fact that Gregg Steinhafel, Target's CEO, was forced to resign within months of the breach. This is significant. Companies have had major breaches for years, but cybersecurity had previously never played a role in a senior executive's resignation. Now it has. Cybersecurity has suddenly become a lot more important in the C suite.

Consider the sudden success of SnapChat, the app that—incorrectly, as it turned out⁵—claimed it deleted photos seconds after their appearance on a recipient's screen. Someone—actually lots of someones—wanted such an app. Share a photo, but have it be ephemeral. There was interest in having email and text be ephemeral too. This is a desire to return to have communications largely be the way they were before digital transmission and storage made the most trivial utterance permanent.

Finally, consider the impact of the Snowden leaks on the U.S. government, which is now having serious, somewhat public, discussions about limits on collection. In January, President Obama announced he would move bulk metadata collection out of the government's hands.^a The president also committed to restraining the use of communications collected for foreign-intelligence purposes in criminal cases and to restricting the length of time a National Security Letter, which "require[s] companies to provide specific and limited information to the government without disclosing the orders to the subject of the investigation" can be kept secret. There are some that would say the president did not go far enough in curbing intelligence collection. But these are substantive changes, and there may be even more forthcoming.

^a "Remarks by the President on Review of Signals Intelligence," January 17, 2014; <http://1.usa.gov/1awEWY8>.

Thus that failing grade for privacy and security may not be an accurate description of where we are on privacy and security, though we do certainly appear to be at a nadir. Where we are headed is more than a little difficult to discern. Before I respond to that issue, we must first understand how we arrived at the present point. As security and privacy present substantively different issues, I will address them separately.

For security, partially the problem arose from the mistaken belief that security was all about confidentiality, integrity, and availability, and partially from the belief that security was merely a technical issue. But while the fact that both of these assumptions were wrong was obvious at least a dozen years ago, developing more complex responses to computer security was not easy. It may be difficult to design a secure cryptosystem; it has proven much harder to develop incentives that cause companies to adopt secure systems. The former is purely a technical problem,^b the latter requires developing an economic model, perhaps including government regulations or laws, to pressure an entity into adopting a more secure—and likely more expensive—system.

For privacy, the situation is even more complicated. The first need was developing an understanding of privacy in the electronic age. In the 1960s, Alan Westin defined privacy as the ability "to determine for themselves when, how, and to what extent information about them is communicated."¹⁷ In the 1970s, the "Ware Re-

^b I am describing this as purely technical because I am focusing on the mathematical algorithm and not the protocol that implements it.

**It is a good time
to ask how privacy
and security have
fared over the last
half-dozen years.**

port” operationalized and extended this into the Fair Information Practice Principles, which have been widely adopted throughout the world.¹⁶ But the principles of “Collection Limitation, Data Quality, Purpose Specification, Use Limitation, Security Safeguards, Openness, Individual Participation, and Accountability”⁸ do not cover a world in which a user may install hundreds of apps on her smartphone, each of which requests access to multiple permissions. And that is only the phone: the user also must confront her work desktop and its browsers, her shared family tablet, her laptop, and ...

This brief reprise of the current communication and computing world tells us that handling security and privacy involves understanding economics (for developing incentive schemes to build security into applications and to help users manage their privacy if they so desire), anthropology (for understanding the different ways people approach their electronic gadgets), psychology (for developing models that give people the computer interactions they want), design (for getting the colors, sizes, and signals right for people to achieve what they want), law (for balancing different interests in society), and so forth. While human-computer interaction has been an established field for at least three decades, the Workshop in Economics of Information Security is only in its 13th year and the Symposium on Usable Privacy and Security is just 10 years old. And while there is a Privacy Legal Scholars Conference that focuses, to a large extent, on digital privacy, there is no equivalent Cybersecurity Legal Scholars Conference. The point? Scholarship in these human aspects of privacy and security is nascent. That means translation into actual use is even more so.

Where does that put us? Being always connected is increasing and we are moving to the Internet of Things; both of these will create additional privacy and security risks. It is likely that for many people—especially those living under repressive regimes and those with less opportunity either as a result of economics or because of a lack of capabilities—privacy and security incursions will mount. In such

I think we have done a reasonable job covering a broad and complex set of topics.

situations, digital communications technologies appear to be leading to increasingly oppressive and invasive situations for many people.

On the other hand, there are changes in the air. danah boyd describes how American teens have learned to hide their actions in plain sight,¹ the U.S. government is seriously considering certain limitations in its collection of communications, both domestically and abroad, the Europeans are attempting to put limits on what information is available about individuals (though the European Court decision on “the right to be forgotten” decision⁴ seems to me to confuse data being accessible on the network—the real issue—with data being linked through search engines). There appears to be considerably more interest by private individuals in securing and protecting their data; Tor usage, for example, is approximately twice what it was in May 2013.^{c,15}

Whether the latter will be a lasting change is, of course, crucial. It is a question I cannot answer. But I can address how well this column has handled broadly addressing the issues surrounding privacy and security over the last six years. And there I think the answer is, “Reasonably well.”

Long before the Snowden leaks, we had columns on electronic surveillance in Sweden⁹ and a proposed code of ethics for U.S. intelligence officers.¹⁴ We had more technical columns focused on de-anonymizing anonymized networks¹² and on why air gaps do not work for SCADA networks.³ We have had some on the so-

cial science side, including on the role of emotions in making complex security decisions,¹⁰ and on the impact that Fear, Uncertainty, and Doubt play in determining how to fight back against computer crime.⁶ We have covered the value, or lack thereof, of professionalizing the cybersecurity workforce.² We have had researchers, implementers, engineers, computer scientists, lawyers, social scientists of many flavors, and management experts write—and there have been authors from three continents and at least 27 different policy persuasions. I think we have also done a reasonable job covering a broad and complex set of topics.

With this, I hand the mantle of *Communications Privacy and Security* columns to Carl Landwehr, for 30 years a leading cybersecurity researcher. Carl is a National Cyber Security Hall of Fame inductee and a former editor-in-chief of *IEEE Security and Privacy*. Many thanks for the pleasurable run. 

References

1. boyd, d. *It's Complicated*. Yale University Press, 2014.
2. Burtley, D., Eisenberg, J., and Goodman, S. Would cybersecurity professionalization help the cybersecurity crisis? *Commun. ACM* 57, 2 (Feb. 2013), 24–27.
3. Byres, E. The air gap: SCADA's enduring security myth. *Commun. ACM* 56, 8 (Aug. 2013), 29–31.
4. European Court, Judgement of the Court, Grand Chamber. *Google Spain SL Google Inc. v. Agencia Espanola Proteccion de Datos Mario Costeja Gonzalez*, May 13, 2014; <http://bit.ly/1prYAfk>.
5. Federal Trade Commission. SnapChat Settles FTC Charges that Promises of Disappearing Messages were False. Press Release, May 8, 2014.
6. Florencio, D., Herley, C., and Shostack, A. FUD: A plea for intolerance. *Commun. ACM* 57, 6 (June 2014), 31–33.
7. Garside, J. Vodafone reveals existence of secret wires that allow state surveillance. *Guardian* (June 5, 2014).
8. Gellman, R. Fair Information Practices: A History, Version 2.11 (Apr. 2014); <http://bobgellman.com/rg-docs/rg-FIPShistory.pdf>.
9. Irion, K. Communications networks tapped for intelligence-gathering. *Commun. ACM* 52, 2 (Feb. 2009), 26–28.
10. McDermott, R. Emotion and security. *Commun. ACM* 55, 2 (Feb. 2012), 35–37.
11. Mirani, L. Google's sneaky new privacy change affects 85% of iPhone users—but most of them won't have noticed. *Quartz* (Apr. 3, 2014); <http://bit.ly/1fKRDB2>.
12. Narayanan, A. and Shmatikov, V. Myths and fallacies of “personally identifiable information”. *Commun. ACM* 53, 6 (June 2010), 24–26.
13. Riley, M., Elgin, B., Lawrence, D., and Mattlack, C. Missed alarms and 40 million stolen credit card numbers: How target blew it. *BloombergBusinessweek Technology*, (Mar. 14, 2014).
14. Snow, B. and Brooks, C. An ethics code for U.S. intelligence officers. *Commun. ACM* 52, 8 (Aug. 2009), 30–32.
15. Tor Metrics Portal, Users; <http://bit.ly/1oM0zJA>.
16. U.S. Secretary's Advisory Committee on Automated Personal Data Systems. *Records, Computers and the Rights of Citizens*, 1973.
17. Westin, A. *Privacy and Freedom*, Atheneum, 1967.

Susan Landau (susan.landau@privacyink.org) is a professor of cybersecurity policy in the Department of Social Science and Policy Studies at Worcester Polytechnic Institute in Worcester, MA.

Copyright held by author.

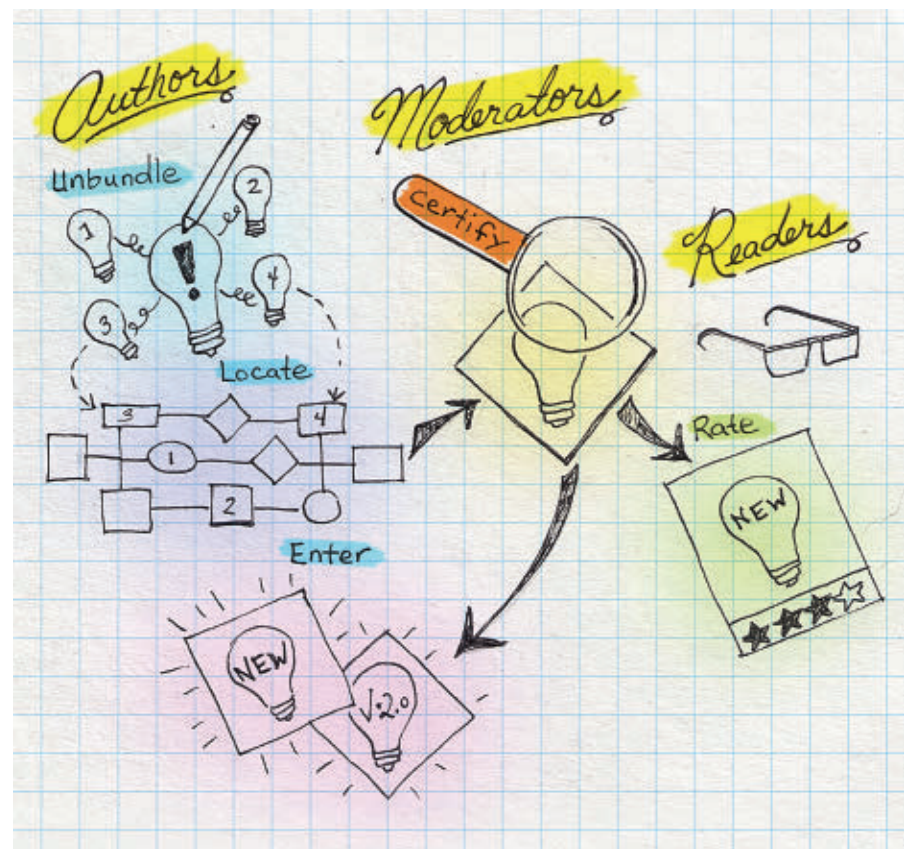
c The reason for this is not entirely clear, since a major spike occurred in August 2013; numbers are substantially down from that high point.

Viewpoint

An Embarrassment of Riches

A critical review of open innovation systems.

OPEN INNOVATION SYSTEMS represent an emerging collective intelligence success story. In such systems, a customer describes a problem they want to solve (for example, “we want ideas for new beverage products”) and provides an online tool that allows the crowd to submit proposed solutions, as well as rate (and sometimes critique) other people’s proposed solutions. Many open innovation platforms have emerged (such as ideascap, spigit, and imaginatik) and have been used widely in contexts that range from IBM to Starbucks, from the Danish central government to the White House. One recent survey² found that one in four companies plan to utilize open innovation systems within the next 12 months, and this figure is growing. Such systems have proven they can elicit substantive contributions at a very large scale and very low cost. In the early weeks of his first term, for example, President Obama asked U.S. citizens to submit and vote on questions on the website change.gov, and promised to answer the top five questions in each category in a major press conference. This initiative engaged over 100,000 contributors, who submitted over 70,000 questions and four million votes. Google’s 10 to the 100th project received over 150,000 suggestions on how to channel Google’s charitable contributions. In IBM’s Idea Jam in 2006, 46,000 ideas for possible IBM products and services were generated by 150,000 contributors. Such large-scale participation en-



A deliberation map¹ can help enable better organizational decision making.

ables, in turn, such powerful emergent phenomena as:

- *The long tail:* Crowds can generate a much *greater diversity* of ideas, including potentially groundbreaking “out of the box” contributions, than we could easily access otherwise.

- *Idea synergy:* Crowds can rapidly develop huge volumes of novel ideas by recombining and refining the ideas

proposed by other participants.

- *Many eyes:* Crowd participants can check and correct each other’s contributions, enabling remarkably high-quality results very inexpensively.

- *Wisdom of the crowd:* Crowds can collectively make better judgments than the individuals that comprise them, often exceeding the performance of experts.

Open innovation systems face, however, serious challenges that, paradoxically, are largely a result of how *successful* they have been at eliciting huge volumes of participation. In this Viewpoint, we review these challenges and propose some promising directions for moving forward.

Challenges

The challenges faced by open innovation systems occur both with idea generation and idea evaluation. Key challenges with *idea generation* include:

- *Harvesting costs*: Open innovation engagements tend to generate idea corpuses that are large, disorganized, and highly redundant. Pruning such a list to find the best ideas can be a massive undertaking. Google's 10 to the 100th project, for example, had to engage 3,000 employees to prune the 150,000 ideas they received, putting them nine months behind schedule. IBM flew 100 senior executives into New York from around the world to prune the results of their Idea Jam.

- *Unsystematic coverage*: Open innovation systems have no inherent mechanism for ensuring the ideas submitted comprehensively cover the most critical facets of the problem at hand, so the coverage is hit-or-miss and may not align with the customer's needs.

- *Shallowness*: Open innovation systems tend to generate large numbers of relatively shallow ideas. A major reason for this, we believe, is that collaborative idea development, and accurate credit assignment, is typically not well supported in current tools.

Open innovation systems also face challenges with crowd-sourced *idea evaluation*: there is often a disconnect between what the customer wants and what the crowd selects. This can occur for several reasons:

- *Shallow evaluations*: Little support is provided for the crowd building upon each other's evaluative expertise, since users usually do not examine and correct each other's facts and reasoning.

- *Rating lock-in*: When there are thousands of ideas, many potentially valuable ideas may not end up being evaluated in sufficient depth, and the system can quickly "lock" into a fairly static, and arbitrary, ranking, where the winning ideas are inferior to others in the list.

The challenges faced by open innovation systems occur both with idea generation and idea evaluation.

Open innovation systems thus face critical challenges in terms of ensuring the potentially massive contributions of the crowd provide high value to the customer without incurring prohibitive harvesting costs.

Promising Directions

How can we meet these challenges and more fully achieve the promise of open innovation systems? Progress will require, we believe, advances on the following two key fronts.

Better open innovation processes.

New open innovation processes are needed that provide more guidance about how the crowd can best contribute, help crowd members build on each other's inputs, and make it easier to harvest their contributions,

- *Collaborative idea definition*: Helping the crowd make more deeply considered contributions will require progress on incentive schemes and collaborative authoring structures. Participants, for example, can be asked to structure their contributions as *deliberation maps*¹—as trees made up of problems to solve, potential solutions for these problems, and the arguments for and against each potential solution, all single-authored. Participants can then compose proposals from the best solution ideas in each map. Credit assignment becomes straightforward because each proposal is built from components with clear authorship.

- *Novel rating mechanisms* can help ensure the crowd evaluates ideas quickly and accurately with respect to the criteria the customer cares about. One possibility, for example, is to use a kind of *prediction market* where par-

Calendar of Events

November 16–19

Interactive Tabletops and Surfaces, Dresden, Germany, Sponsored: SIGCHI, Contact: Raimund Dachzelt, Email: raimund.dachzelt@tu-dresden.de

November 16–22

22nd ACM SIGSOFT Symposium on the Foundations of Software Engineering, Hong Kong, China, Sponsored: SIGSOFT, Contact: Shing-Chi Cheung, Email: sccheung@cse.ust.hk

November 30–December 2

13th ACM International Conference on Virtual Reality Continuum and Its Applications in Industry, Shenzhen, China, Sponsored: SIGGRAPH, Contact: Jin Xiaogang, Email: jin@cad.zju.edu.cn

December 2–5

Conference on Emerging Networking Experiments and Technologies, Sydney, Australia, Contact: Aruna Seneviratne, Email: a.seneviratne@unsw.edu.au

December 8–12

15th International Middleware Conference, Bordeaux, France, Contact: Laurent Reveillere, Email: reveillere@labri.fr

December 13–17

47th Annual IEEE/ACM International Symposium on Microarchitecture, Cambridge, U.K., Sponsored: SIGMICRO, Contact: Krisztian Flautner, Email: krisztian.flautner@arm.com

January 8–10

9th International Conference on Ubiquitous Information Management and Communication, Bali, Indonesia, Sponsored: SIGAPP, Contact: Sukhan Lee, Email: lsh@ece.skku.ac.kr

ticipants are given a limited budget to buy and sell stocks, each representing a different idea, and get a payoff when ideas they “bought” are successfully implemented by the customer. This gives users incentives to evaluate ideas carefully from the customer’s perspective. Deliberation maps can also help here, by allowing crowds to check and build upon each other’s reasoning by creating chains of supporting and rebutting arguments.

► *Creativity enhancement techniques* (which have been developed, to date, almost exclusively for face-to-face team settings) can be used offline to feed open innovation engagements as well as adapted to crowd-scale online contexts.

► *Interleaving ideation and evaluation* across multiple open innovation rounds. The crowd can be asked to create new ideas built upon those that survived the previous round of selection, so idea generation is more likely to focus on what the customer wants.

Deeper computer support. Crowds (of people) and clouds (computers) have synergistic capabilities. Crowds are able to create, understand and evaluate ideas in ways that computers cannot match, but are best suited for performing relatively small and quick tasks that require little context. Computers, by contrast, excel at rapid analysis of large swaths of data to get “the big picture” of what is (and is not) happening in a crowd. Combining these strengths will require bridging the semantic gap between the natural language that crowds use, and the formal languages that computers require. Someday, this will be achieved by advanced algorithms that allow computers to deeply understand natural language. But this achievement seems to remain far off. In the meantime our goal, we believe, must be to find ways that crowds can do the *minimum formalization* needed to enable significant computer support, for example by:

► *Semantic tagging:* Crowds can annotate natural language idea corpuses with semantic cues (for example, idea and argument boundaries, topic keywords). This can be a mixed initiative process, wherein computers propose possible tags that are corrected by

As the semantic gap between crowds and clouds narrows, we can create powerful new forms of computer support for open innovation.

crowd members, and where machine learning can be used to improve the computer algorithms over time based on this human feedback.

► *Design tools* can allow users to express their ideas as semi-formal models built on domain-specific primitives, rather than just as natural language text. Design tools aimed at the masses (for example, Google sketchup) are already becoming ubiquitous, but have yet to be incorporated into open innovation platforms.

As the semantic gap between crowds and clouds narrows, we can create powerful new forms of computer support for open innovation, such as:

► *Analysis tools* that take advantage of semi-formal idea representations to help evaluate the strengths and weaknesses of contributed ideas;

► *Semantic compression algorithms* that remove duplicates, and cluster related ideas, in order to compress idea corpuses; and

► *Visualization tools* that summarize what the crowd has done so far, so customers can determine what the gaps/promising areas are and use this information to guide future crowd contributions, for example via focused incentives.

A Call to Arms

Open innovation systems, as we have seen, have the potential to harness the collective intelligence of the crowd for problem solving in areas ranging from business to government, from science to education. This potential

is far from fully realized, however, largely because of our inability to deal effectively with the massive levels of user contributions that these systems can elicit. Advances in this area will require contributions from many disciplines, including computer science, cognitive science, social psychology, computational linguistics, and economics. Will you join us in addressing these important challenges? **C**

Further Reading

Bailey, B.P. and Horvitz, E.
What's your idea? A case study of a grassroots innovation pipeline within a large software company. In *Proceedings of CHI 2010*, ACM Press, NY, 2010.

Bason, C.
Leading Public Sector Innovation: Co-creating for a Better Society. Policy Press, 2010.

Bjelland, O.M. and Chapman Wood, R.
An inside view of IBM's innovation jam. *MIT Sloan Management Review* 50, 1 (2008), 32–40.

Chesbrough, H., Vanhaverbeke, W., and West, J., Eds.
Open Innovation: Researching a New Paradigm. Oxford University Press, Oxford, U.K., 2006.

Gulley, N.
Patterns of innovation: A web-based MATLAB programming contest. *Human Factors in Computing Systems* (2001), 337–338.

Jouret, G.
Inside Cisco's Search for the Next Big Idea. *Harvard Business Review* 87, 9 (2009), 43–45.

Lakhani, K.R. and Jeppesen, L.B.
Getting unusual suspects to solve R&D puzzles. *Harvard Business Review* 85, 5 (2007), 30–32.

von Hippel, E.
Democratizing Innovation. MIT Press, 2005.

References

1. Klein, M. and Iandoli, L. Supporting collaborative deliberation using a large-scale argumentation system: The MIT collaboratorium. *Directions and Implications of Advanced Computing: Conference on Online Deliberation (DIAC-2008/OD2008)*. University of California, Berkeley, 2008.
2. Thompson, V. *IDC MarketScape: Worldwide Innovation Management Solutions 2013 Vendor Analysis*, (2013); http://idcdocserv.com/240823_spigit.

Mark Klein (m_klein@mit.edu) is a principal research scientist at the MIT Center for Collective Intelligence, an affiliate at the MIT Computer Science and AI Lab, the New England Complex Systems Institute, and the Dynamic and Distributed Information Systems Group at the University of Zurich in Switzerland.

Gregoria Convertino (gconvertino@informatica.com) is a senior user researcher at Informatica Corporation in Redwood City, CA.

Copyright held by authors.

Viewpoint

Designing User Incentives for Cybersecurity

How to encourage better user security practices and behavior.

THE TRADITIONAL “PATCHING” approach to managing software vulnerabilities and cybersecurity risk has been less effective than desired. In theory, once a vulnerability is discovered, software patches should be quickly developed and released by producers and then expeditiously applied by users. Successful completion of this process would help to maintain secure systems. However, what has been consistently observed in practice is that this process instead breaks down.⁷ Of particular concern is the failure of the current approach to adequately address the economic incentives that underlie users’ decisions to patch their systems. We propose a simple adaptation to software producer offerings (“versions”) involving users’ patching rights and argue why this change would make a patching approach more effective.

In particular, we advocate that users should be charged for the *right* to control patching of their own individual copies of software. That is, a user’s ability to choose whether to install patches should no longer be an implicit right. Rather, a user who prefers to retain the ability to choose whether and when to patch would need to pay a certain price; one can think of this fee as the price to cause additional security risk (perhaps temporarily) which has negative consequences on other users and the software producer. A user who chooses



to forgo this right, and not pay the premium, has his or her system automatically updated with security patches as soon as they are released. Framed differently, a user can choose whether to purchase a “discounted,” default version that is automatically updated or a “premium” version that includes the right to patch at one’s convenience or even not at all.

Although on the surface this prescription is simple, the design of an incentives-based approach to improve cybersecurity is a difficult task because the level of risk that realizes on a given system or network is a complex outcome of the behaviors of many stakeholders: government, critical infrastructure providers, technology producers, malicious (“black hat”) hack-

ers, and users. With regard to users, how has the current approach failed? For more than 40 years, malware has posed a significant challenge to cybersecurity. The commercialization of the Internet in the 1990s exacerbated this situation tremendously as millions of users began connecting vulnerable devices to networks. Today, the situation is even more problematic. It is common to find an end user with a desktop computer, laptop, tablet, and smartphone running both system software and applications that are not secure. The problem is not limited to companies like Microsoft, who has released on the order of hundreds of security updates to its Windows 7 operating system. There are already one billion smartphones in the U.S. and this number is expected to double by 2015.¹² Many of these devices are running the Android operating system that is being modified by the carriers. These devices are so insecure the American Civil Liberties Union (ACLU) has filed a complaint with the Federal Trade Commission (FTC) requesting action be taken to require vendors to provide more frequent and timely updates to their systems.¹¹

But a patching approach to mitigate cybersecurity risk will only work if users actually apply the patches. Historically, many users have lacked sufficient incentives to properly maintain computing systems by applying these patches. As an example, Code Red is notorious malware that attacked instances of Microsoft's Internet Information Services (IIS) Web server software by exploiting a particular vulnerability. Microsoft had developed a security patch for this vulnerability and released it approximately three weeks before Code Red erupted. Yet, even with the technical fix being made publicly available, most users failed to patch their installations in time, and nearly 360,000 servers were struck by the worm.⁷ Code Red is not an exception. Similar patching windows existed in the cases of other high-profile attacks by SQL Slammer, Blaster, and Sasser. To effectively address the security problem, one must understand how to incentivize improved user behavior.

Users are heterogeneous in the total value they derive from deploying and protecting a system running a given

A patching approach to mitigate cybersecurity risk will only work if users actually apply the patches.

software product. For example, one can consider the differences between two users of enterprise application software: an organization that highly values a system enabled by the software and an organization that values the software just enough to purchase it. A prudent approach to patching involves costly activities including testing, staging, and a controlled roll-out of patches to production systems. Because of the higher value the former organization derives from the software (for example, perhaps the software supports an e-commerce website), it also possesses stronger economic incentives to incur these patching costs and protect its systems. In contrast, the organization that derives less value has much weaker incentives. Not only would such an organization likely not allocate resources to follow the extensive patching process described earlier, even potential inconvenience costs associated with a patch failure may sway it to not deploy patches in a cursory manner either. Although the example with two types of users described here is informative, in reality, there is a continuum of users varying in their patching preferences. But, the point remains: some users tend to have less incentive to patch while other users who utilize software for mission-critical purposes have stronger incentives to patch and protect their systems at the expense of time and resources.

Unfortunately, cybersecurity risk is characterized by detrimental network externalities. That is, when users behave insecurely on a network, they increase the risk faced by everyone else connected to the network.^{1,2} Nevertheless, users in the segment who choose not to patch can still be targeted by eco-

nomic incentives that align their protection decisions more closely with the objective of keeping systems secure. The approach we propose is for software producers to version their products based on users' patching rights. One can logically reason through how this approach would affect different types of users. Organizations and other high-value users will opt for premium versions and pay for the right to patch according to their own timeframes, thereby causing additional security risk on others while remaining unpatched. However, because of their inherent incentives for patching, most of these users will indeed patch at the end of their internal, systematic processes. For users who value the software enough to use it but not enough to protect their systems, they will now be forced to decide whether the right to remain unpatched in a network environment is worth paying the price. Because they may likely belong to the more price-sensitive segment, many of them will prefer not to pay the premium and instead have their systems automatically updated.

Considering that people make daily decisions on whether to install security updates for operating systems, Web browsers, productivity tools, antivirus software, and a long list of other application software, the impact of well-crafted incentives that target user behavior can be quite substantial. Will some consumers prefer discounted versions of popular products like Microsoft Windows and Microsoft Office that automatically update, no longer include the right to be unpatched, and remove the hassle of deploying patches? The answer is likely "yes," and if such a versioning strategy can convert a sizable percentage of unprotected systems from the status quo, then cybersecurity can improve considerably.

Given all of the stakeholders involved, our recommended approach to version based on patching rights also faces counterarguments. Users may contend they should be endowed with the right to choose whether to patch. By not having this right, they would be forced to incur inconvenience costs associated with loss of control, system rebooting, and even system instability due to poorly designed patches. These concerns are valid and underscored by

the not-so-distant memories of Windows XP Service Pack 2. However, we argue the social concern should tilt in favor of improved security. In fact, similar trade-offs exist in other settings where it is common for “users” of products and services to be required to protect others’ interests. For example, most states require that children be immunized before being permitted to attend school. Similarly, many states have requirements that vehicle owners regularly have their vehicles inspected and possibly corrected to meet emission standards. In our case, we are advocating an even milder approach. Users may always retain the ability to choose whether to patch—mandates are not necessary. What is important is that users internalize the cost of causing greater security risk as is reflected through a higher price for retaining patching rights.

Software producers may push back on our recommended approach because they will either lose customers at the low end due to these inconvenience costs or be forced to cut prices to keep them as paying users. While this may be true, producers may also

find versioning on patching rights to be beneficial for several reasons. First, organizations and other high end users will derive greater value from more secure software. If the unpatched population shrinks considerably, these organizations will bear less security risk during the time it takes them to undergo their patching processes. For this lower risk, software vendors can charge a higher price that helps make up for revenue losses associated with the user segment that forgoes patching rights. Second, the often employed argument by software producers that they should not be held liable for security vulnerabilities because they make patches available is somewhat fragile if patches are not being applied.⁵ By providing better economic incentives such that the patching approach they subscribe to actually leads to more secure outcomes, software producers can strengthen their arguments against governmental intervention through means such as liability. There is an opportunity for future research to formally examine these trade-offs using economic models of the decision problem faced by software producers;

such research can yield useful and important insights.

One noteworthy risk of our approach concerns the reaction of black hat hackers. Like other economic agents, black hats have typically found it more cost effective to reverse engineer security patches and develop attacks to exploit the vulnerabilities these patches aim to fix. In this sense, black hats are leveraging the fact that users’ lack of incentives to deploy patches leads to large exploitable populations. Our recommended approach would likely force black hats to redirect their efforts toward other endeavors such as finding unknown vulnerabilities and exploiting them with zero-day attacks. What is important to note is that by their revealed attack preferences, these endeavors appear to be costlier for the same economic return on effort. Hence, the extent of their efforts may be partially reduced. Nevertheless, zero-day attacks can cause considerable economic losses. In anticipation of how black hats respond, both organizations and end users will necessarily need to adjust their defense-in-depth strategies such



Your next job is waiting.

Many of today’s senior tech jobs go unfilled as candidates lack the advanced skills and training needed. Demonstrate your commitment to get ahead and earn CEUs by completing non-credit courses from **Georgia Tech’s Master of Science in Computer Science** online program.

Each course is just \$399.

For more information or to register, visit
pe.gatech.edu/omscs-acm

Georgia Tech Professional Education

ACM Transactions on Reconfigurable Technology and Systems



This quarterly publication is a peer-reviewed and archival journal that covers reconfigurable technology, systems, and applications on reconfigurable computers. Topics include all levels of reconfigurable system abstractions and all aspects of reconfigurable technology including platforms, programming environments and application successes.

www.acm.org/trets
www.acm.org/subscribe



Association for
Computing Machinery

that the economic gains from a world with versioning on patching rights are not overshadowed by losses in the “equilibrium” state that arises. More research studies that endogenize black hat behavior can help to better predict the actual outcomes.

The U.S. government would likely be a strong advocate of the approach we outline. Several federal agencies including the National Science Foundation (NSF) have recognized that innovative policies are needed to help reduce security risks currently faced by the U.S.^{4,8} Along with President Obama’s executive order to develop a cybersecurity framework, the Department of Commerce was also directed to determine what types of economic incentives will cost-efficiently help facilitate adoption of the framework and whether additional legislation may be required.⁹ The government seems to implicitly favor a *voluntary* approach toward improving cybersecurity. For example, whether software producers should be held liable for the economic losses incurred by their users due to poor security has been heavily debated over the last decade, but with little legislative action taken by the government.^{6,10} In spirit, the idea is that holding a company like Microsoft liable will ultimately hurt its bottom line and thus finally provide incentives for greater investments in making its products more secure. This outcome may indeed be the case. But, other undesirable outcomes can certainly arise instead. In particular, Microsoft may make strategic choices to limit its liability. One way to do so is to serve fewer users because a smaller network of users corresponds to reduced security risk due to the network externality. Specifically, all software users benefit in terms of security when there are fewer users exhibiting insecure behaviors, such as not protecting their individual systems. Under a liability policy, Microsoft would, in turn, benefit by not paying out as much to cover users’ losses. If this latter effect of a liability policy is strong, Microsoft may in fact reduce its investments and/or raise its prices to achieve a smaller user population.³

Instead, an approach where software producers begin versioning their products based on patching rights

seems to strike a balance across the interests of government, software producers, and users. Unlike government-imposed liability, this approach is more consistent with how the government has thus far attempted to nudge stakeholders toward better cybersecurity outcomes. Furthermore, targeting user incentives to protect their machines can be a more direct and effective approach in comparison to liability schemes that software producers would prefer to avoid. In fact, if producers are able to charge higher prices from users who appreciate the increased security and it thereby leads to increased producer profitability, there is the potential for win/win outcomes that also substantially improve the economic value associated with software to society. **C**

References

1. Anderson, R. and Moore, T. The economics of information security. *Science* 314, (2006), 610–613.
2. August, T. and Tunca, T. Network software security and user incentives. *Management Science* 52 (2006), 1703–1720.
3. August, T. and Tunca, T. Who should be responsible for software security? A comparative analysis of liability policies in network environments. *Management Science* 57 (2011), 934–959.
4. Department of Defense, *Department of Defense Strategy for Operating in Cyberspace* (2011); <http://www.defense.gov/news/d20110714cyber.pdf>.
5. Espiner, T. EC wants software makers held liable for code. *ZDNet* (2009).
6. Heckman, C. Two views on security software liability: Using the right legal tools. *IEEE Security & Privacy* 1, (2003), 73–75.
7. Moore, D., Shannon, C., and Brown, J. Code-Red: A case study on the spread and victims of an Internet worm. In *Proceedings of the ACM SIGCOMM/USENIX Internet Measurement Workshop* (2002), 273–284.
8. National Science Foundation. *Secure and Trustworthy Cyberspace (SaTC) Program Solicitation NSF 12-596* (2012); www.nsf.gov/pubs/2012/nsf12596/nsf12596.pdf.
9. Obama, B. Executive Order—Improving Critical Infrastructure Cybersecurity. The White House, Office of the Press Secretary, Washington, D.C., 2013.
10. Ryan, D. Two views on security software liability: Let the legal system decide. *IEEE Security & Privacy* 1, (2003), 70–72.
11. Satter, R. ACLU: Slow smartphone updates are privacy threat. Associated Press (2013).
12. Yang, J. Smartphones in use surpass 1 billion, will double by 2015. *Bloomberg* (2012).

Terrence August (taugust@ucsd.edu) is an associate professor in the Rady School of Management at the University of California, San Diego and a visiting Iljin Professor at the Korea University Business School, Seoul, Korea.

Robert August (augur@lake.ollusa.edu) is an associate professor emeritus in the School of Business and Leadership at Our Lady of the Lake University, San Antonio, TX.

Hyoduk Shin (hdshin@ucsd.edu) is an assistant professor in the Rady School of Management at the University of California, San Diego.

This Viewpoint is based upon work supported by the National Science Foundation under Grant No. CNS-0954234.

Copyright held by authors.

Are you looking for your next IT job?

Do you need Career Advice?

The **ACM Career & Job Center** offers ACM members a host of career-enhancing benefits:

- A **highly targeted focus** on job opportunities in the computing industry
- **Access to hundreds** of industry job postings
- Resume posting **keeping you connected** to the employment market while letting you maintain full control over your confidential information
- **Job Alert system** that notifies you of new opportunities matching your criteria
- **Career coaching** and guidance available from trained experts dedicated to your success
- **Free access** to a content library of the best career articles compiled from hundreds of sources, and much more!



Visit **ACM's Career & Job Center** at:
<http://jobs.acm.org>



Association for
Computing Machinery

Advancing Computing as a Science & Profession

The **ACM Career & Job Center** is the perfect place to begin searching for your next employment opportunity!

Visit today at **<http://jobs.acm.org>**

Article development led by [acmqueue](http://acmqueue.queue.acm.org)
queue.acm.org

Better education needed to develop the discipline.

BY ELLEN CHISA

Evolution of the Product Manager

SOFTWARE PRACTITIONERS KNOW product management is a key piece of software development. Product managers talk to users to help figure out what to build, define requirements, and write functional specifications. They work closely with engineers throughout the process of building software. They serve as a sounding board for ideas, help balance the schedule when technical challenges occur—and push back to executive teams when technical revisions are needed. Product managers are involved from before the first code is written, until after it goes out the door.

Due to this close relationship, it is often aligned with the engineering teams. At Microsoft, for example, it is considered to be part of the core software-engineering organization (note, Microsoft uses the term “program manager” to refer to the

responsibilities called “product management” in this document).¹⁰ At Google, a computer science/computer engineering (CS/CE) degree is required to apply for an associate product management (APM) role.⁶

While a technical degree is an asset to a product manager, it is not enough to excel in the role. The industry needs to push for better education tailored to aspiring product managers. They currently have a few ways to learn and develop their product management skills—on the job, in education boot camps, or in graduate programs—but more needs to be done.

One reason product management has not traditionally been included in engineering curricula is because it did not start as an engineering role. Its earliest form was *brand management*, a term coined by a young advertising manager named Neil McElroy, who in 1931 wrote a memo to the executive team at Procter & Gamble proposing the idea of a “Brand Man”—an employee who would be responsible for a product, rather than a business function.³ The role had many similarities to modern-day product management. His memo called out the need to promote processes that work and outline solutions to problems. Above all, it called for the “Brand Man” to take full responsibility for the product. From McElroy’s memo:

- ▶ “Where brand development is ... progressing, examine carefully the combination of effort that seems to be clicking and try to apply his same treatment to other territories that are comparable.”

- ▶ “Where brand development is light ... study the territory personally at first hand ... find out the trouble ... develop a plan ... outline this plan ... prepare all other necessary material for carrying out the plan ... keep whatever records are necessary.”

- ▶ “Take full responsibility, not simply for criticizing individual pieces of the printed word copy, but also for the general printed word plans for his brands.”

Between the 1930s and the 1980s

PRODUCT MANAGER

#partylikeaproductmanager



What my friends think I do



What my mom thinks I do



What society thinks I do



What my colleagues think I do



What I think I do



What I actually do

many firms adopted brand-management principles. Consumer-product companies still tend to group these responsibilities under the title of product manager. Some graduate business programs focus on what they call “brand and product management.”

As the software industry matured, its products became more consumer-targeted, and companies began to see the wisdom in applying some of the same principles of consumer product management to their products. Intuit was among the first tech companies to do so. Its founder, Scott Cook, was previously a brand manager at P&G and implemented those principles at Intuit. The company focused on user needs to develop its software.

Other software companies came to the same discipline from a different direction. Microsoft started out as a software company without the product-management discipline (called *program management* in the Microsoft vocabulary). As time went on, engineers alone

could not meet technical demands and create a cohesive product for users.

Former Microsoft executive Steven Sinofsky describes the program-management role at the company as filling in gaps that software engineers could not:¹²

“Program managers got started at Microsoft while developing Excel for the Macintosh. The challenge the company saw was that we were stretched thin trying to push the state of the art in technology ... The developers were very busy just trying to make things like printing, display, graphics, etc. work as well as trying to develop new algorithms for spreadsheets and modeling. That meant that the place where we were not focusing enough attention was on the usability or the scenarios for how people would use the software ... So at Microsoft a new role was created in program management with the explicit goal of partnering with development and working through the entire product cycle as the advocate for end users and customers.”

Similarly, Joel Spolsky, who once held the title of program manager on the Microsoft Excel team, recalls that Jabe Blumenthal, a programmer on the Mac Excel team in the late 1980s, “noticed that software development was getting so complicated that none of the programmers had the time to figure out how to make software that was either usable or useful. The marketing team was ranting and raving about customer needs and nobody had time to talk to them or translate their MBA-speak into actual features. There was a lot of product design stuff that took a lot of work: talking to users, running usability tests, reviewing competitive products, and thinking hard about how to make things easier, and most programmers just didn’t have the time (nor were they particularly good at it). Blumenthal took the title ‘program manager ...’” a role that involved designing user interfaces, writing functional specs, coordinating teams, and serving as

the customer advocate.¹³

Regardless of whether the role comes out of the realm of brand management or engineering, product managers end up with similar roles. They are responsible for figuring out a plan, and making sure that plan is implemented.

Current State of Product Management

Product management as a job has evolved organically. Universities still lack accredited product-management programs (some graduate programs exist for brand and product management, though they typically focus on the consumer-goods industry, not the computing industry). Since there is no shared training for product managers, many companies have evolved their own doctrines.

One favored resource is Ben Horowitz's document, "Good Product Manager/Bad Product Manager,"⁸ created for his work at Netscape. It echoes McElroy's memo when it states, "Good product managers take full responsibility and measure themselves in terms of the success of the product."

Similarly, Paula Gray, anthropologist-in-residence at the Association of International Product Marketing and Management, states that product managers "must be able to envisage the product from start to finish and to have the ability to ensure that this vision and strategy of the organization is realized."⁷

The role differs, however, in the way it sits within the organization. Horowitz emphasizes the product manager as part of the engineering team, as is common in software companies. "Engineering teams don't consider good product managers a 'marketing resource.' Good product managers are the marketing counterparts of the engineering manager."

That is the key distinction between the origin of product management and its role in the software industry. In software it is not just about executing the advertising strategy. Because the products change so quickly, it is also about deciding what to build, and executing that.

Other respected pieces on how the discipline has been implemented include: Microsoft's JobsBlog article on the Zen of PM;⁹ First Round Capital's

articles on Google's product strategy⁴ and an explanation of good product management at Facebook;⁵ and Julie Zhuo's essay on how to work with product managers.¹⁵

While it varies by company, the role of product manager generally encompasses three areas:

- *Experience* (design). This is the user-facing aspect of the product. It means deciding which features to build for the users—not necessarily which features will make money but which ones make for a better product.

- *Technology* (engineering, project management). This involves understanding the implementation of the product. At the least, it means managing the schedule and checking in on accomplishments. In a more technical product-manager role it might involve working directly with developers to create an API specification.

- *Strategy* (business). This is the piece most aligned with brand management. Strategy means deciding which business areas the product needs to grow in and why. It also means running A/B tests and other experiments to help optimize the performance and revenue of the product.

Types of product managers. Since the role varies by company and has three distinct parts, not every product manager job is the same. Most product managers are stronger in one or two areas. Their areas of strength are often influenced by their fields of study and experience before entering the career.

Christina Wodtke has proposed a framework in which product managers are identified by their areas of strength.¹⁴ She suggests that because product managers enter from another discipline, they take on the flavor of that discipline—be it design, engineering, or business.

A highly technical company may want product managers with more of an engineering focus. A company that sells to enterprise clients might prefer product managers who have more of a business focus. A company striving to find market fit might want more experience-focused product managers.

Current Education for Product Management

Since there is no standard way to be-

come a product manager, people take a variety of approaches. A few of the most common approaches are:

- Going to engineering school and then entering straight into a product-management role and learning on the job.¹

- Going to engineering or design school, starting in that discipline, and then transitioning into product management as needed on the team (much the way product management was originally created during work on Mac for Excel).

- Working in an adjacent nontechnical role such as marketing, account management, or customer service. Nontechnical employees sometimes transition into product management as a way of getting closer to the product.²

- Going to an MBA program.¹¹ In fact, as product management becomes a more desired career for MBAs, MBA programs are creating new classes. One of the best is Tom Eisenmann's PM101 class at Harvard Business School, which focuses on relevant content for both consumer goods and software product management (<http://bit.ly/1mlNZn4>).

Recently, programs such as Startup Institute and General Assembly are offering a more targeted education for people trying to move into product management. Startup Institute is a full-time eight-week program to retrain people to be good early-stage startup employees. One of its four tracks focuses on product and design (<http://www.startupinstitute.com/programs/curriculum/product-and-design>), teaching a combination of classic design skills (typography), modern design skills (responsive design), and front-end engineering. It touches on product management.

This is a valid way of teaching product management since it involves many disciplines and helps an individual understand all of the pieces. Yet, attempting to learn to design, write software, and understand product management in eight weeks is a daunting endeavor.

General Assembly (<https://general-assembly.ly/education/product-management>) is another emerging program for product-management education. Its business fundamentals curriculum includes a 10-week, four-hour-per-week product-management

course. Rather than focusing on engineering, design, and product management, it focuses more on the skills a product manager will frequently need, such as creating and prioritizing a features list, creating wireframes, and making a roadmap. The curriculum also includes some classic business education that is not commonly used for software product managers, such as Porter's Five Forces. General Assembly is also starting a new "immersive" product-management class that will allow students to focus on the topic full-time for 10 weeks.

Continuing Education for Product Managers

Once people enter product management via one of the routes stated here, the only way left to learn is on the job. There are a few ways computing professionals can help product managers continue to learn and improve in the field.

Mentorship from experienced engineers. As Horowitz stated in PM101, the product manager should be a foil to the engineering lead. An experienced engineering lead has likely worked with great product managers before and has expectations of the product manager's abilities.

The engineering lead is one of the core people a software product manager works with. They often have overlapping responsibilities. In many companies, both help to keep track of schedules, breakdowns, and progress.

Because of the significant overlap, engineering leads can proactively help new product managers grow into their roles. First, engineering leads can set clear expectations of what they need from product managers. This will help new product managers prioritize their own work and create a more solid relationship with the engineering team.

If the bandwidth is available, engineering leads and product managers can absorb their shared responsibilities. Particularly, if a new product manager is overwhelmed with just trying to decide "what" to build, it may make sense for the engineering lead to take on the project-management responsibility. Then, as the product manager becomes more comfortable, the engineering lead can transfer



Observing a design critique is a good way to see how designers view their work. The key is that at first the product manager should be only observing—and paying careful attention to the designer's feedback.



more work to the product manager over time.

This type of relationship does not need to be restricted to the engineering lead. Often, senior engineers have built many products and worked with many product managers and should not be afraid to give feedback to new product managers.

Engineers can also be welcoming to new, especially nontechnical, product managers by offering to help with technical education. Explaining why a particular feature will take longer to build will help the product manager develop a better intuition for what is or is not feasible.

Other cross training. A benefit of the Startup Institute program is it includes engineering, design, and product management. Product managers can grow in their roles by learning more about the areas in which they are not comfortable.

For nontechnical product managers, usually the most daunting piece is learning more about technical systems. The suggestions already stated for working with engineering teams can help.

Similarly, all product managers should learn more about design and business angles of building a product. They can do so by working with their design and business counterparts. Observing a design critique is a good way to see how designers view their work. The key is that at first the product manager should be only observing—and paying careful attention to the designer's feedback.

Product managers might also want to sit in with the strategy or business development teams to get a better feel for the company strategy, and as a result the product strategy. They might even go on a visit to a customer with an account manager.

Sometimes internal cross training is not an option. Another way to get experience from other disciplines is to use an educational stipend. If a stipend is available, it is wise for a product manager to take a design or business class. Similarly, a technical product manager could opt to attend a design or business conference to learn more about those perspectives.

If a budget for education is out of the question, free online options are

available. These could be accessed during work time without distracting other employees. For nontechnical product managers wanting to learn about software, this could be as simple as trying out an independent training program such as Codecademy (codecademy.com), Treehouse (team-treehouse.com), or Dash (dash.generalassemb.ly). For technical product managers looking to learn about design, HackDesign (hackdesign.org) is a good option.

Conferences and meetups. Another way product managers can improve their skills is to learn from other product managers. Conferences and meetups are good places to find experienced product managers. These are the people who would qualify as “faculty” if product management were an academic discipline.

At conferences, product-management leaders share their experience on an array of topics. This can be a good way to get a product manager thinking about different challenges in building social features, thinking about internationalization, and launching new product lines, for example.

Unfortunately, there are not that many product-management conferences. Wodtke postulates it is because product management is more focused on success of the product than on how to make products. “PM is a different job because how they do it is a distant second to their success than results. Perhaps that’s why there is so little written about how to be a PM compared to the other disciplines, and so few conferences.”

Available conferences include #ProductSF (<https://www.eventbrite.com/e/productsf-building-better-products-tickets-10613205383>), Product Camp (productcamp.org), and the Product Management Festival (productmanagementfestival.com).

While meetups are informal, they also are opportunities for product managers to get together and discuss the challenges they face. Many software startups have few product managers, and not many mutual experiences to draw on. A broader network is a valuable resource, especially for new product managers.

Peer mentoring. Another trend in product-management training is peer

mentoring. This is similar to the role of conferences and meetups, but with more of a one-on-one experience with another product manager.

One way to find a peer mentor is through an existing network. Product managers at Union Square Ventures-funded companies can apply to its peer-mentoring program. Other firms may start similar programs.

If you are a manager with connections across the industry, consider talking to other managers to help connect your product managers to those in other companies. (You could also do the same for junior engineers!)

Early exposure. Another way to get better product managers is simply by training more people in school and ensuring technical students have the opportunity to pursue product management. Students should be made aware of the opportunity and have the chance to practice before entering the workforce.

This does not mean engineering students should be doing less technical work. Product-management skills can be introduced as part the work students are already doing, and a substantial benefit can be made with small adjustments.

Technical professionals who have entered the product-management discipline should spread the word. They can return to their schools and give talks on their work as product managers. Companies can be sure to mention product management (or sponsor talks on the discipline) during industry events.

Within academic programs, requiring more projects will allow for more product-management experience. Two programs that already do this are the Olin College of Engineering and the Rochester Institute of Technology.

Conclusion

Product management started out as brand management, and in the software industry was redeveloped to fill a gap in engineering. It then proliferated across companies. Unfortunately, product-management education has not caught up to its prevalence in the industry. The field does not have consistent product-management education. This makes it difficult to have a consistent set of expectations. A more

standard curriculum could add significant value to the discipline. Computing professionals need to work proactively to improve the quality of education and continuing education for product managers. 

Related articles on queue.acm.org

Getting What You Measure

Eric Bouwers, Joost Visser,
and Arie Van Deursen
<http://queue.acm.org/detail.cfm?id=2229115>

Best Practice (BPM)

Derek Miers
<http://queue.acm.org/detail.cfm?id=1122688>

Outsourcing: Devising a Game Plan

Adam Kolawa
<http://queue.acm.org/detail.cfm?id=1036501>

References

- Chisa, E. How I became a PM. Ellen's Blog, 2014; <http://blog.ellenchisa.com/2014/01/10/how-i-became-a-pm/>.
- Daily Muse. How 5 product managers got their start, 2013; <https://www.themuse.com/advice/how-5-product-managers-got-their-start>.
- Duffy, S. The McElroy “Brand Man” memo turns 80. Brand Rants, 2011; <http://www.brandrants.com/brandrants/2011/05/mcelroy-brand-man-memo.html>.
- First Round Capital. 42 rules to lead by from the man who defined Google's product strategy. First Round Review; <http://firstround.com/article/42-rules-to-lead-by-from-the-man-who-defined-googles-product-strategy>.
- First Round Capital. Top hacks from a PM behind two of tech's hottest products. First Round Review; <http://firstround.com/article/Top-Hacks-from-a-PM-Behind-Two-of-Techs-Hottest-Products>.
- Google; <https://www.google.com/about/careers/search#t=jo&jid=26535001&>.
- Gray, P. Business anthropology and the culture of product managers. Association of International Product Marketing and Management; <http://www.aipmm.com/html/newsletter/archives/BusinessAnthroAndProductManagers.pdf>.
- Horowitz, B. Good product manager/bad product manager; <http://benhorowitz.files.wordpress.com/2010/05/good-product-manager.pdf>.
- Microsoft JobsBlog. Zen of PM, 2014; <http://microsoftjobsblog.com/zen-of-pm/>.
- Microsoft. University careers; http://careers.microsoft.com/careers/en/us/grad-software-jobs.aspx#tab_pm.
- Sanchez, L. Genevieve Sheehan on creating her own path, 2014; <https://medium.com/dropbox-makers/genevieve-sheehan-on-creating-her-own-path-401192c44f6d>.
- Sinofsky, S. PM at Microsoft, 2005; <http://blogs.msdn.com/b/techtalk/archive/2005/12/16/504872.aspx>.
- Spolsky, J. How to be a program manager, 2009; <http://www.joelonsoftware.com/items/2009/03/09.html>.
- Wodtke, C. 27 thoughts on product management. ElegantHack, 2013; <http://www.eleganthack.com/27-thoughts-on-product-management/>.
- Zhuo, J. How to work with PMs. The Year of the Looking Glass, 2013; <https://medium.com/the-year-of-the-looking-glass/how-to-work-with-pms-3e852d5eccf5>.

Ellen Chisa is an MBA candidate at the Harvard Business School, Cambridge, MA. Previously, she was a product manager at Kickstarter, focusing on backer experiences and was part of the Microsoft team working in program management for Office Mobile.



Conditional dependency resolution.

BY ALEX LIU

JavaScript and the Netflix User Interface

IN THE TWO decades since its introduction, JavaScript has become the de facto official language of the Web. JavaScript trumps every other language when it comes to the number of runtime environments in the wild. Nearly every consumer hardware device on the market today supports the language in some way. While this

is done most commonly through the integration of a Web browser application, many devices now also support Web views natively as part of the operating system user interface (UI). Across most platforms (phones, tablets, TVs, game consoles), the Netflix UI, for example, is written almost entirely in JavaScript.

Despite its humble beginnings as a language intended to be Java's "silly little brother,"⁴ JavaScript eventually became a key component in enabling the Web 2.0 evolution. Via the introduction of Ajax, this evolution added an element of dynamism to the Web, creating the concept of a living and social Web that is now taken for granted. Today the language's influence continues to grow as it makes its way into the server landscape via Node.js. For all of its shortcomings, arguably JavaScript has more than successfully achieved the "write once, run anywhere" motto

that Sun Microsystems often used to tout as one of the benefits of Java.

With increasingly more application logic being shifted to the browser, developers have begun to push the boundaries of what JavaScript was originally intended for. Entire desktop applications are now being rebuilt entirely in JavaScript—the Google Docs office suite is one example. Such large applications require creative solutions to manage the complexity of loading the required JavaScript files and their dependencies. The problem can be compounded when introducing multivariate A/B testing, a concept that is at the core of the Netflix DNA. Multivariate testing introduces a number of problems that JavaScript cannot handle using native constructs, one of which is the focus of this article: managing conditional dependencies. Despite this, engineering ingenuity has enabled Netflix to build highly com-

plex and engaging UIs in a rapid and maintainable fashion.

A/B Testing Netflix.Com

Netflix is steeped in a culture of A/B testing. All elements of the service, from movie personalization algorithms to video encoding, all the way down to the UI, are potential targets for an A/B test. It is not unusual to find the typical Netflix subscriber allocated into 30 to 50 different A/B tests simultaneously. Running the tests at this scale provides the flexibility to try radically new approaches and multiple evolutionary approaches at the same time. Nowhere is this more apparent than in the UI.

While many of the A/B tests are launched in synchrony across multiple platforms and devices, they can also target specific devices (phones or tablets). The tests allow experimentation with radically different UI experiences

from subscriber to subscriber, and the active lifetime of these tests can range from one day to six months or more. The goal is to understand how the fundamental differences in the core philosophy behind each of these designs can enable Netflix to deliver a better user experience.

A/B testing on the Netflix website tends to add new features or alter existing features to enhance the control experience. Many of the website tests are designed from the outset to be cross-allocation friendly—in other words, stackable with other A/B tests. This ensures newly introduced functionality can coexist with other tests. Thus, while one Netflix subscriber's homepage can look similar on the surface to another subscriber's homepage, various bits and pieces of functionality are added or modified throughout the page to make the end product feel

different. It should be noted the testing encompasses all pieces of the Netflix UI (HTML, CSS, JavaScript), but the focus here is on using JavaScript to reduce the scope of the problem.

Facets to Features to Modules

The HTTP Archive estimates the average website in 2014 includes approximately 290KB of JavaScript across 18 different files.² By comparison, the Netflix homepage today delivers, on average, a 150KB payload in a single JavaScript file. This file actually consists of 30 to 50 different files concatenated together, with their inclusion in the payload being dictated by one or many of the hundreds of personalization facets generated by Netflix's recommendation algorithms. These facets can be commonly derived via a subscriber's A/B test allocations, country of signup, viewing tastes, and sharing preferences (Facebook integration), but can be conceivably backed by any piece of arbitrary logic. These facets act as switches, a method by which the UI can be efficiently pivoted and tweaked. This has driven the website into a distinctly unique predicament: how to manage packaging and delivery of many different UIs in a maintainable and performant manner.

It is useful first to draw a clear line connecting the personalization facets and their impact on the UI. A simple example can help illustrate this relationship. Let's imagine that today we want to A/B test a search box. For this test, we may have a control cell, which is the traditional experience that sends users to a search-results page. To accommodate for regional differences in user experiences, we also have a slight variation of that control cell depending on whether the subscriber is located within the U.S. The first test cell provides autocomplete capability, and is available to all subscribers allocated in cell 1. Allocation in this scenario means the subscriber was randomly selected to participate in this test. A secondary test cell provides search results right on the current page by displaying results as the user types. Let's call this *instant search*, and it is available to all subscribers allocated in cell 2. These are three distinct experiences, or "features," with each one being gated by a set of very specific personalization fac-

Table 1. The search-box test.

Experience	Desired Behavior	Requirements
Search Box Test Cell 0 (Control)	Search Results Page	Must be a U.S. subscriber
Search Box Test Cell 0 (Control)	Search Results Page Type 2	Not a U.S. subscriber
Search Box Test Cell 1	Autocomplete	Subscriber allocated to Cell 1
Search Box Test Cell 2	Instant Search	Subscriber allocated to Cell 2

Figure 1. AMD module popularized by libraries such as RequireJS.

```
define('depA', function() {
  // execute some code
});
define('depB', function() {
  // execute some code
});
define('foo', ['depA', 'depB', 'depC'], function(a, b, c) {
  // this function is an async callback.
  // code here only runs after A, B, C have loaded.
  // A and B are loaded as part of page payload,
  // but C will be fetched asynchronously since it has not yet
  // been defined by the time we reach module foo.
  // thus, code here is NOT run until C has been retrieved
  // and executed.
  // the value of this return statement is what this module exports.
  // it is the value imported by other modules when this module is
  // requested as a dependency
  return ...;
});

/*
and a visual representation of the dependency tree:

foo
|__depA
|__depB
|__depC (fetched asynchronously)

*/
```

ets. Thus, users are presented only one of these search experiences when they are allocated to the test and when their facets fulfill the test's requirements (see Table 1). Other parts of the page, such as the header or footer, can be tested in a similar manner without affecting the search-box test.

Under this test stratagem, it is imperative to separate each functional section of the website into discrete sandboxed files known as *modules*. Modules have become a common best practice in JavaScript as a way to sandbox and group relevant features safely in a discrete and cohesive unit. This is desirable for various technical reasons: it reduces the reliance on implied globals; it enables the use of private/public methods and properties; and it allows for the existence of a true imports/exports system. Imports/exports also opens the door for proper dependency management.

In this case, there is yet another driving force behind modules. They allow seamless feature portability from one page to the next. Division of a Web page into smaller and smaller pieces should be done until it is possible to compose new payloads using existing modules. If functionality must be broken out from a previous module to achieve that, it is a likely indicator the module in question had too many responsibilities. The smaller the units, the easier they are to maintain, test, and deploy.

Finally, using modules to encapsulate features provides the ability to build an abstraction layer on top of the personalization facets that gate the A/B tests. Since eligibility for a test can be mapped to a specific feature, and a feature can then be mapped to a module, the JavaScript payload can be effectively resolved for a given subscriber by simply determining that subscriber's eligibility for each of the currently active tests.

Dependency Management

Modules also allow more advanced techniques to come into play, one of which is critically important for complex applications: dependency management. In many languages, dependencies can be imported synchronously, as the runtime environment is colocated on the same machine as the requested dependencies. The complexity in managing browser-side JavaScript dependencies, however, is



Netflix is steeped in a culture of A/B testing. All elements of the service, from movie personalization algorithms to video encoding, all the way down to the UI, are potential targets for an A/B test.



the runtime environment (browser) is separated from its source (server) by an indeterminate amount of latency. Network latency is arguably the most significant bottleneck in Web application performance today,¹ so the challenge is in finding the balance between bandwidth and latency for a given set of indeterministic constraints that may differ per subscriber, per request.

Through the years, the Web community devised several methods to handle this complexity, with varying degrees of success. Early solutions simply included all dependencies on the page, regardless of whether or not the module would be used. While simple and consistent, this penalized users across the board, with bandwidth constraints often exacerbating already long load times. Later solutions relied on the browser making multiple asynchronous requests back to the server as it determined missing dependencies. This, too, had its drawbacks, as it penalized deep dependency trees. In this implementation, a payload with a dependency tree N nodes deep could potentially take up to $N - 1$ serial requests before all dependencies were loaded.

More recently, the introduction of asynchronous module definition (AMD) libraries such as RequireJS allows users to create modules, then preemptively generate payloads on a per-page basis by statically analyzing the dependency tree. This solution combined the best of both previous solutions by generating specific payloads containing only the things needed by the page and by avoiding unnecessary penalization based on the depth of the dependency tree. More interestingly, users can also opt out entirely from the static-analysis step and fall back on asynchronous retrieval of dependencies, or they can employ a combination of both. In Figure 1, a module called `foo` has three dependencies. Because `depC` is fetched asynchronously, $N - 1$ additional request(s) are made before the page is ready (where $N=2$, and N is the depth of the tree). An application's dependency tree can be built using static-analysis tools.

Conditional Dependencies

The problem with AMD and similar solutions is their assumption of a static-dependency tree. In situations where

Figure 2. A header module that loads all three search features but uses only one of them.

```

define('header', ['controlSearch', 'autocompleteSearch',
  'instantSearch', 'facets'],
  function(control, autocomplete, instant, facets) {
    // facets is a dependency that contains personalization facets
    if (facets.inSearchTest && facets.geo !== 'US') {
      autocompleteSearch.init();
    } else if (facets.inSearchTest && facets.geo === 'US') {
      instantSearch.init();
    } else {
      controlSearch.init();
    }
  }
);

/*
even though the header loads all 3 possible dependencies, it really only
needs one of the three:
header
|__controlSearch
|__autocompleteSearch
|__instantSearch
*/

```

Figure 3. A header module that loads only the modules needed.

```

define('header', ['facets'], function(facets) {
  // facets is a dependency that contains personalization facets
  if (facets.inSearchTest && facets.geo !== 'US') {
    // fetch the autocomplete dependency over the wire
    // search cannot be used until this file and all of
    // autocomplete search dependencies are loaded.
    require(['autocompleteSearch'],
      function(autocompleteSearch) {
        // async callback
        autocompleteSearch.init();
      });
  } else if (facets.inSearchTest && facets.geo === 'US') {
    // fetch the instant search dependency over the wire
    // search cannot be used until this file and all of
    // instant search dependencies are loaded.
    require(['instantSearch'], function(instantSearch) {
      // async callback
      instantSearch.init();
    });
  } else {
    // fetch the control search dependency over the wire
    // search cannot be used until this file and all of
    // control search dependencies are loaded.
    require(['controlSearch'], function(controlSearch) {
      // async callback
      controlSearch.init();
    });
  }
});

```

the runtime environment is colocated with the source code, it is common to import all possible dependencies but exercise only one code path, depending on the context. Unfortunately, the penalty for doing so in the browser is much more severe, especially at scale.

The problem can be better visualized by recalling the previous search-box A/B test, which has three distinct search experiences. If the page header

depends on a search box, how do you load *only* the correct search box experience for that given user? It is possible to add all of them to the payload, then have the parent module add logic that allows it to determine the correct course of action (see Figure 2). This is unscalable, however, as it bleeds knowledge of A/B test features into the consuming parent module. Loading all possible dependencies also increases

the payload size, thereby increasing the time it takes for a page to load.

A second option of fetching dependencies just-in-time is possible but may introduce arbitrary delays in the responsiveness of the UI (see Figure 3). In this option, only the modules that are needed are loaded, at the expense of an additional asynchronous request. If any of the search modules has additional dependencies, there will be yet another request, and so on, before search can be initialized.

Both options are undesirable and have proven to have a significant negative impact on the user experience.³ They also do not take into account the possibility that certain personalization facets are available only on the server and for security reasons cannot be exposed to the JavaScript layer.

Big Numbers Change Everything

The Netflix website repository counts more than 600 unique JavaScript files and more than 500 unique Cascading Style Sheets (CSS) files. A/B tests account for the vast majority of these files. A guesstimate of the amount of different JavaScript payloads the website deals with can be made using the formula for unique combinations:

$$n!/r!(n-r)!$$

Assuming a total bucket of 600 modules, and estimating the average JavaScript payload includes about 40 modules, you arrive at the following number of possible combinations:

$$600!/40!(560!) = 433518929550349486086117218185493567650720611537099964754232870$$

This number is eye-catching, though not entirely honest. Of the 600 different modules, most are not independently selectable. Many of those modules depend on other common platform modules that then depend on third party modules. Furthermore, even the largest of A/B tests usually affects fewer than three million users. This seems like a large population to test on, but in reality it is still a small percentage of the total 50-plus million subscriber base. This information leads to some early conclusions: first, the allocation of the tests is not large enough to spread

evenly over the entirety of the Netflix subscriber base; and second, the number of independently selectable files is extremely low. Both of these will contribute to a significantly reduced number of unique combinations.

Rather than attempt to adjust the formula, it might be more practical to share some empirical data. The website deploys a new build on a weekly cycle. For every build cycle, the website generates approximately 2.5 million unique combinations of JavaScript and CSS payloads.

Given this huge number, it is tempting to go the route of letting the browser fetch dependencies as the tree is resolved. This solution works for small code repositories, as the additional serial requests may be relatively insignificant. As previously mentioned, however, a typical payload on the website contains 30 to 50 different modules because of the scale of A/B testing. Even if the browser's parallel resource fetching could be leveraged for maximum efficiency, the latency accumulated across a *potential* 30-plus requests is significant enough to create a suboptimal experience. In Figure 4, even with a significantly simplified example with a depth of only five nodes, the page will make four asynchronous requests before the page is ready. A real production page may easily have 15-plus depth.

Since asynchronous loading of dependencies has already been disqualified for this particular situation, it becomes clear the scale of A/B testing dictates the choice to deliver a single JavaScript payload. If single payloads are the solution, this might give the impression these 2.5 million unique payloads are generated ahead of time. This would necessitate an analysis of all personalization facets on each deployment cycle in order to build the correct payload for every possible combination of tests. If subscriber and A/B testing growth continues on its correct trajectory, however, then preemptive generation of the payloads becomes untenable. The number of unique payloads may be 2.5 million today, but five million tomorrow. It is simply not the correct long-term solution for Netflix's needs.

What the A/B testing system needs, then, is a method by which conditional dependencies can be resolved

without negatively affecting the user experience. In this situation, a server-side component must intervene to keep the client-side JavaScript from buckling under its own complexity. Since we are able to determine all possible dependencies via static analysis, as well as the conditions that trigger the inclusion of each dependency, the best solution given our requirements is to resolve all conditional dependencies *when the payload is generated* just-in-time.

Just-in-Time Dependency Resolution

Let's add another column to the search-box test definition (see Table 2). This table now represents a complete abstraction of all data needed to build the payload. In practice, the final column mapping exists only in the UI layer, not in the core service that provides the A/B test definition. Often, it is up to

the consumers of the test definitions to build this mapping since it is most likely unique for each device or platform. For the purposes of this article, however, it is easier to visualize the data in a single place.

Assume the payload contains the files for the homepage shown in Fig-

Figure 4. Letting the browser fetch dependencies.

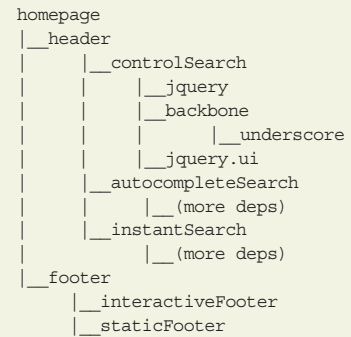


Table 2. The search-box test with additional mapping to the modules.

Experience	Desired Behavior	Requirements	Module name
Search Box Test Cell 0 (Control)	Search Results Page	None	controlSearch
Search Box Test Cell 1	Autocomplete	Not a US subscriber	autocompleteSearch
Search Box Test Cell 2	Instant Search	Must be a US subscriber	instantSearch
Search Box Test Cell 2	Instant Search	Instant Search	Subscriber allocated to Cell 2

Figure 5. The homepage payload.

```

define('search', function() {
    // search has no dependencies.
    // which search module goes here?
});
define('header', ['search'], function(search) {
    search.init();
});
define('footer', function() {
    // footer has no dependencies.
    ...
});
define('homepage', ['header', 'footer'], function(header, footer) {
    ...
});

/*
and a visual representation of the dependency tree:
homepage
|__header
|   |__search
|__footer
*/
  
```

ure 5. The browser has asked for the homepage JavaScript payload. There is a dependency tree created as a result of static analysis, and there is a table that maps the search module to three potential implementations. Since the header cares only about the inclusion of a search module, but not its implementation, we can plug in the correct search module by ensuring all implementations conform to a specific contract (that is, a public API), as in Figure 6.

Having variations of a single experience conform to a similar public API allows us to change the underlying implementation by simply including the correct search module. Unfortunately, because of JavaScript's weakly typed nature, there is no way to enforce this contract, or even to verify the validity of any modules claiming to conform to said contract. The responsibility to do the right thing is often left up to the developers creating and consum-

ing these shared modules. In practice, nonconforming modules are not game breakers; “drop-in” replacements as in the previous example are typically entirely self-contained with the exception of a single entry point, which in this case is the exposed `init()` method. Modules with complex public APIs tend to be shared common libraries, which are less likely to be A/B tested in this manner.

It is also worth noting the number of differences between each of these A/B experiences can often drive whether or not doing a drop-in replacement is even possible. In some cases where the new experiences are designed to be intentionally and maybe even radically different, it can make sense to have differences in the public API. This almost certainly increases complexity in the consuming parent modules, but that is the accepted cost of running radically different experiences concurrently. Other strategies can

help mitigate the complexity, such as returning module stubs (see Figure 7), rather than attempting a true drop-in replacement. In this scenario, the module loader can be configured to return an empty object with a stub flag, indicating it is not a true implementation. This strategy can be useful if the A/B experiences in question share almost nothing in common, and would benefit very little, if at all, from a common public API.

Continuing with the example of the homepage payload, when a request comes in asking for the homepage payload (see Figure 8), we already know all the possible files the subscriber may receive, as a result of static analysis.

As we begin appending files to the payload, we can look up in the search-box test table (Table 2) whether or not this file is backed by an eligibility requirement (that is, whether the subscriber is eligible for that feature). This resolution will return a Boolean value, which is used to determine if the file gets appended (Figure 9).

Using a combination of static analysis to build a dependency tree, which is then consumed at request time to resolve conditional dependencies, we are able to build customized payloads for the millions of unique experiences across Netflix.com. It is important to note this is only the first step in a chain of services that finally delivers the JavaScript to the end user.

For performance reasons, it is never desirable to deliver the entire payload via an inline script. Inline scripts cannot be cached independently from the HTML content, so the benefits of browser-side caching are lost immediately. It is much more desirable to deliver it via a script tag that points to an URL representing this payload, which a browser can easily cache. In most cases, this is a CDN (content delivery network)-hosted URL whose origin server points back to the original server that generated this payload. Thus, everything discussed up to this point is merely responsible for generating the uniqueness of the payload.

It is not sufficient, however, simply to cache the unique payload with a randomly generated identifier. If the server has multiple instances running for load balancing, any one of those instances could receive the incoming

Figure 6. Three search experiences that return different implementations using the same public API.

```
// controlSearch.js
define('search', function() {
  // the control search implementation here
  return { init: function(){...}};
});

// autocompleteSearch.js
define('search', function() {
  // the autocomplete search implementation here
  return { init: function(){...}};
});

// instantSearch.js
define('search', function() {
  // the instantSearch search implementation here
  return { init: function(){...}};
});
```

Figure 7. A strategy of returning module stubs.

```
define('autocompleteSearch', function() {
  // the autocomplete search implementation here
  return { init: function(){...}};
});

// note that controlSearch is not in this payload.
// the module system here will return a stub when it is asked for.
define('header', ['controlSearch', 'autocompleteSearch'],
function(search) {
  if (!controlSearch.stub) {
    controlSearch.init();
  }
  if (!autocompleteSearch.stub) {
    autocompleteSearch.init();
  }
});
```

request for this payload. If the request goes to an instance that has not yet generated (or cached) that unique payload, it cannot resolve the request. To solve this issue, it is critically important the payload's URL is reverse resolvable; any instance of your server must be able to resolve the files in a unique payload by simply looking at the URL. This can be solved in a few ways, most often by representing a file by referencing the file name directly in the URL or by using a combination of unique hashes, where each chunk of the hash can be resolved to a specific file.

Future Optimizations

Though we have optimized for a single payload, there is potential to use parallel browser requests for additional performance gains. We want to avoid unbundling the entire payload, which forces us to take the route of making 30-plus requests, but we could split our single payload into two, with the first containing all common third-party libraries or shared modules, and the second bundle containing page-specific modules. This would allow the browser to cache common modules from page to page, further decreasing the upper limit of time to page ready as the user moves through the site. This strikes a nice balance between the bandwidth and latency constraints that Web browsers must typically deal with.

Looking Forward

For all of its shortcomings, JavaScript has become the de facto language of the Web, and as the industry grows it will continue to be used across countless devices and platforms. The problems identified in this article are just the tip of the iceberg, especially as applications grow in size and complexity. The reality is JavaScript is still primarily a *client*-side language, whose runtime environment is primarily the browser. What this means is most libraries or tools designed to address complex problems such as conditional dependencies have approached and attempted to solve the problem from the browser domain.

Constriction of the approach to solving issues from within the browser limits the possibilities of a richer end-to-end solution. Though the updates

Figure 8. The exploded dependency tree of the example homepage payload.

```
homepage
|__header
|   |__controlSearch
|   |__autocompleteSearch
|   |__instantSearch
|__footer
```

Figure 9. The dependency tree with ineligible files removed at request time.

```
homepage
|__header
|   |__controlSearch
|   |__autocompleteSearch
|   |__instantSearch
|__footer
```

in the upcoming ECMAScript 6 revision have provisions for both native JavaScript modules and a module loader, it, too, suffers from the same problem of constricted scope. Even the most fully complete module systems today address the problem only from within the browser domain.

The wild-card constraint, as we have discovered, is that browser runtime environment is too “far” from the location of the source code (server). Historically, larger Web-development teams have refrained from developing solutions that tightly integrate the server and browser domains. The reason for this is most likely simplicity, or perhaps the desire for a clearer separation of concerns between client-side and server-side code. Conditional dependencies, however, make the existence of this constraint painfully clear. Any solution that fails to account for this will inevitably leave some performance on the table. As a result, the most performant JavaScript packaging solutions for resolving conditional dependencies will require a server-side component, at least for the foreseeable future.

With the rise of Node.js and JavaScript on the server, it is entirely possible the problems we face today will gain more exposure. In many enterprise environments, the server has been a completely separate domain, owned by engineers with a different

set of skills. Node.js, however, opens the door for many front-end engineers to move to the server side, expanding not only the role of the traditional front-end engineer, but also the set of tools available to solve UI-specific problems. This paradigm shift, along with the newly expanded role of the front-end engineer, does give some hope for the future. With front-end engineers owning the UI server, they can control end-to-end for all UI-specific concerns.

Solutions such as the one discussed in this article were born out of this exact environment, and it is a good sign for the future. Despite JavaScript lacking the conventions to natively handle some of the problems of today's Web-application-heavy world, some creative engineering can fill in the gaps in the meantime. Solving complex problems such as conditional dependencies makes it possible for Netflix to continue to build large, complex, and engaging UIs—all in JavaScript. The website's recent adoption of Node.js could not be a better endorsement of the language and Netflix's commitment to the open Web. **C**

Related articles on queue.acm.org

Reveling in Constraints

Bruce Johnson

<http://queue.acm.org/detail.cfm?id=1572457>

Multitier Programming in Hop

Manuel Serrano and Gérard Berry

<http://queue.acm.org/detail.cfm?id=2330089>

The Antifragile Organization

Ariel Tseitlin

<http://queue.acm.org/detail.cfm?id=2499552>

References

1. Grigorik, I. Latency: The new Web performance bottleneck; <https://www.igvita.com/2012/07/19/latency-the-new-web-performance-bottleneck/>.
2. HTTP Archive. Trends; <http://httparchive.org/trends.php?s=All&minlabel=Nov+15+2010&maxlabel=Jun+15+2014>.
3. Nielsen, J. Response times: The three important limits (1993 updated 2014); <http://www.nngroup.com/articles/response-times-3-important-limits/>.
4. Severance, C. JavaScript: Designing a language in 10 days. *Computer* 45, 2 (2012), 7-8; <http://www.computer.org/csdl/mags/co/2012/02/mco2012020007.html>.

Alex Liu is a senior UI engineer at Netflix and part of the core team leading the migration of Netflix.com to Node.js. He has spent his career building JavaScript applications on the desktop, browser, and server, and is still surprised by all the new and creative ways in which the community leverages JavaScript.

Copyright held by owner/author. Publications rights licensed to ACM. \$15.00

Looking at the design and benefits of X10.**BY JOHN T. RICHARDS, JONATHAN BREZIN, CALVIN B. SWART,
AND CHRISTINE A. HALVERSON**

A Decade of Progress in Parallel Programming Productivity

IN 2002, THE U.S. Defense Advanced Research Projects Agency (DARPA) launched a major initiative in high-productivity computing systems (HPCS). The program was motivated by the belief that the utilization of the coming generation of parallel machines was gated by the difficulty of writing, debugging, tuning, and maintaining software at peta scale.

As part of this initiative, DARPA encouraged work on new programming languages, runtimes, and tools. It believed by making the expression of parallel constructs easier, matching the runtime models to the heterogeneous processor architectures under development, and providing powerful integrated

development tools, that programmer productivity might improve. This is a reasonable conjecture, but we sought to go beyond conjecture to actual measurements of productivity gains.

While there is no established method for measuring programmer productivity, it is clear a productivity metric must take the form of a ratio: programming results achieved over the cost of attaining them. In this case, *results* are defined as successfully creating a set of parallel programs that ran correctly on two workstation cores. This is a long way from peta scale, but since new parallel software often starts out this way (and is then scaled and tuned on ever larger numbers of processors), we viewed it as a reasonable approximation. Moreover, results found with two cores should be of interest to those coding nearly any parallel application, no matter how small. *Cost* was simpler to handle once results was defined, since it could reasonably be approximated by the time it took to create this set of parallel programs.

The purpose of this study was to measure programmer productivity, thus defined, over the better part of the decade starting in 2002, the beginning of the HPCS initiative. The comparison was primarily focused on two approaches to parallel programming: the SPMD (single program multiple data) model as exemplified by C/MPI (message-passing interface), and the APGAS (asynchronous partitioned global address space) model supported by new languages such as X10 (<http://x10-lang.org>), although differences in environment and tooling were also studied. Note that the comparison was *not* between C/MPI as it has come to be with X10 as it is now. Rather, it was a historical contrast of the way things were in 2002 with the way things are now. Indeed, C++ with its exceptions and MPI-2 with its one-sided communication protocol likely enhance programmer productivity and are worthy of study in their own right.

Given our objective, we sought to replicate as closely as possible the



programming environment found in 2002 for users of C/MPI. This included the gdb debugger, along with a typical set of command-line tools. For X10, we used Eclipse (<http://www.eclipse.org>) with the X10 plug-in as it was found in 2010, the date of this study. X10 was developed as part of the HPCS initiative. It combines the succinctness of newer languages such as Scala (<http://www.scala-lang.org>) with a model of concurrent programming that maps nicely to the PGAS model of modern parallel machines. After a decade, the language is still evolving,

but the basic model was stable when this study began. Importantly, no X10 debugger was available until after this study was completed, so current users of X10 can expect to see gains even larger than those reported here.

Prior Work

This is not the first study to attempt to measure the productivity of programming in X10. Ebcioglu et al.⁵ and Danis and Halverson³ describe an experiment in which undergraduates (novices with respect to parallel programming) were trained in either

C/MPI, UPC (Unified Parallel C), or (a very early version of) X10 and then attempted to parallelize the string-matching algorithm of SSCA 1 (Scalable Synthetic Compact Application 1), as described in Bader et al.² Average time to completion was approximately 510 minutes in C/MPI, 550 minutes in UPC, and 290 minutes in X10. When the differences in code-execution time during testing were removed, these times were approximately 360 minutes for C/MPI, 390 minutes for UPC, and 180 minutes for X10. Interpretation of this roughly twofold pro-

Days to first successful parallel run on two cores and lines of code in completed programs.

Programs	X10 Days	X10 Lines	C/MPI Days	C/MPI Lines
SSCA 1	9.0	1625	20	1225
SSCA 2	7.9	1150	28	3000
Consumer-Producer	5.0	1275	24	1750
Unbalanced Tree Search	7.5	1125	30	2200
Floyd's Algorithm	3.8	520	12	1400
Discrete Fourier Transform	5.6	500	15	670
Total	38.8	6195	129	10,245

ductivity gain was somewhat complicated, however, by the absence of clear success criteria for code completion.

In a subsequent study, Halverson and Danis⁶ added the Eclipse programming environment to the tools available to X10 programmers. Participants in this study were more experienced, having had some parallel programming training in earlier coursework. A productivity gain for X10 and Eclipse over C/MPI was found here as well, but computing the size of this gain was complicated by the fact that only one of the seven C/MPI participants completed the task (in 407 minutes) vs. five of the seven X10 participants in an average of 321 minutes.

In both of these earlier studies, only one program was coded (a portion of SSCA 1), only relatively novice parallel programmers were tested, and the task was limited to parallelizing an embarrassingly parallel algorithm that was provided to the test participants in working serial code. All three of these deficiencies are addressed in the assessment reported here.

The Experimental Approach

In a long-running study we measured the time to develop six parallel programs in each of two languages for a total of 12 programs in all. Each program was developed from a problem statement and/or algorithm description without benefit of any preexisting code samples. The time to completion ended with the first successful parallel run using both cores of a two-core workstation.

Two of the six programs were defined in phase 2 of the DARPA HPCS program, SSCA 1 and SSCA 2, and are described in Bader et al.² SSCA 1 is the string-matching problem used in the two previously referenced studies. In

this study, however, two versions of SSCA 1 were developed in each language: the embarrassingly parallel version that works well when the target string is much shorter than the search string; and a more complex anti-diagonal version that is better suited for cases where the two strings are of roughly equal length. SSCA 2 was new to this assessment, with all four kernels being coded in each language.

In addition to these two SSCAs, four more problems were defined, based on an analysis of the Berkeley motifs.¹ We will describe all six problems.

Developing these 12 programs took nearly a year. As desirable as it would be to have multiple skilled programmers perform this task, it would have greatly exceeded available funds. Instead, study costs were contained by having only a single programmer—a member of our productivity assessment team and one of the authors of this report—serve as the test participant. A Ph.D. mathematician by training, he has been programming in C professionally since 1979, wrote the front end of IBM's first C compiler, and has been writing MPI for multi-core systems since 2007. He began programming X10 in 2007 as well, developed the Halverson and Danis⁶ study test bed, and wrote several sections of the X10 tutorial. While it is tempting to view the results from a single programmer as anecdotal, the observed productivity gains were within the range of those found for novice programmers on simpler tasks and are likely quite conservative given the programmer's greater proficiency with C and the lack of an X10 debugger during the study period.

The programs. Six programming problems were defined for this study, each representing a class of parallel

application. Collectively, they span a substantial range of typical small parallel programs:

► *SSCA 1 (first kernel)*. This involves a string-matching problem that was motivated by chromosome and protein studies. One is asked to find the best approximate matches between substrings of a pair of strings of uppercase letters.

► *SSCA 2 (all four kernels)*. This problem involves the efficient representation of a sparse graph whose edge set is known a priori to be randomly distributed across the set of available processors. The first kernel constructs the graph in a format to be used as input by all subsequent kernels. The second kernel extracts edges by weight from the graph representation and forms a list of the edges with maximal weight. The third extracts a series of subgraphs formed by following paths of specified length from a starting set of initial vertices. The final kernel computes a network centrality metric that identifies vertices of key importance along the shortest paths of the graph.

► *Consumer-producer*. This problem uses one process as a server managing a queue for multiple client processes that randomly add to and remove from the queue. Since one buffer is the target for all of the operations, the first programming problem is to minimize contention. Since adding elements to and removing elements from the queue occur at opposite ends, this is the main opportunity for avoiding contention; you can safely do these two operations simultaneously if the queue is large enough.

► *Unbalanced tree search*. In contrast to the previous problem in which coordination is managed through a central server process, here a set of peer processes each manages its own queue of work and, as the need arises, shares work with other processes. The overall task is a breadth-first search of a tree that is known to have leaves at highly varying depths from the root. Each processor is managing its own queue of unvisited tree nodes but may be asked by an idle peer to provide some nodes for it to pursue.

► *Floyd's algorithm*. Given a directed graph with weighted edges that is either acyclic or has only non-negative

weights, we compute the summed weights of the shortest paths between all pairs of vertices.

► *Discrete Fourier Transform.* The algorithm implemented is the original Cooley-Tukey Fast Fourier Transform. The input is an array of 2^n complex numbers, as is the output. Until at least $n=3$ and $2^n=8$, the conventional serial algorithm beats Cooley-Tukey, which means you can profitably distribute the array across $2^n/8$ processes.

Observed Productivity Gain

The accompanying table summarizes the number of days to first successful parallel run and the number of lines of code written for each of the 12 programs. The six programs in X10 required a total of 39 days to develop. The six programs in C with MPI required a total of 129 days to develop. Overall productivity gain due to language (and, secondarily, environment) was, therefore, in excess of 3x. Over the 39 days of writing X10, 6,195 lines of code were written at an overall rate of 159 lines of code per day. Over the 129 days of writing C with MPI, 10,245 lines of code were written at an overall rate of 79 lines of code per day. While we did not measure the performance of these programs, a study by Saraswat et al.⁷ examined the excellent parallel performance of an X10 implementation of an unbalanced tree search program quite similar to ours.

Productivity Contributors

What accounts for the threefold productivity gain found in this study? While some of the gain is a result of the integrated tooling provided by Eclipse and the X10 plug-in, most is attributable to features of the X10 language. These include the task and data partitioning provided by activity and place, the flexible task support provided by *async* and *finish*, and a nicely integrated exception-handling mechanism. In combination, these simplified the expression of parallelism and allowed the program structure to mirror the natural problem structure more closely. In addition, automatic garbage collection removed the chore of managing memory, and object orientation simplified the task of dealing with

complex structures while avoiding the miscalculation of memory references. The following sections provide examples of how these features contributed to productivity.

Expressing parallelism. X10 programs begin with a single activity executing at a single place. An activity is a single sequential thread of execution, and a place is a piece of the system that combines processing power with a share of the partitioned global memory. Any activity can, as needed, initiate new activities at the place where it is currently running or at any other convenient place.

Local threading. As a simple example, SSCA 1 requires reading a pair of

character strings from some peripheral source. If the strings are very long (as they might well be since this problem is designed to scale to arbitrarily long strings), then it is desirable to read the two concurrently. If both streams are to be accessed from a single processor, the X10 code you need is shown in Figure 1.

Simple fork-join parallelism is done in X10 using an *async* statement to do the fork, surrounding the set of activities with a *finish* block to implement the join. The *try* block that surrounds the whole computation catches any errors that might have occurred in the file I/O or activity startup/shutdown.

Figure 1. X10 code needed to read two streams in parallel.

```
try {
    val seq1: Array[Byte](1); // seq1 and seq2 are arrays of bytes
    val seq2: Array[Byte](1); // indexed by a single integer
    finish { // control leaves only when both activities complete
        async{ seq1 = readASequence(path1); } // spawns an activity
        seq2 = readASequence(path2); // runs in parallel
    }
    compare(seq1, seq2); // both sequences are available here
}
catch(e: Exception) { handleTheProblem(e); }
```

Figure 2. C code needed to read two streams in parallel.

```
typedef struct {
    size_t size; /* of the data array in bytes */
    int rc; /* return code for the i/o, normally 0 */
    char buffer[]; /* either the string or an error message */
} ByteArray;

static void do_the_read(void *arg) {
    char *path = (char *)arg;
    pthread_exit(read_sequence(path));
}

...

int rc;
if (my_id == 0) { /* am I the processor with MPI rank 0? */
    pthread_t thread; /* yes, so I am in charge */
    ByteArray *seq1, *seq2;
    rc = pthread_create(&thread, NULL, do_the_read, path1);
    if (rc != 0) perror("Thread creation failed");
    else {
        seq2 = read_sequence(path2);
        rc = pthread_join(thread, (void *)&seq1);
        if (rc || seq1->rc || seq2->rc) {
            rc = report_the_problem(rc, seq1, seq2);
        }
    }
}
MPI_Bcast(&rc, 1, MPI_INT, 0, MPI_COMM_WORLD);
if (rc == 0) return compare(seq1, seq2);
else return rc;
```

For C, the essential problem is how to get a second thread with which the originating thread can share memory. Figure 2 shows the same code as it might be done in C using POSIX threads.

The SPMD model's relatively bad fit surfaces in the test for `my_id==0`, a typical C/MPI idiom. If an I/O error occurs, processor 0 has to let the other processors know there was a problem. That is why there *must* be a broadcast after the strings have been read in, whether or not there were any problems. The function `do_the_read` encapsulates the second thread's mission. Even though it is doing exactly the same read as the root thread, it must be separate.

Remote threading. The previous example is so simple it is easy to underestimate its significance. To get a better sense, consider a slight variation. Suppose that, instead of processor 0 being known to handle most of the work, the processors where the data resides are known only at runtime, as they might be if the data is coming in from a network or from a set of databases. How does the code change? Figure 3 shows the X10 code.

Places in an X10 program, like MPI processes, have a unique integer ID. In Figure 3 the two sequences are being read from the places with IDs `seq1Id` and `seq2Id`. The read operations differ from the original only in that an `at` clause has to be provided to say where each read is to take place. All the remaining code is unchanged, including the error handling, *even though multiple processors might now be involved*. This is a consequence of X10's exception model, which is designed to mirror a program's activ-

ity tree. It allows any activity in that tree to handle exceptions thrown by anything in the subtree below it. (An X10 exception does not automatically include the place it was thrown, although a programmer can readily subclass an exception and insert place information. It is also the case that exceptions have no impact on the children or siblings of the activity in which the exception occurs.) Although C++ introduced the `try/catch` style of exception handling for serial code before 2002, none of the threading frameworks carries it over to sets of multiple processes with the same simplicity of X10 in which both single- and multithreaded cases are supported without change.

Access serialization. A single-server queue is a common pattern where serial writes to a stream of items by one set of players are read by a possibly different set. The usual solution is to buffer the items in an array until they are consumed. Items are added at the end of the array but removed from its beginning.

With multiple players, the potential for race conditions in this scenario is obvious. Remove operations from the array must be serialized with respect to one another, as must adds. On the other hand, unless few items remain, an add and a remove need not be serialized with respect to one another, because they affect opposite ends of the array.

How might one wring out the last bit of performance by allowing an add in parallel with a remove when that is safe? Serializing two adds or two removes needs only a lightweight atomic version of C's postfix `++`. When few items remain, all of the operations

have to be serialized, which requires an atomic block of code to guarantee a thread, once active in the block, will be the only thread in the block until it exits the block.

In X10 the class `AtomicInteger` provides atomic update to a single integer-valued variable. Because this uses hardware support, it is a much lighter-weight operation than protecting a whole block. The keyword *atomic* in front of a block of statements serializes access to the block. X10 thereby provides what is needed to write 650 lines of code that read quite naturally. As of 2002, there was no standardized C API analogous to either `Atomic-Integer` or atomic blocks. (See <http://gcc.gnu.org/onlinedocs/gcc-4.1.1/gcc/Atomic-Builtins.html>. This API is essentially what standardized in C11.) The serialization you need from both atomic blocks and atomic updates can be achieved by using synchronous MPI calls such as `MPI_Recv`, but the resulting code is longer—1,100 lines in our implementation—and considerably more difficult to follow.

For those wondering what could take 1,100 lines, only a very small part of the code supports the queue operations. This is a client-server application. The server has to start, it has to stop, clients have to make requests, there has to be access control (not everyone can use every array), and so on. What makes the X10 version so much shorter is simpler error handling and not having to write both a client and a server API.

Peer coordination. In the previous example, coordination is managed through a central server process. What happens when a set of peer processes each manages its own queue and, as the need arises, shares items with other processes? Here each process acts as both client and server for some subset of the active processes. The example we programmed performs a breadth-first search of a tree that is known in advance to be substantially unbalanced. Each processor manages its own queue of unvisited nodes, but, to spread the workload more evenly, a processor may be asked to ship some unhandled nodes to another processor to work on.

The biggest issue here is termina-

Figure 3. X10 code to read using processors known only at runtime.

```
try {
    val seq1: Array[Byte](1); // seq1 and seq2 are arrays of bytes
    val seq2: Array[Byte](1); // indexed by a single integer
    val seq1Home = Place.place(seq1Id); // the id names a Place object
    val seq2Home = Place.place(seq2Id); // that reifies the processor
    finish {
        async { seq1 = at(seq1Home) readASequence(path1); }
        seq2 = at(seq2Home) readASequence(path2);
    }
    compare(seq1, seq2);
}
catch(e: Exception) { handleTheProblem(e); }
```

tion: how does a process know when not only it is out of work, but all the other processes are also out of work? Algorithms to handle this (at least at lower orders of scale) date back some 30 years to Dijkstra et al.⁴ and have been refined a number of times since. The SPMD solution is based on them. The X10 solution is entirely different, relying on threading on demand.

Let's look at X10 first. A single X10 activity begins the action by asking each peer in a set to start searching a part of the tree. If one of these peers runs out of work, it has a list of "nearby" peers it is permitted to ask for more work. It does this by sending a reference to itself to each of its neighbors and then lets its current activity die. The object sending these requests, however, continues to exist. Thus, if one of the idle peer's neighbors has work to spare, it can spawn a new activity at the idle peer's place in which that peer can resume work. A little handshaking is involved if the peer wants to ensure only one neighbor's work is taken on, but otherwise that is the whole story. Since the root activity has spawned the multiple peer activities as *asyncs* within a single *finish* block, it can merely exit that block when none of these *asyncs* remains active thereby finishing the job.

The C solution is entirely different. There is no peer object at each processor ready to receive a remote process call, and even if there were, there is no analog of the X10 *finish* block. The naïve way of handling an idle peer is to let it hang on a synchronous receive, but then who is responsible for knowing when to send that peer the "all done" message? That is exactly the problem Dijkstra et al. were addressing by a clever method of polling all the peers.

In the C solution each processor must perform three activities: the primary work (namely, searching and processing its share of the tree); listening for a termination control message; and listening to its neighbors for shared work requests. By using two nonblocking receives to listen for the control and sharing communications, each processor has, in effect, three concurrent activities: the primary activity and the two pending receives. Each processor's main loop is



The lack of introspection in C is a significant problem as well. In an object-oriented language such as X10, objects know how big they are and what outstanding references exist.



responsible for dealing with incoming control and sharing requests in a timely manner. Most of the complexity is involved in deciding who sends what control information to whom.

Managing memory and handling errors. The issues discussed previously pertain to the fit between the natural threading models of our programs on the one hand and the X10 APGAS and MPI SPMD parallel models on the other. There are also several well-known C-related shortcomings that impacted our productivity in writing serial sections of code. These shortcomings are not present in X10. On average, we encountered about six or more of these well-known problems per 1,000 lines of C code, and the overall impact was significant. Many required additional effort because they did not reveal themselves close to the point where a correction was needed.

Memory leaks. X10, like many recent languages, has automatic garbage collection. The need to explicitly manage storage has long been recognized as one of the serious impediments to C coding productivity. You might think that in a problem as simple as, for example, SSCA 1, with around 1,200 lines of relatively straightforward string-matching code, this would not be an issue—but it is, especially if the code is to run for a long time (or be incorporated into a frequently called library routine). Here memory leaks need to be found and eliminated, a task that can easily take one or two days for the places found in our code where memory was allocated but never freed.

Getting memory references right. The lack of introspection in C is a significant problem as well. In an object-oriented language such as X10, objects know how big they are and what outstanding references exist. In C those calculations must be done by hand. It is not that, using X10, you cannot miscompute an array extent. It is still possible, but X10 will catch an out-of-bounds access to an array at runtime at the point of the access. Even skilled C programmers are unlikely to take the time to bulletproof every array access. Since errant accesses are usually detected far from the point where the code needs to be corrected, these errors are not easily found and fixed.

Error handling. The lack of a convenient exception mechanism in C forces programmers to be more verbose. This surfaced in Floyd's algorithm, for example, when our programmer wanted a generic input stream to read an ASCII stream of numeric values. His API had an entry that tokenizes the stream, converts the tokens to the appropriate numeric type, and assures the value is legal. Clearly there are a number of problems the stream can encounter. The question is how to handle these errors.

In the case of errors in X10, an exception is thrown whose type identifies the problem encountered. An application can avoid providing special code for detecting generic errors such as an unexpected end-of-file that is discovered by lower-level functions, because they, too, can signal errors by throwing exceptions. The application can therefore concentrate on signaling semantic errors in the content.

For most throwaway code, error handling is not a serious issue, but for production code, and for complex communications patterns such as Dijkstra et al.'s termination algorithm, it certainly is. C's signaling mechanism is best suited for expert use. C's problems, however, run even deeper in a multithreaded SPMD world. Consider the standard library routine `strtoll` the stream calls to convert a found token to a long integer. Here is the discussion of `strtoll`'s error indications as given by its "man" page:

"the `strtol`, `strtoll`... functions return the result of the conversion, unless the value would underflow or overflow. If no conversion could be performed, 0 is returned and the global variable `errno` is set to `EINVAL` (the last feature is not portable across all platforms). If an overflow or underflow occurs, `errno` is set to `ERANGE`..."

Consider the code a C application needs in order to deal with the various possible errors. Should the code make sure to zero `errno` before calling `strtoll`? After all, it might be non-zero because of an entirely unrelated earlier problem. For the code that checks `errno` to understand what happened, it is, moreover, not enough just to check that `errno` is non-zero, because the error may be an

I/O error, a parsing error, or a range error. Nor can you be sure `errno` is thread-safe—it is not on all systems. What then? And where in the application should you clear `errno`, whose value is a global? Which of the other processes need to be made aware of the problem, and how should they be made aware?

Conclusion

There are very good reasons for C and MPI's dominance in the parallel-programming community. They are superbly documented, elegant, clean designs that have been carefully implemented. In the hands of an experienced, disciplined professional, they provide a level of control that is simply not available elsewhere. Neither C nor MPI is a standing target: both continue to be improved, even though they are now mature technologies.

How often, however, do the benefits of C/MPI outweigh its costs? Through all three of our studies, and particularly in this final one, we have seen substantial benefits—ranging from two to six times faster development to first successful parallel run—from using a higher-level language and programming model. These productivity benefits might be even greater for large programs that need to be maintained for years.

X10 and its APGAS programming model are now being explored in many research and university settings. While the language may or may not cross over into mainstream use, it is likely the qualities that made it so much more productive for us will likely become well established in the parallel community: flexible threading, automatic garbage collection, runtime type-driven error checking, partitioned global memory, and rooted exception handling are all valuable. We hope our experiments encourage those looking to improve parallel-programmer productivity to seriously study X10's design and its benefits.

Acknowledgments

This work was supported by the Defense Advanced Research Projects Agency under its Agreement No. HR0011-07-9-0002. The authors thank Catalina Danis, Peter Malkin, and John Thomas, members of IBM's

HPCS productivity assessment team, for their invaluable contributions to this research. 

Related articles on queue.acm.org

Unlocking Concurrency

Ali-Reza Adl-Tabatabai, Christos Kozyrakis and Bratin Saha

<http://queue.acm.org/detail.cfm?id=1189288>

The Ideal HPC Programming Language

Eugene Loh

<http://queue.acm.org/detail.cfm?id=1820518>

Software Transactional Memory: Why Is It Only a Research Toy?

Calin Cascaval, Colin Blundell, Maged Michael, Harold W. Cain, Peng Wu, Stefanie Chiras, and Siddhartha Chatterjee

<http://queue.acm.org/detail.cfm?id=1454466>

References

1. Asanovic, K., Bodik, R., Catanzaro, B.C., Gebis, J., Husbands, P., Keutzer, K., Patterson, D., Plishker, W., Shalf, J., Williams, S. and Yelick, K. The landscape of parallel computing research: a view from Berkeley. Technical Report No. UCB/EECS-2006-183. Electrical Engineering and Computer Sciences, University of California at Berkeley, 2006; <http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.pdf>.
2. Bader, D., Madduri, K., Gilbert, J., Shah, V., Kepner, J., Meuse, T. and Krishnamurthy, A. Designing scalable synthetic compact applications for benchmarking high productivity computing systems; <http://www.cse.psu.edu/~madduri/papers/SSCA-CTWatch06.pdf>.
3. Danis, C. and Halverson, C. The value derived from the observational component in an integrated methodology for the study of HPC programmer productivity. In *Proceedings of the Third Workshop on Productivity and Performance in High-End Computing*, (2006), 11–21.
4. Dijkstra, E., Feijen, W. and van Gasteren, A. Derivation of a termination detection algorithm for distributed computations. *Information Processing Letters* 16, 5 (1983), 217–219.
5. Ebcioğlu, K., Sarkar, V., El-Ghazawi, T. and Urbanic, J. An experiment in measuring the productivity of three parallel programming languages. In *Proceedings of the Third Workshop on Productivity and Performance in High-End Computing*, (2006), 30–36.
6. Halverson, C. and Danis, C. Towards an ecologically valid study of programmer behavior for scientific computing. In *Proceedings of the First Workshop on Software Engineering for Computational Science and Engineering*, (2008).
7. Saraswat, V.A., Kambadur, P., Kodali, S., Grove, D. and Krishnamoorthy, S. Lifeline-based global load balancing. In *Proceedings of the 16th ACM Symposium on Principles and Practice of Parallel Programming*, (2011), 201–212.

John Richards is a research manager in IBM's Watson Group and holds an appointment as Honorary Professor in the School of Computing at the University of Dundee, Scotland.

Jonathan Brezin was trained as a mathematician and held positions at the University of Minnesota and the University of North Carolina in the 1960s and 1970s. He later joined IBM Research, from which he retired this year.

Cal Swart joined IBM Research in 1982 and is currently a senior technical staff member in the IBM Watson Group. He is a member of Watson Life Research, exploring new applications of cognitive computing.

Christine Halverson is an independent consultant in Silicon Valley. Formerly she worked at IBM Research where she spent five years working on the DARPA HPCS initiative studying parallel programmers.



SIGAI

ACM SPECIAL INTEREST GROUP ON ARTIFICIAL INTELLIGENCE

Announcing SIGAI's Career Network and Conference: CNC

The first annual SIGAI CNC will showcase high-quality research from graduating Ph.D.s and postdocs. CNC will also include a wide range of opportunities for career development and mentoring.

Join us January 26, 2015 in Austin, Texas, close by AAAI-2015.

See the call for finishing grad students, postdocs, and employers at

<http://sigai.acm.org/cnc/cfp.html>

AI Matters: The ACM SIGAI Newsletter

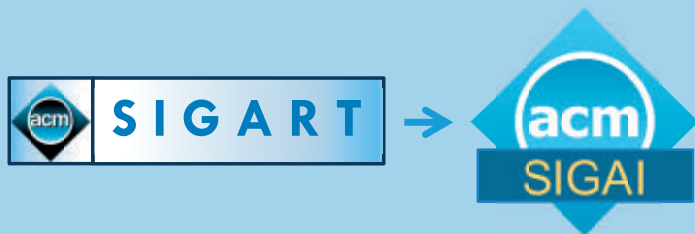
AI Matters is SIGAI's new quarterly newsletter. It features work and announcements of interest to the AI community.

<http://sigai.acm.org/aimatters/>

Our mission:

- **Promote and support** the growth and application of AI principles and techniques throughout computing
- **Sponsor** or co-sponsor high-quality, AI-related conferences
- **Sponsor** recognized AI awards
- **Support** important journals in the field
- **Provide scholarships** to student members to attend conferences
- **Promote AI** education and publications through various forums and the ACM digital library

SIGART is now SIGAI



Join us today to support the increasingly interdisciplinary AI community!

Find us at <http://sigai.acm.org/>

Pixels labeled with a scene's semantics and geometry let computers describe what they see.

BY STEPHEN GOULD AND XUMING HE

Scene Understanding by Labeling Pixels

PROGRAMMING COMPUTERS TO automatically interpret the content of an image is a long-standing challenge in artificial intelligence and computer vision. That difficulty is echoed in a well-known anecdote from the early years of computer-vision research in which an undergraduate student at MIT was asked to spend his summer getting a computer to describe what it “saw” in images obtained from a video camera.³⁵ Almost 50 years later researchers are still grappling with the same problem.

A scene can be described in many ways and include details about objects, regions, geometry, location, activities, and even nonvisual attributes (such as date and time). For example, a typical urban scene (see Figure 1) can be described by specifying the location of the foreground car object and background grass, sky, and road regions. Alternatively, the image could be summarized as a street scene. We would like a computer to be able to reason about all these aspects

of the scene and provide both coarse image-level tags and detailed pixel-level annotations describing the semantics and geometry of the scene. Early computer-vision systems attempted to do just that by using a single unified model to jointly describe all aspects of the scene. However, the difficulty of the problem soon overwhelmed this unified approach, and, until recently, research into scene understanding has proceeded along many separate trajectories.

Along one of them, researchers aim to provide a high-level summary or categorization of a scene using a small number of tags (such as city and forest) without explicitly identifying the objects within it. Another, known as “object detection,” aims to locate discrete objects (such as cars or pedestrians) in a scene by placing a bounding box around the objects. Face-detection algorithms in today’s digital cameras and smartphones perform this task. However, these approaches do not provide detailed object outlines and fail to reason about the image as a whole.

Perhaps closest to the long-term scene-understanding goal is yet another trajectory that aims to produce annotations for the entire image, at the pixel level. Such a pixel-labeling, or semantic-segmentation, approach to scene understanding is our main focus

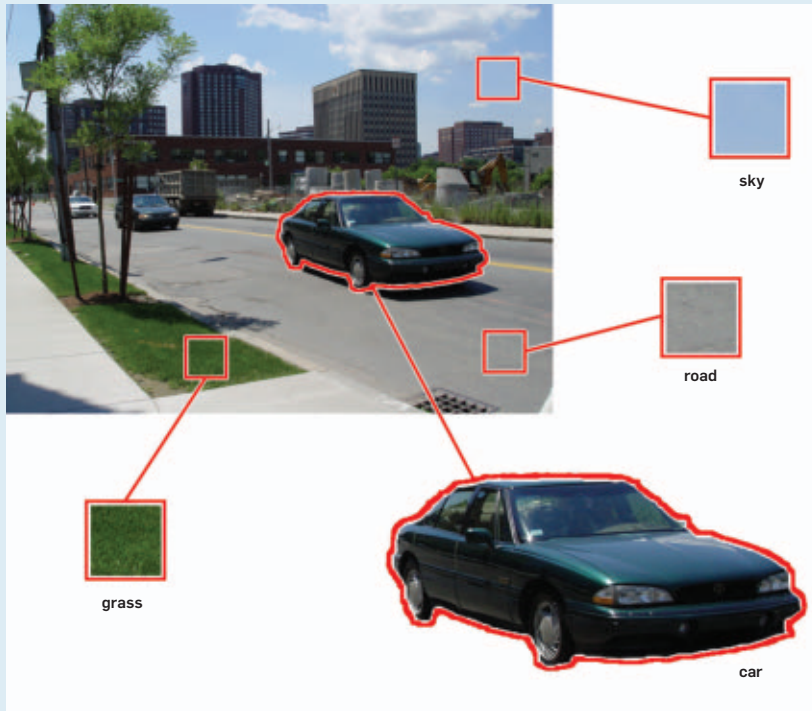
» key insights

- Recent progress on image understanding, a long-standing challenge of AI, is enabling numerous new applications in robot perception, surveillance and environmental monitoring, content-based image search, and social-media summarization.
- One approach to image understanding is to label every pixel in an image with a category label using probabilistic models known as CRFs that can handle uncertainty and propagate contextual information across the image.
- Improved machine-learning techniques, more-powerful machines, and ever-growing volume of data are getting us closer to machines that are able to see and understand the world the way humans do.



Figure 1. Understanding a typical urban scene in terms of foreground objects (“things”) and background regions (“stuff”).

The background regions are easily recognized from local color and texture cues, whereas foreground objects like the car require complex shape and appearance models.



here. Objects and background classes are segmented into discrete nonoverlapping regions and a label provided for each region, or, equivalently, all the pixels within it. In addition to labeling each pixel with a class label, different instances of each object can also be labeled with a unique identifier, so two adjacent cars are treated as, for example, disjoint objects. Such multi-class/multi-instance approaches are at the cutting edge of contemporary scene-understanding research. Hierarchical segmentation, so-called “scene parsing,”³⁷ can produce an even more refined view of the scene by breaking objects into component parts; for example, a car can be broken down into wheels, body panels, and windows.

These pixel-labeling approaches work with a predefined set of class labels that dictates the categories of objects and types of scenes the model can recognize. The labels can be semantic (such as grass, road, sky, and car) or geometric (such as horizontal, vertical, and slanted) and tuned for different scene types; for example, describing an indoor scene requires

very different semantic and geometry classes from an outdoor scene. Scaling up the number and diversity of recognizable categories is a core thrust in contemporary research.

Most methods for pixel labeling use a probabilistic model known as a conditional Markov random field, or CRF, which provides a formal framework for encoding the complex relationship between the visual appearance of a scene and the underlying semantic (or geometric) labels. Moreover, the formalism admits efficient inference algorithms and allows model parameters to be learned from data. Some recent research (such as Heitz et al.,¹⁵ Hoiem,¹⁸ Li et al.,²⁸ and Yao et al.⁴⁰) has sought to reunite the different research trajectories for scene understanding into a single coherent model incorporating high-level scene labeling, object segmentation, and geometric reasoning. This unified view allows, for example, constraints (such as object support and scale) to be expressed naturally by linking the semantic and geometric aspects of a scene. These state-of-the-art holistic models build on the pixel-

labeling approach and employ the CRF framework to integrate the various aspects of the scene.

Here, we outline CRF model variants for scene understanding, showing how they exploit various assumptions (such as that cars typically appear on roads) about real-world scenes. Note, however, the problem of scene understanding remains wide open, with new innovations being introduced regularly. We also offer example results from a baseline CRF model on standard scene understanding datasets to demonstrate the capabilities and weaknesses of today's scene-understanding models.

Pixel Labeling

In pixel labeling, each pixel in an image is assigned a class label from a predefined set (such as grass, tree, road, car, and person). The assumption is that each pixel belongs to a single category of interest and that category can be identified unambiguously.

One approach is to classify each pixel individually without regard to the label assigned to other pixels in the image. However, as we show, treating each pixel independently can produce highly inconsistent results. A more sophisticated approach adopted by state-of-the-art scene-understanding algorithms is to consider the labeling of the pixels jointly by defining a random field over them.^{13,34} Such an approach has enjoyed much success in recent years due to effective inference and learning algorithms.^{3,22,31}

Conditional Markov random fields.

CRFs were first introduced in the area of natural-language processing but have since found application in a range of machine-learning tasks.²⁵ Their key benefit is to provide a principled probabilistic framework for describing the relationship between related output variables (such as labels for pixels in an image) as a function of observed features (such as pixel colors). They are thus ideal for integrating multiple visual cues and combining related scene-understanding problems. Moreover, CRFs admit a compact representation and provide efficient (approximate) algorithms for inference and learning.

Formally, let $\mathbf{y} = (y_1, \dots, y_n)$ be a vector of discrete random variables (output variables) we are interested in predict-

ing from some observed features $\mathbf{x}$. We assume each variable can take a label from a predefined finite set of labels $y_i \in L$. In pixel labeling the variable y_i denotes the label assigned to the i -th pixel in the image and n is the total number of pixels in the image. Typically, the features $\mathbf{x}$ are represented by real-valued vectors associated with the individual pixels or image regions. For example, $\mathbf{x}$ may encode color and texture cues, as discussed later.

The CRF model defines a probability distribution over the output variables, given the observed features via an energy function $E(\mathbf{y}; \mathbf{x})$ as follows

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \exp\{-E(\mathbf{y}; \mathbf{x})\} \quad (1)$$

where $Z(\mathbf{x})$ is the so-called partition function that ensures the probability distribution is normalized correctly, or sums to one. Note, in general, computing the partition function is intractable since computation involves summing over the exponentially many assignments to $\mathbf{y}$. Fortunately, computation of the partition function is not necessary for inferring the most likely labeling, as we show here.

CRFs are compactly represented by decomposing the energy function $E(\mathbf{y}; \mathbf{x})$ as the sum of smaller clique potentials

$$E(\mathbf{y}; \mathbf{x}) = \sum_c \psi_c(\mathbf{y}_c; \mathbf{x}). \quad (2)$$

Here, each clique potential $\psi_c(\mathbf{y}_c; \mathbf{x})$ is a real-valued function defined over a subset of the random variables. We use the shorthand $\mathbf{y}_c$ to indicate the subset of variables in the clique, or the scope of the “clique potential.”^a Roughly, a clique potential encodes a numerical score for every joint assignment to the variables within its scope (ignoring all other variables in the model). Probabilistic influence propagates across the model via clique potentials that share variables, or have overlapping scope.

The number of variables within a clique potential defines the order of the model. Higher-order models can contain a very large number of variables within each clique. However, without appropriate structure, these models result in intractable inference problems. A common model for pixel

labeling involves only unary and pairwise terms

$$E(\mathbf{y}; \mathbf{x}) = \sum_{i=1}^n \psi_i^u(y_i; \mathbf{x}) + \sum_{ij \in \mathcal{E}} \psi_{ij}^p(y_i, y_j; \mathbf{x}). \quad (3)$$

There is one unary term associated with each pixel in the image, while the pairwise terms are defined over a set of pixel pairs $\mathcal{E}$. The set is usually sparse, containing only pairs of pixels that are neighbors in the image. Figure 2 is an example of a grid structured CRF defined over a four-connected neighborhood. In such a model, the label for each variable is influenced by some local features, as well as by the labels from surrounding variables. This influence can be captured by the unary and pairwise terms, respectively. In particular, the pairwise terms permit the encoding of smoothness assumptions; that is, a pixel in an image is likely to belong to the same object as its neighbors. Some recent research considers fully connected graphs in which a pairwise term exists between all pairs of pixels. This term between pairs of pixels allows long-range interactions to be captured but is tractable only when the pairwise terms take a specific form involving Gaussian kernels.²³

Researchers are often interested in

only the most likely interpretation of a scene; probabilistically, this is known as maximum a posteriori, or MAP, inference and requires solving

$$\hat{\mathbf{y}} = \operatorname{argmax}_{\mathbf{y}} P(\mathbf{y}; \mathbf{x}). \quad (4)$$

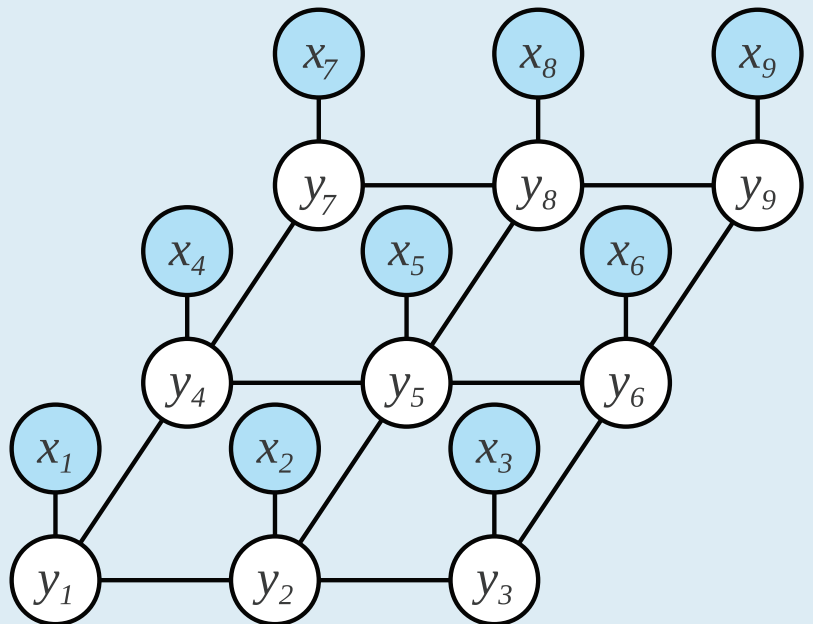
Since the partition function $Z(\mathbf{x})$ does not change with different assignments to $\mathbf{y}$, the most likely scene interpretation can be found by solving the equivalent energy-minimization problem

$$\hat{\mathbf{y}} = \operatorname{argmin}_{\mathbf{y}} E(\mathbf{y}; \mathbf{x}). \quad (5)$$

For CRFs found in pixel-labeling problems, very fast algorithms have been developed for approximate energy minimization.^{3,21,22} The most ubiquitous is a method known as “graph-cuts,”^{3,21} a move-making algorithm that starts with an initial assignment to the variables and then iteratively solves a sequence of binary optimization problems that progressively improve on the solution at hand.

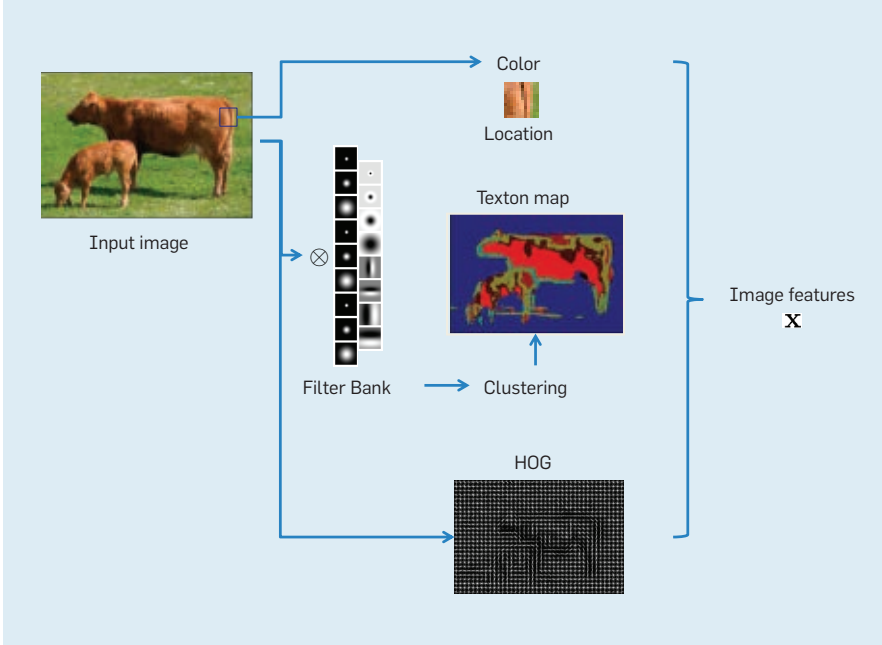
The clique potentials $\psi_c(\mathbf{y}_c; \mathbf{x})$ can be specified in a variety of ways but often include parameters that control the value of the potential as a function of the variables and observed features; for example, the potential could

Figure 2. Graphical representation for a grid-structured pairwise CRF over nine random variables $\{y_1, \dots, y_9\}$, with each variable represented by a node, direct pairwise correlations between variables by edges, and observed features $\{x_1, \dots, x_9\}$ as shaded nodes.



^a Formally, c indexes a sparse subset of elements from the power set of $\{1, \dots, n\}$.

Figure 3. Typical image features used in the potentials of CRF models, including color, image location, filter outputs, and HOG; the filter outputs can be used to generate “texton features,” or texture prototypes.



be defined as a linear combination of feature values and parameterized by a weight vector. Alternatively, the potential could be defined as the output of a decision-tree classifier and parameterized by node splits and leaf probabilities. The main difficulty in designing CRF models for scene understanding is in learning these parameters. Maximum-likelihood approaches cannot be applied due to the intractability of calculating the partition function. Many approximate learning techniques have been attempted. Popular among them are “piecewise learning with cross-validation”³⁴ and “pseudo-likelihood learning.”¹² In recent years, max-margin learning approaches have shown success.³¹ However, the most effective strategy for CRF parameter learning is still an open research question.

Features. One advantage of the CRF formulation is that a variety of image features can be used in the clique potentials and treated as observed values, or fixed for each image. Treating image features as observed values provides flexibility in scene representation and simplifies the model structure, as the model does not need to explicitly capture the probability distribution over these features. Different types of image features are tailored to the visual properties of the object classes of interest.

Computer-vision researchers af-

fectionately divide object classes into two broad categories: “things” and “stuff.” Background classes (stuff) are easily recognized from local appearance; objects (things) need additional cues (such as shape). Figure 1 reflects this distinction with three examples of background classes—grass, road, and sky—and a foreground object—car.

The local appearance features are color and texture cues. Two commonly used color spaces are RGB for its simplicity and CIE-Lab for color similarity closer to human perception.³⁵ For image features associated with regions, color histograms or summary statistics are often adopted to describe their overall appearance. The texture cues aim to capture repetitive local patterns in images and are usually extracted by filter banks and the distribution of their outputs. One widely used example is the texton, or small texture prototype, feature^{30,34} that first clusters the filter bank outputs into a texton codebook, then represents an image patch as a histogram over the texton code words; see Figure 3 for several types of commonly used image features.

For foreground objects, however, local-appearance features usually have much greater variation than background classes. More distinctive and stable features are needed for reliable recognition. These features look

beyond individual pixels and encode structural information; for example, object-shape information can be represented using the histogram of oriented gradient, or HOG, feature.⁵ In addition, other higher-level features derived from top-down processing, discussed later, can be used to represent object properties. Moreover, global appearance features (such as average image color) can provide context for certain objects; for example, sheep tend to appear in greenish images, whereas cars tend to appear in grayish ones.

Image features can be concatenated as a feature vector $\mathbf{x}$ for each pixel and incorporated into the potential functions in CRF models. In particular, the unary potentials, which play a critical role in mapping image cues into labels, can be defined like this

$$\psi_i^U(y_i = k; \mathbf{x}) = \mathbf{w}_k^T \phi_i(\mathbf{x}) \quad (6)$$

where w_k is a weight coefficient for label class k and $\phi_i(\mathbf{x})$ is a (non-)linear mapping of the features for pixel i . The weights w_k are determined from training data using, for example, structured prediction learning algorithms.³¹

The role of context. Context plays an important role in natural human recognition of objects and scene understanding. Consider the urban scene in Figure 4; we instantly recognize the street and somewhat surprisingly each of the cars in it despite the cars being fewer than 32 pixels high. This recognition is because the weak local evidence for the cars is compensated by strong contextual cues based on the scene’s spatial layout.

Contextual information can be incorporated into a scene-understanding model in a number of ways. The simplest form of context is the statistical correlation between features and class labels; for example, blue pixels are more likely to be sky or water than grass or trees. Likewise, green pixels are more likely to be grass or trees than sky or water. This is exactly the type of context captured through the unary potentials in a CRF model. However, much more sophisticated contextual assumptions can also be encoded.

The co-occurrence of object classes (such as that cars often co-occur with road) can also be included in scene-understanding models.^{32,40} Here, the

relative positioning of the two objects is irrelevant; all that is important is they often appear together. This sort of context can be incorporated into a CRF by including binary variables z_k that indicate the presence or absence of each object class. Pairwise potentials linking these object variables z_k to the pixel variables y_i enforce consistency; that is, if an object is present, then some pixels must be labeled as that object and vice versa. Pairwise potentials between two object variables z_k and z_l then capture co-occurrence preferences. However, note that the resulting CRF graph is no longer a regular grid over pixels. More important, the very efficient graph-cut technique cannot be used for inference, so methods must resort to slower inference algorithms; see, for example, Yao et al.⁴⁰

Class co-occurrence is a crude measure of context and does not capture the relative position of the objects. It can thus result in false detections placed inconsistently with respect to one another (such as a car floating in the trees). The spatial location of objects relative to the adjacent background¹⁶ or relative location of objects across the scene¹² allows such errors to be corrected automatically by designing potential functions that encode this information in the model. Unfortunately, modeling the relative location of objects directly requires non-local reasoning that can make inference intractable. A global entity (such as the location of the horizon for outdoor scenes¹⁹ or box-structure layout of a room for indoor scenes^{14,39}) can be used to link objects indirectly and simplify reasoning. Here, additional variables (such as v_{hz} for the location of the horizon) are introduced into the CRF formulation (Equation 2) and linked via pairwise potentials to the pixel label variables y_i .

Pixels versus superpixels. An image is represented within a computer as a rectangular array of pixels, so associating labels with individual pixels for scene understanding is a natural choice. Moreover, the regular image structure is convenient for constructing CRFs with pairwise terms over adjacent pixels. However, pixels themselves are an artifact of the imaging process and do not necessarily reflect the underlying structure or complexity

of a scene; for example, a one-megapixel photograph and an eight-megapixel photograph convey essentially the same information for the purpose of scene understanding despite the latter containing eight times as many pixels, hence requiring a bigger model. Pixels are noisy indicators of class; for example, in Figure 5, two adjacent pixels can have very different colors despite belonging to the same object.

A potential way to avoid an unnecessarily large number of variables is to model the image in terms of superpixels, or small contiguous regions of consistent appearance. Here, y_i in Equation 3 denotes the label assigned to the i -th superpixel, and n is the number of superpixels, as in Fulkerson et al.⁹ A superpixel representation of an image can be generated by clustering adjacent pixels with similar color to produce numerous small regions.^{4,8,33} Since superpixels are often much smaller than the objects of interest, the process is called “oversegmentation.” Traditional oversegmentation methods are often slow and can result in highly irregular regions that do not honor object boundaries. A number of recent research efforts have attempted to produce superpixels with a more regular shape¹ and better alignment to true boundaries in the scene.

Many algorithms provide a way to control the number of superpixels. Figure 5 includes three different oversegmentations of an image into superpixels. The first (Figure 5b) produces too few superpixels, and much detail in the image is lost. However, the last (Figure 5d) maintains all essential structure in the image and achieves a 46 times re-

duction in the number of entities needed to describe the scene, from 68,160 pixels to 1,476 superpixels.

In addition to reduced model size, superpixels provide spatial support for computing features,^{9,17} making these features less susceptible to noise. However, there are also a number of drawbacks to representing an image with superpixels rather than pixels. First, superpixels do not conform to a regular neighborhood structure, making it difficult for a programmer to weight the influence of each pairwise term in a CRF model. Second, superpixels commit to region boundaries as a pre-processing step, and these boundaries may not coincide with the true object boundaries.

The best of both worlds—pixels and superpixels—can be achieved by representing the image in terms of pixels but using superpixels to provide additional features or enforce higher-order consistency constraints.²⁰ For example, the model can penalize labelings where pixels within a superpixel disagree on their class label. Such penalties can be implemented in a CRF model through a higher-order Potts potential

$$\psi_c^H(\mathbf{y}_c) = \begin{cases} 0 & \text{if } \exists \ell \in \mathcal{L} \text{ s.t. } y_i = \ell \text{ for all } i \in c \\ \lambda & \text{otherwise} \end{cases} \quad (7)$$

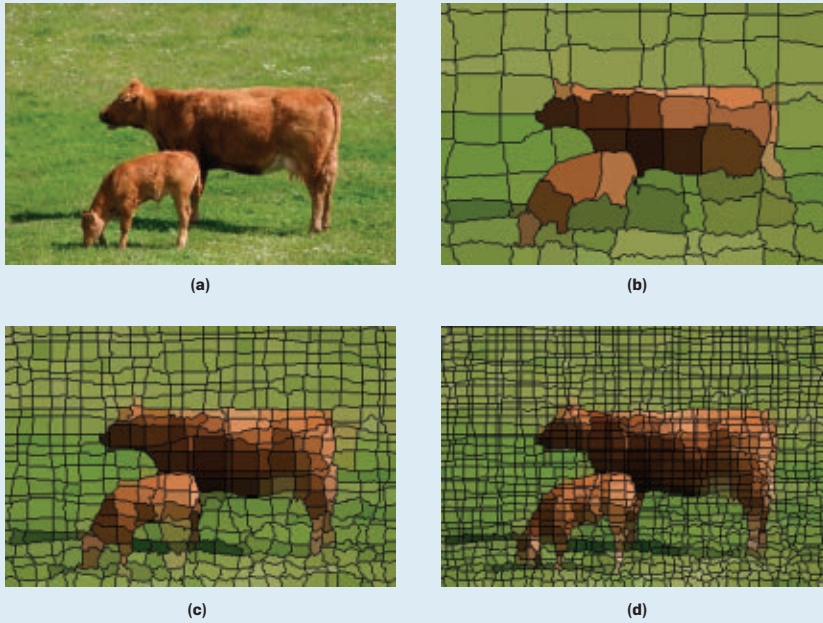
where a cost of $\lambda > 0$ is paid if not all variables within clique c take the same label. In this way, the algorithm prefers boundaries defined by the superpixels, but, given enough evidence to the contrary, this assumption can be overridden.

Superpixels can also be used within iterative inference procedures (such as move-making algorithms^{10,11}) or data-

Figure 4. Example low-resolution image showing the importance of context for object detection; recognizing the highlighted object is difficult in isolation but in the context of a street scene is easily identified as a car.



Figure 5. An image and various oversegmentations of the image into superpixels; the number of superpixels controls the trade-off between model complexity and representation accuracy.



driven Markov chain Monte Carlo algorithms.³⁸ Here, the superpixels guide the inference procedure by proposing changes to the boundaries of objects for the current interpretation of the scene at hand.

Top-down vs. bottom-up. CRF models provide a unified framework for integrating image cues and prior assumptions over scene-label configurations. Image cues used for defining unary clique potentials or forming superpixel representations are usually referred to as “bottom-up information.” These bottom-up cues describe the similarity between pixels and can be used to generate informative proposals for the labeling task (such as larger homogeneous regions). We view the class-dependent unary potentials as part of the bottom-up process, though they also include the label information. The most common prior assumptions over labels are defined on neighboring (super)pixels (via pairwise terms in the CRF) and represent a soft smoothness constraint on region labels. While this constraint is an effective prior on region labels, it lacks the expressive power needed for object-specific information (such as global shape and pose). We refer to the priors on these latter properties as “top-down information,”

as they encode knowledge at a level beyond pixels and simple image regions.

Object shape is probably the most important type of object-level information. Shape priors are usually represented as either rigid or deformable masks that assign each pixel a cost of belonging to the foreground object corresponding to the mask^{2,27} (see Figure 6). Deformable masks can capture pose variation suitable for object classes (such as people and animals) and are often implemented through a part-based model where individual components of the mask can move relative to one another. As shape is a global property of foreground regions, shape priors naturally lead to large cliques and therefore more complex inference algorithms. In practice, some mask representations decompose into a sum of local unary or pairwise terms; for instance, given a matched shape mask, pixels lying within the mask can be encouraged to take specific labels by modifying costs within the corresponding unary potentials.

One common strategy for incorporating object priors is to use object detection to generate object-instance proposals (such as bounding boxes) and define object-shape potentials on those regions.^{11,24} The resulting energy func-

tion usually has higher-order terms, as in Equation 2, in which the clique c is defined by pixels within the object-support regions, and the term favors coherent labeling of those pixels.

Higher-order energy functions can be optimized by generalized graph-cut inference algorithms²⁰ or transformed into pairwise models by adding auxiliary variables. Alternatively, both bottom-up and top-down information can be used in the design of the move-making algorithms that iteratively propose energy-reducing changes to variable assignments; for example, object-detection outputs can generate powerful top-down moves, overcoming fragmented object-labeling results from pure bottom-up-driven segmentation.¹¹

As discussed earlier, some recent approaches take a holistic view toward scene understanding by integrating scene segmentation with multiple scene-analysis tasks (such as scene classification, object detection, and depth/layout estimation^{15,18,40}). In such a framework, several levels of top-down information, including geometry, candidate object bounding boxes, and scene category, are incorporated into a single CRF energy function. Concretely, variables are introduced to represent these quantities, and their relationships to pixel labels are modeled through pairwise or higher-order potential functions. The overall problem can be decomposed into smaller tasks and solved in an alternating fashion or treated as a unified objective to be optimized jointly.

Datasets and Software

Many software implementations of scene-understanding algorithms are freely available. Two notable examples are the Darwin software framework^b and the Automatic Labeling Environment (ALE),^c providing infrastructure for scene understanding via pixelwise CRF models, as described earlier. The most basic models consist of unary and pairwise potentials. The unary potentials are constructed using local and global appearance features. The pairwise clique potentials are defined on either four-connected

^b <http://drwn.anu.edu.au>

^c <http://cms.brookes.ac.uk/staff/PhilipTorr/ale.htm>

or eight-connected neighborhoods around each pixel and often involve a contrast-sensitive smoothness prior³⁴ that discourages adjacent pixels from taking different labels when the pixels are similar in color.

Moreover, many well-labeled datasets are readily available, with many researchers using them to develop and compare scene-understanding algorithms. To give a flavor of the results that can be achieved, consider the following results on two standard datasets from the Darwin software framework:

*The Stanford Background Dataset (SBD)*¹⁰ consisting of 715 images of rural, urban, and harbor scenes. Images are labeled from two different label sets: the first captures semantic class and includes seven background classes (sky, tree, road, grass, water, building, and mountain) and a single foreground object class; the second captures scene geometry (sky, vertical, and horizontal). Each image pixel is allocated two labels, one semantic and one geometric; and

*The Microsoft Research Cambridge (MSRC) dataset*³⁴ consisting of 591 images. Pixels are labeled with one of 23 different classes. However, due to the rare occurrence of the horse and mountain class, they are often discarded. Pixels not belonging to one of the remaining 21 categories are ignored during training and evaluation. One drawback of this dataset is the ground-truth labeling is rough and often incorrect near object boundaries. Nevertheless, the dataset contains a diverse set of images and is widely used.

As scene-understanding research matures, larger and more diverse datasets are becoming more important for applying existing scene-understanding algorithms and inspiring new ones. The PASCAL Visual Object Classes (VOC) dataset⁶ is a very large collection of images annotated with object-bounding boxes and pixelwise segmentation masks for 20 different (foreground) object categories. It contains approximately 20,000 images organized into numerous challenges, with training, validation, and evaluation image sets pre-specified. Another large dataset of interest to scene-understanding researchers is the SIFT Flow dataset,²⁹ a subset of outdoor images from the LabelMe im-

age repository (<http://labelme.csail.mit.edu>), which contains 2,688 images annotated using 33 diverse object and background categories. Performing well on both these datasets requires a combination of many of the techniques described earlier.

Accuracy of scene-understanding algorithms can be evaluated by many measures, including sophisticated boundary-quality metrics and intersection-over-union (Jaacard) scores. The simplest measure computes the percentage of pixels that were correctly labeled by the model on a “hold out,” or separate, set of images, referred to as the “test set” or “evaluation set.” As is standard practice when evaluating machine-learning algorithms, these images should not be viewed during training of the model parameters. Formally, we can write

$$\text{acc} = \frac{\sum_{i=1}^n \mathbb{I}[\hat{y}_i = y_i^*]}{n} \quad (8)$$

where $\hat{y}_i$ is the label for pixel i predicted by the algorithm, y_i^* is the ground-truth label for pixel i , and $\mathbb{I}[\cdot]$ is the indicator function taking value one when its argument is true and zero otherwise. An alternative evaluation metric that better accounts for performance on rare categories is class-averaged accuracy, defined as

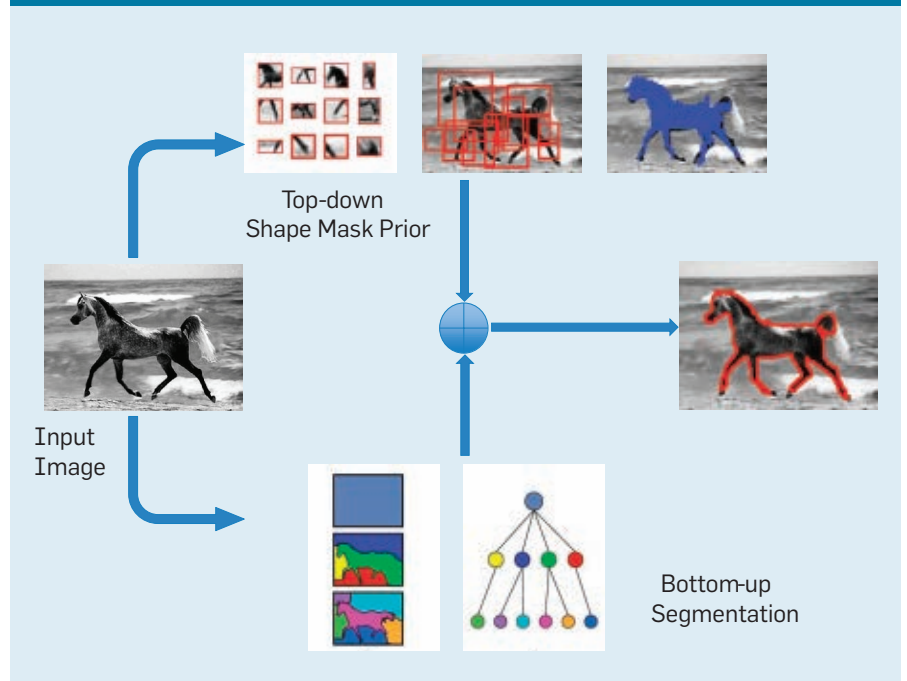
$$\text{acc}^{\text{class-avg}} = \frac{1}{|\mathcal{L}|} \sum_{\ell \in \mathcal{L}} \frac{\sum_{i=1}^n \mathbb{I}[\hat{y}_i = \ell] \wedge (y_i^* = \ell)}{\sum_{i=1}^n \mathbb{I}[y_i^* = \ell]} \quad (9)$$

The different accuracy measures defined by Equation 8 and Equation 9 are often referred to in statistics as “micro averaging” and “macro averaging,” respectively.

State-of-the-art performance on the semantic categories of the MSRC and Stanford Background datasets is approximately 86% and 77% pixelwise accuracy, respectively; class-averaged accuracies are typically 5%–10% less. On larger datasets, performance can be quite poor without top-down and contextual cues, especially on the less frequently occurring classes.

Illustrating the effects of different aspects of a scene-understanding model, Figure 7 includes results on an example image from the MSRC dataset. Classifying pixels independently (left results column) produces very noisy predictions, as shown. Adding a pairwise smoothness term helps remove the noise (right side). However, when the features are weak (top row), the algorithm cannot correctly classify the object in the image, though the background is easily identified using local features. Stronger features, including local and global cues, as discussed, coupled with the pairwise smoothness term, produce the correct

Figure 6. Combining bottom-up and top-down information in a CRF framework.²
The bottom-up process forms region hypotheses and maps them to semantic labels. The top-down process represents part-based shape constraints on object instances.



labeling result (bottom right).

These examples of semantic segmentation are indicative of more general trends in scene-understanding algorithms. More sophisticated features that incorporate contextual information (such as pixel location and global and shape-based features) perform much better than local appearance features, in general. Moreover, CRF models, with their pairwise smoothness priors, improve performance over independent pixel classification, but the benefit decreases as the sophistication of the features used by the independent classifiers increases. This trade-off is to be expected, as these features allow both the baseline performance to increase and the features to encode contextual information that can act as a surrogate for the smoothness assumption.

The qualitative results from the Darwin software framework (see Fig-

ure 8) also highlight a few points; as shown, the accuracy of the predictions is generally good, and the model is able to identify the boundary between object categories quite well. The labeling of foreground objects occasionally leaks into the background. This leakage is more prominent in the MSRC results and can be attributed to, in part, rough ground-truth labeling in that dataset. In models that use superpixels, these boundary errors can also be caused by inaccurate over-segmentations.

An interesting result is the labeling of the ducks in Figure 8 (MSRC, left column, third row down). Here, the water is classified correctly, but both ducks are labeled incorrectly. The white duck is mislabeled as water by the model due to both confusion of its local appearance with that of water and a strong smoothness assumption preferring to label it

with the same class as the surrounding background. The second duck is mislabeled as a boat. While this may seem absurd to humans, it is a reasonable mistake for an algorithm when considering the context is correct (boats also co-occur with water) and the model was trained on only a handful of images containing ducks.

Conclusion

We have explored scene understanding as a pixel-labeling task, including a number of technical challenges facing scene-understanding algorithms and a glimpse at current trends toward addressing them. Active research along these lines and the growing availability of high-quality datasets reflect current scene-understanding research; for example, better low-level feature representations are being learned automatically from large volumes of data rather than engineered by hand.²⁶ Researchers are also looking toward mid-level visual cues (also called “attributes”) to overcome some of the limitations of scarce training data; for example, knowing an object has feathers narrows the range of possible labels for that object.⁷

Moreover, improved learning algorithms based on structured-prediction models³¹ means large numbers of parameters can be tuned simultaneously. This results not only in more optimal parameters but enables use of richer models (such as those with parameterized higher-order terms). Other models being studied are hybrid models (such as grid-structured

Figure 7. Example semantic segmentation for an image from the MSRC dataset.

Shown are the original image (left) and color-coded pixel labels (right) from different scene-understanding models. The models vary by features (local appearance versus local and global appearance) and model complexity (independent pixel classification versus a CRF model with pairwise term); see Figure 8 for the related color legend.

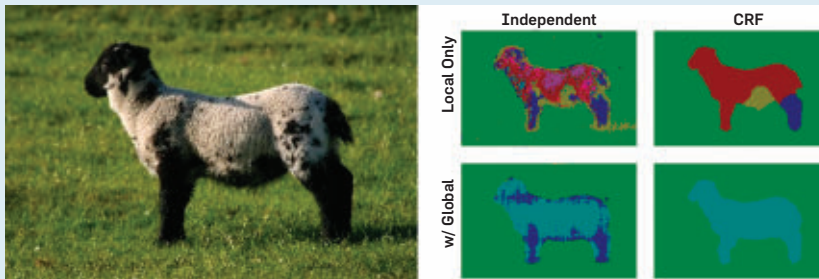
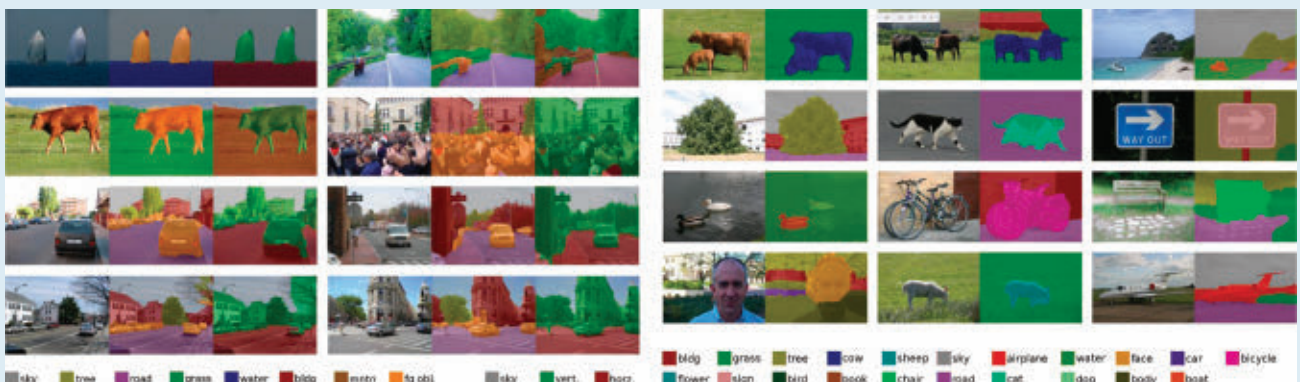


Figure 8. Representative results on two standard scene-understanding datasets produced by the Darwin software library; shown are the original image and predicted class labels overlay; best viewed in color.



CRFs) combined with a shape-constraining counterpart (such as restricted Boltzmann machines, or RBMs, and multi-scale patterns). Here, the CRFs capture appearance and smoothness, while the RBM and its variants encourages global consistency over the shape of objects. The models are still in their infancy but show great promise.

Another exciting direction is the use of non-parametric label-transfer techniques to allow for greater scalability in terms of the size of the dataset and the diversity of object categories.^{29,36} These techniques overcome (somewhat) the assumption of a closed-world set of labels implicit in the CRF formulation but introduce other complications (such as more-expensive test-time computation) and the need to resolve language ambiguities (such as whether “water” and “river” are semantically equivalent and, more important, refer to the same object).

One may ask whether it is necessary to label every pixel in an image. Indeed, for some scene-understanding tasks (such as face detection), a rough bounding box may suffice. However, for a detailed description of a scene, along the lines envisaged by early computer-vision researchers, pixel-level labeling seems inevitable. Advances on this front (such as those discussed here) and their integration into coherent holistic models are getting us closer to when a computer is indeed able to describe what it sees. ■

References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Sussstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 11 (Nov. 2012), 2274–2282.
- Borenstein, E., Sharon, E., and Ullman, S. Combining top-down and bottom-up segmentation. In *Proceedings of the IEEE Workshop on Perceptual Organization in Computer Vision at the IEEE Conference on Computer Vision and Pattern Recognition* (Washington, D.C., June 27–July 2). IEEE Computer Society Press, 2004, 46–46.
- Boykov, Y., Veksler, O., and Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 11 (Nov. 2001), 1222–1239.
- Comaniciu, D. and Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (May 2002), 603–619.
- Dalal, N. and Triggs, B. Histograms of oriented gradients for human detection. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (San Diego, CA, June 20–25). IEEE Computer Society Press, 2005, 886–893.
- Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., and Zisserman, A. The Pascal visual object classes challenge. *International Journal of Computer Vision* 88, 2 (June 2010), 303–338.
- Farhadi, A., Endres, I., Hoiem, D., and Forsyth, D. Describing objects by their attributes. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (Miami, FL, June 20–25). IEEE Computer Society Press, 2009, 1778–1785.
- Felzenszwalb, P.F. and Huttenlocher, D.P. Efficient graph-based image segmentation. *International Journal of Computer Vision* 59, 2 (Sept. 2004), 167–181.
- Fulkerson, B., Vedaldi, A., and Soatto, S. Class segmentation and object localization with superpixel neighborhoods. In *Proceedings of the 12th International Conference on Computer Vision* (Kyoto, Japan, Sept. 29–Oct. 2). IEEE Computer Society Press, 2009, 670–677.
- Gould, S., Fulton, R., and Koller, D. Decomposing a scene into geometric and semantically consistent regions. In *Proceedings of the 12th International Conference on Computer Vision* (Kyoto, Japan, Sept. 29–Oct. 2). IEEE Computer Society Press, 2009, 1–8.
- Gould, S., Gao, T., and Koller, D. Region-based segmentation and object detection. In *Advances in Neural Information Processing Systems 22* (Vancouver, B.C., Canada, Dec. 6–11). Curran Associates, Inc., 2009, 655–663.
- Gould, S., Rodgers, J., Cohen, D., Elidan, G., and Koller, D. Multi-class segmentation with relative location prior. *International Journal of Computer Vision* 80, 3 (Dec. 2008), 300–316.
- He, X., Zemel, R.S., and Carreira-Perpinan, M. Multiscale conditional random fields for image labeling. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (Washington, D.C., June 27–July 2). IEEE Computer Society Press, 2004, 695–702.
- Hedau, V., Hoiem, D., and Forsyth, D. Recovering the spatial layout of cluttered rooms. In *Proceedings of the International Conference on Computer Vision* (Kyoto, Japan, Sept. 29–Oct. 2). IEEE Computer Society Press, 2009, 1849–1856.
- Heitz, G., Gould, S., Saxena, A., and Koller, D. Cascaded classification models: Combining models for holistic scene understanding. In *Advances in Neural Information Processing Systems 21* (Vancouver, B.C., Canada, Dec. 8–13). Curran Associates, Inc., 2008, 641–648.
- Heitz, G. and Koller, D. Learning spatial context: Using stuff to find things. In *Proceedings of the European Conference on Computer Vision* (Marseille, France, Oct. 12–18). Springer, Berlin, Heidelberg, 2008, 30–43.
- Hoiem, D., Efros, A.A., and Hebert, M. Recovering surface layout from an image. *International Journal of Computer Vision* 75, 1 (Oct. 2007), 151–172.
- Hoiem, D., Efros, A.A., and Hebert, M. Closing the loop on scene interpretation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (Anchorage, AK, June 23–28). IEEE Computer Society Press, 2008, 1–8.
- Hoiem, D., Efros, A.A., and Hebert, M. Putting objects in perspective. *International Journal of Computer Vision* 80, 1 (Oct. 2008), 3–15.
- Kohli, P., Ladicky, L., and Torr, P.H. Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision* 82, 3 (May 2009), 302–324.
- Kolmogorov, V. and Zabih, R. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 2 (Feb. 2004), 147–159.
- Komodakis, N., Paragios, N., and Tziritas, G. MRF optimization via dual decomposition: Message-passing revisited. In *Proceedings of the International Conference on Computer Vision* (Rio de Janeiro, Oct. 14–21). IEEE Computer Society Press, 2007, 1–8.
- Krahenbuhl, P. and Koltun, V. Efficient inference in fully connected CRFs with gaussian edge potentials. In *Advances in Neural Information Processing Systems 24* (Granada, Spain, Dec. 12–17). Curran Associates, Inc., 2011, 109–117.
- Ladicky, L., Russell, C., Kohli, P., and Torr, P.H. Graph cut-based inference with co-occurrence statistics. In *Proceedings of the 11th European Conference on Computer Vision* (Crete, Greece, Sept. 5–11). Springer, Berlin, Heidelberg, 2010, 239–253.
- Lafferty, J.D., McCallum, A., and Pereira, F.C.N. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the International Conference on Machine Learning* (Williamstown, MA, June 28–July 1). Morgan Kaufmann, San Francisco, 2001, 282–289.
- Le, Q.V., Ranzato, M., Monga, R., Devin, M., Chen, K., Corrado, G.S., Dean, J., and Ng, A.Y. Building high-level features using large scale unsupervised learning. In *Proceedings of the International Conference on Machine Learning* (Edinburgh, Scotland, June 26–July 1). Morgan Kaufmann, San Francisco, 2012.
- Levin, A. and Weiss, Y. Learning to combine bottom-up and top-down segmentation. *International Journal of Computer Vision* 81, 1 (Sept. 2008), 105–118.
- Li, L.-J., Socher, R., and Fei-Fei, L. Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (Miami, FL, June 20–25). IEEE Computer Society Press, 2009, 2036–2043.
- Liu, C., Yuen, J., and Torralba, A. Nonparametric scene parsing via label transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (Dec. 2011), 2368–2382.
- Malik, J., Belongie, S., Shi, J., and Leung, T. Textons, contours and regions: Cue integration in image segmentation. In *Proceedings of the International Conference on Computer Vision* (Corfu, Greece, Sept. 20–25). IEEE Computer Society Press, 1999, 918–925.
- Nowozin, S. and Lampert, C.W. Structured learning and prediction in computer vision. *Foundations and Trends in Computer Graphics and Vision* 6, 3–4 (May 2011), 185–365.
- Rabinovich, A., Vedaldi, A., Galleguillos, C., Wiewiora, E., and Belongie, S. Objects in context. In *Proceedings of the International Conference on Computer Vision* (Rio de Janeiro, Oct. 14–21). IEEE Computer Society Press, 2007, 1–8.
- Ren, X. and Malik, J. Learning a classification model for segmentation. In *Proceedings of the International Conference on Computer Vision* (Nice, France, Oct. 13–16). IEEE Computer Society Press, 2003, 10–17.
- Shotton, J., Winn, J., Rother, C., and Criminisi, A. TextonBoost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In *Proceedings of the European Conference on Computer Vision* (Graz, Austria, May 7–13). Springer, Berlin, Heidelberg, 2006, 1–15.
- Szeliski, R. *Computer Vision: Algorithms and Applications*. Springer, Berlin, Heidelberg, 2011.
- Tighe, J. and Lazebnik, S. SuperParsing: Scalable nonparametric image parsing with superpixels. In *Proceedings of the European Conference on Computer Vision* (Crete, Greece, Sept. 5–11). Springer, Berlin, Heidelberg, 2010, 352–365.
- Tu, Z., Chen, X., Yuille, A.L., and Zhu, S.-C. Image parsing: Unifying segmentation, detection and recognition. *International Journal of Computer Vision* 63, 2 (July 2005), 113–140.
- Tu, Z. and Zhu, S.-C. Image segmentation by data-driven Markov chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (May 2002), 657–673.
- Wang, H., Gould, S., and Koller, D. Discriminative learning with latent variables for cluttered indoor scene understanding. In *Proceedings of the 11th European Conference on Computer Vision* (Crete, Greece, Sept. 5–Sept. 11). Springer, Berlin, Heidelberg, 2010, 497–510.
- Yao, Y., Fidler, S., and Urtasun, R. Describing the scene as a whole: Joint object detection, scene classification and semantic segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition* (Providence, RI, June 16–21). IEEE Computer Society Press, 2012, 702–709.

Stephen Gould (stephen.gould@anu.edu.au) is a fellow (senior lecturer) in the Research School of Computer Science in the College of Engineering and Computer Science at the Australian National University, Canberra, and visiting researcher in the Machine Learning Research Group at the National ICT Australia, Canberra.

Xuming He (xuming.he@nicta.com.au) is a senior researcher in the Computer Vision Research Group at the National ICT Australia, Canberra, and an adjunct fellow (senior lecturer) in the Research School of Engineering in the College of Engineering and Computer Science at the Australian National University, Canberra.

Strong ties connect individuals in the same community; weak ties connect individuals in different communities.

**BY PASQUALE DE MEO, EMILIO FERRARA,
GIACOMO FIUMARA, AND ALESSANDRO PROVETTI**

On Facebook, Most Ties Are Weak

PERVASIVE SOCIO-TECHNICAL NETWORKS bring new conceptual and technological challenges to developers and users alike. A central research theme is evaluation of the intensity of relations linking users and how they facilitate communication and the spread of information. These aspects of human relationships have been studied extensively in the social sciences under the framework of the “strength of weak ties” theory proposed by Mark Granovetter.¹³ Some research has considered whether that theory can be extended to online social networks like Facebook, suggesting interaction data can be used to predict the strength of ties. The approaches being used require handling user-generated data that is often not publicly available due to privacy concerns.

Here, we propose an alternative definition of weak and strong ties that requires knowledge of only the topology of the social network (such as who is a friend of whom on Facebook), relying on the fact that online

social networks, or OSNs, tend to fragment into communities. We thus suggest classifying as weak ties those edges linking individuals belonging to different communities and strong ties as those connecting users in the same community. We tested this definition on a large network representing part of the Facebook social graph and studied how weak and strong ties affect the information-diffusion process. Our findings suggest individuals in OSNs self-organize to create well-connected communities, while weak ties yield cohesion and optimize the coverage of information spread.

Analysis and understanding of OSNs like Facebook takes a theoretical foundation in social network analysis from Kleinberg.¹⁵ However, studying a real OSN poses computer science challenges, given the size, distribution, and organization (such as privacy and visibility rules) of data available to regular OSN subscribers.¹

In this context, analysis of large subsets of OSNs should generate a series of statistically robust measurements that are the basis of understanding OSN structure and evolution. Indeed, aggregate measures are very valuable in data management, privacy management, and online marketing. A challenging problem is how to evaluate the intensity of relations that bind users and determine how they facilitate the spread of information. These aspects of user connectedness and social communication have been studied in social sciences before, notably through Granovetter’s theory.¹³

» key insights

- Online social networks can facilitate the growth of a friendship network consisting of distant communities; severing weak ties has a strong effect on the connectivity of the overall network.
- We redefine the strong-versus-weak ties relationship to reflect the structure and evolution of massive online social networks, including Facebook.
- Interesting properties emerge when users self-organize into communities within online social networks.



Weak ties are connections between individuals belonging to distant areas of the social graph, or the ones that happen to have most of their relationships in different national, linguistic, age, or common-experience groups. Weak ties are a powerful way to transfer information across large social distances and to wide segments of the population. Strong ties are contacts between trusted/known persons (such as family ties and close friendships).

Are Granovetter's weak ties also found in OSNs like Facebook in the form he envisioned, or connections between individuals belonging to different areas of the social graph? Answering is difficult for at least two reasons: Facebook is organized mainly around the recording of just one type of relationship—friendship—implying Facebook friendship captures (and compresses) several degrees and nuances of human relationships dif-

ficult to separate and characterize through analysis of online data. And second, as Facebook grows in size and complexity, its friendship network grows denser, not sparser.¹

As OSNs become increasingly interconnected, testing Granovetter's theory poses scalability challenges. Early research, as in Gilbert and Karahalios,⁹ took a supervised approach, where a panel of Facebook users was asked to assess the strength of their own

friendship ties. Large-scale studies of Granovetter's theory, as in Gilbert and Karahalios,⁹ would arguably be difficult to conduct, given the size of today's OSNs. Other approaches, notably Bakshy et al.,² access Facebook data on user activities and compute tie strength as a function of type and frequency of user interactions. However, a cutoff threshold is required to distinguish strong ties from weak ties, and the tuning of that threshold has a crucial effect on identification of weak ties.

Here, we propose a new definition of weak ties rooted in analysis of large OSNs and in light of related computational challenges. We first identify communities within the network, then classify as weak ties those edges that connect users located in different communities; strong ties are those edges between users in the same community. Identifying weak ties through our definition is quick due to the efficiency of recent algorithms for finding communities in large networks⁸ and robust because no threshold has to be defined. We performed extensive experimental analysis on a public dataset on Facebook friendship collected and released by Gjoka et al.¹⁰ and a null-model comparison against randomly generated graphs. We deployed two well-known community-detection algorithms: the Louvain Method (LM)³ and Infomap.²⁰

Our analysis of the experimental results produced the following insights:

Independent of algorithm. The weak ties discovered through LM tend to coincide with those found through Infomap; our definition of weak ties is thus largely independent of the choice of community-detection algorithm;

Weak outnumber strong. Our community-defined weak ties outnumber strong ties;

More frequent. Weak ties occur more frequently in communities of small size; and

Spread of information. Weak ties identified through our approach play a crucial role in spreading information over a network, and their removal reduces the portion of the network that can be reached through information diffusion.

Weak and Strong Ties

Mark Granovetter introduced his now-classic definition:¹³ “The strength of a



We suggest classifying as weak ties those edges linking individuals belonging to different communities and strong ties as those connecting users in the same community.



tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services that characterize the tie.” It introduces important features of social ties like intensity of connection, or frequency of contacts, and reciprocity. Granovetter considered identification of strong and weak ties through topological information related to the social network structure. He thus introduced the concept of “bridge”: “A bridge is a line in a network which provides the only path between two points. Since, in general, each person has a great many contacts, a bridge between A and B provides the only route along which information or influence can flow from any contact of A to any contact of B.” Note, all bridges are weak ties. Unfortunately, with OSNs, the definition of bridge is restrictive and unsuitable. The well-known “small world effect,” or presence of short paths connecting a pair of vertices, and the “scale-free degree distribution,” or presence of “hubs” maintaining the whole network, make it unlikely that an edge could be found whose deletion would completely disconnect its terminal vertices.

To adapt Granovetter's definition to OSNs, we could redefine a “shortcut bridge” as a link connecting a pair of vertices whose deletion would cause an increase in the distance (defined as the length of the shortest path linking two vertices) between them. However, this otherwise sensible definition is more controversial than it might seem. First, it depends on the notion of shortest path, but, unfortunately, identification of all-pairs shortest paths is computationally unfeasible, even on networks of modest size. Second, even the alternative definition of distance, or “weighted path” (computed as the sum of the weights assigned to each connection being used), would remain computationally prohibitive for OSNs. Our goal here is to explore a novel and computationally feasible definition of weak ties that suits analysis of large-scale OSNs. Instead of a relational definition based on the intensity of the interactions between two users, we propose a community-based definition. We define weak ties

as those edges that, after dividing up the network into communities (thus obtaining so-called “community structure”), connect vertices belonging to different communities. Finally, we classify intra-community edges as strong ties.

One of the most important features of our weak ties is that those that are bridges (in the Granovetter sense) create more, and shorter, paths; their deletion would indeed be more disruptive than the removal of a strong tie, from a community-structure perspective. This may even be the reason weak ties (albeit defined in a slightly different fashion) have proved effective in the diffusion of information.^{5,21}

Benefits of our definition. Our definition of weak tie has four appealing features:

Weaker than Granovetter. The fact that vertices linked by a weak tie belong to different communities does not imply the edge between them is a bridge; its deletion may not disconnect its vertices, and, in practice, almost never does;

Based on topological information. It enables weak/strong classification on the basis of only topological information. Zhao et al.²¹ used topological information but only locally, on the neighbors of the two terminal vertices, whereas our definition handles global information, as it relies on the partitioning of the whole network;

Binary. It labels each edge in the network as either weak or strong; as a consequence, there is no need to need to “tune” any threshold, below which edges are classified as weak, the upshot being we cannot compare two edges on the basis of their strength; and

Accuracy of community discovery. Our experiments show our definition of weak tie is robust with respect to the choice of a particular community-detection algorithm.⁸

Related Approaches

Several research projects have examined the strength of weak ties in terms of nontopological information. A nontopological approach to determining degrees of strength or weakness requires researchers to choose some measurable variables by which the strength of the relation binding two users can be deduced. In such a

scenario, weak ties are intended in a relational sense; that is, they connect acquaintances who do not interact frequently and therefore do not influence each another strongly. For “offline” social networks, Marsden and Campbell¹⁷ identified some variables (such as communication reciprocity and the presence of at least one mutual friend) as indicators of a weak tie.

In the context of the social Web, Gilbert and Karahalios⁹ and Petroczi et al.¹⁹ studied small OSNs—35 participants and 56 participants, respectively—assigning weights to the relationships on the basis of several measures of strength (such as intimacy/closeness, reciprocity, and sociability/conviviality) assessed through direct questions to participants. Gilbert and Karahalios⁹ extended Marsden-Campbell’s method to Facebook by identifying as many as 74 variables as potential predictors of strength. They then modeled strength as a linear combination of the variables, computing the weights by means of a variant of ordinary least-square regression. To validate their model, Gilbert and Karahalios⁹ recruited 35 users and asked them to rate their Facebook friends. They achieved accuracy of approximately 85%, but their performance is difficult to replicate on large OSN fragments. Their method requires collecting a large amount of information on user behavior; due to privacy concerns and the limitation on use of proprietary data, most such required data is unlikely to be available for academic studies.

Zhao et al.²¹ formalized the strength of ties with a weight w_{ij} assigned to the edge connecting vertices i and j as follows

$$w_{ij} = \frac{c_{ij}}{k_i + k_j - 2 - c_{ij}}$$

where k_i and k_j are the degrees of i and j and c_{ij} are the numbers of their mutual acquaintances. Numerical simulations by Zhao et al.²¹ showed that by gradually deleting ties with lower w values, information coverage drops sharply.

Zhao et al.’s results corroborate the hypothesis that weak ties are in fact the key to information diffusion. A 2012 paper on weak ties in Facebook due to Bakshey et al.² considered four parameters to measure the strength of a

connection: frequency of private messages; frequency of public comments left on each other’s posts; number of times they jointly appear in a photo; and number of times they jointly commented on a third-party post.

Unlike us, Bakshy et al.² did not consider, at least not explicitly, the topological aspects of Facebook social connections; moreover, their analysis requires access to proprietary data and records of user activity, so the methodologies are not easily compared. Our work is more like Blondel et al.² and Zhao²¹ in that weak ties are understood to be useful connectors favoring the spread of information and was (as in Blondel et al.²) tested on Facebook.

However, one difference emerges: Blondel et al. and Zhao et al. both assigned scores to ties and classified them according to a threshold value. In contrast, we classify ties as weak or strong, depending on whether or not they connect vertices located in different communities; our classification scheme is binary and does not use scores and threshold, which may be difficult to set up and tune properly. Another approach is due to Grabowicz et al.,¹² who, like us, used information about network topology to identify weak ties. However, they focused on Twitter, which can be modeled as a directed network where user relationships are mostly asymmetric. The Twitter “retweet” and “mention” functions have a major effect on identification of weak ties; the studies are thus not easily compared.

Results

The results of the tests we carried out highlight the pros and cons of our definition of a weak tie. For the preliminary community-detection phase we deployed two popular algorithms: LM³ and Infomap.²⁰ We considered two testbeds, comparing them with a fragment of the Facebook network collected by Gjoka et al.¹⁰ with 957,000 users and 58.4 million friendship connections^a and “null-model” networks consisting of Erdős-Rényi random graphs with up to 2,048 vertices; we varied the

a For a complete description of the dataset, see supplementary material at <http://informatica.unime.it/weak-ties/>.

Figure 1. The complementary cumulative distribution function associated with (a) weak ties and (b) strong ties in our Facebook dataset.

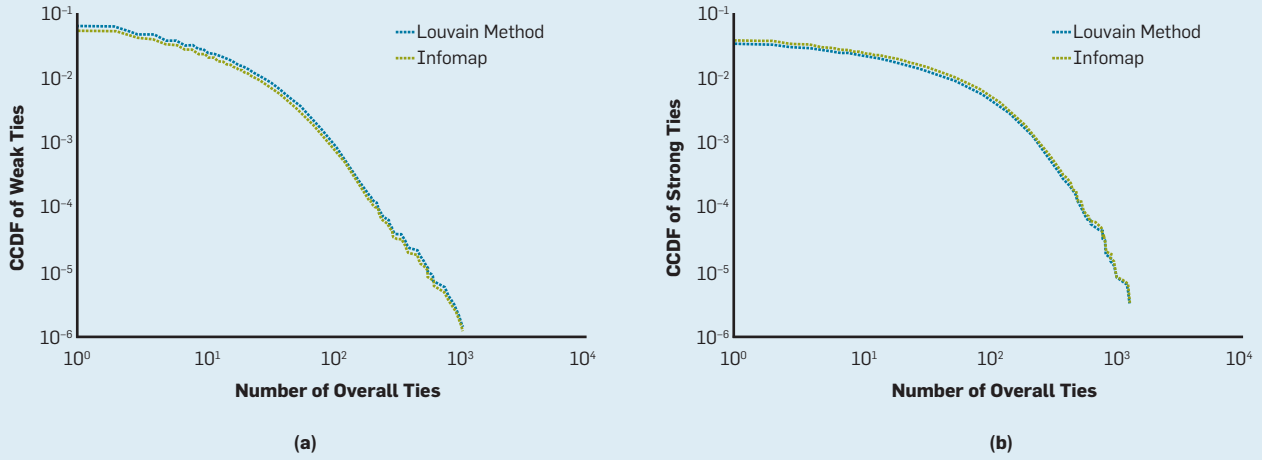


Figure 2. The R_{avg} ratio for Erdős-Rényi random graphs clustered with (a) LM and (b) Infomap; instances with $|V| = 128, 512, 1,024$, and $4,096$.

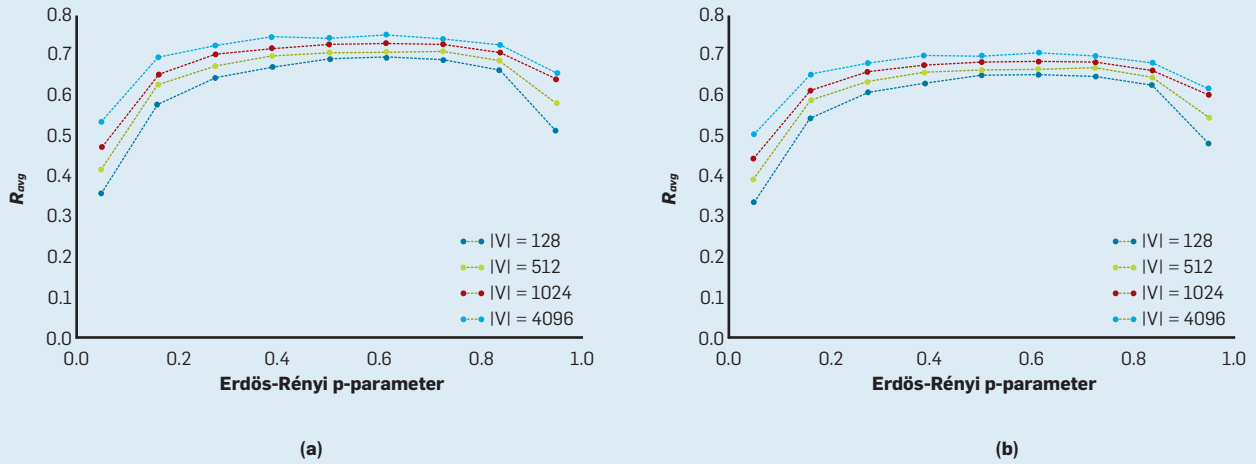
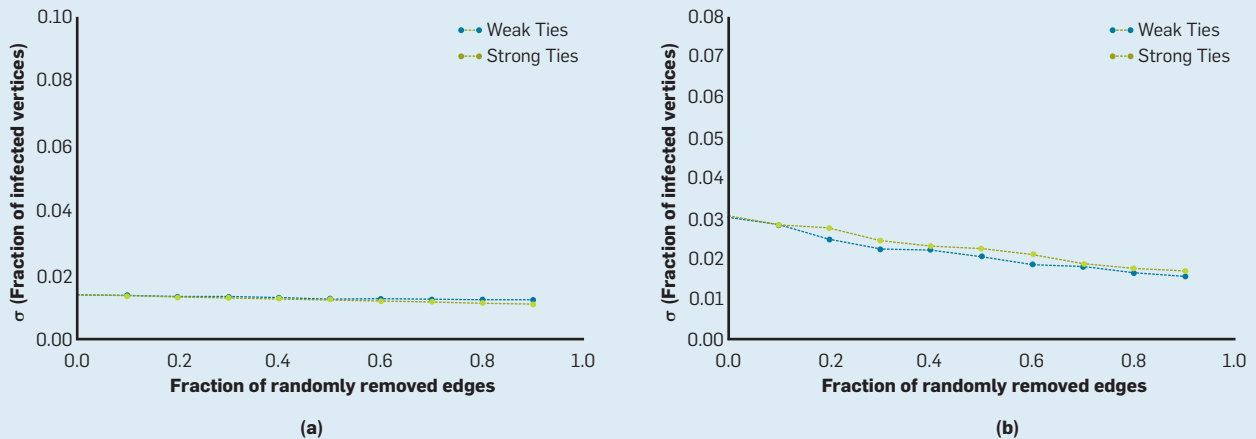


Figure 3. Coverage in graphs with different values of τ : (a) with Erdős-Rényi's random graphs ($|V| = 512$, $p_{link} = 0.05$, $p_{inf} = 0.01$) when τ ranges from 0 to 0.9; (b) with the Facebook dataset generated with τ ranging from 0 to 0.9. Charts (a) and (b) report results for the LM.



probability of having an edge between an arbitrary pair of vertices uniformly from 0.05 to 0.95, with a 0.05 step.

Robustness of definition. As a preliminary experiment, we studied the robustness of our definition of weak ties with respect to the method one might adopt for finding communities. Different community-detection methods might produce different results, albeit slightly, hence our weak/strong ties classification could become dependent on the community-detection method. We ran both LM and Infomap on our Facebook sample, comparing the community structures found by the two algorithms by applying the normalized mutual information (NMI). We found an NMI of approximately 0.9, informing us that the level of disagreement between the two algorithms is quite low.^b

We conclude that on Facebook our definition of weak ties is robust because it does not depend on a specific community-detection algorithm. However, note that the communities discovered by LM and Infomap might differ in subtle ways yet still be characterized by a high NMI. We next offer a detailed, graphical comparison of the results we achieved by applying LM and Infomap to better understand the practical consequences of adopting a particular community-detection algorithm.

Strong and weak ties on Facebook. Here, we consider the distribution of weak and strong ties in Facebook. We computed the “complementary cumulative distribution function,” or CCDF, of weak (as opposed to strong) ties, or the probability of finding a vertex with more than k weak (as opposed to strong) ties (see Figure 1).

The CCDF associated with strong ties decreases quicker than that for weak ones. At $k = 4$, we see a tipping point after which weak ties quickly outnumber strong ties; that is, the latter are much less frequent than the former in higher-degree vertices.

The results of our experiments agree fairly well with Granovetter’s “vi-



Our findings suggest individuals in OSNs self-organize to create well-connected communities, while weak ties yield cohesion and optimize the coverage of information spread.



sion” of weak ties. Sociological theories (such as triadic closure,¹³ cognitive balance,¹⁴ and homophily¹⁸) suggest individuals tend to aggregate in small communities.

We thus confirm that the intensity of human relations is very tight within small groups but decreases for individuals belonging to distant communities. Most connections are weak in the sense of Granovetter, with few contacts and infrequent interactions.

Weak and strong ties in random graphs. Here, we offer a comparative analysis of weak/strong ties distribution in Erdős-Rényi random graphs; that is, we check whether the results we obtained on Facebook data still hold on graphs where the structure is known in advance and lacks, by construction, a clear community structure. We computed the ratio R_{avg} of the number of weak ties to the total number of ties; Figure 2 outlines R_{avg} for different values of $|V|$ and when the probability p_{link} of having an edge between an arbitrary pair of vertices varies uniformly from 0.05 to 0.95.

We thus derive several insights:

Weak outnumber strong. R_{avg} is always greater than 0.6; that is, weak ties still outnumber strong ties, even in randomly generated graphs;

Stable and independent. R_{avg} is relatively stable and independent of p_{link} ; that is, the “sparseness” of G has a limited effect on the number of weak ties; and

Limited effect. $|V|$ has a limited effect on R_{avg} ; when $|V|$ goes from 128 to 4,096, or increases by a factor of 32, R_{avg} increases by only 17.14%.

Weak ties in information diffusion. We also studied how weak ties influence the information-diffusion process, clarifying the connection between our definition and Granovetter’s, since we view weak ties as providing specific links between individuals who would otherwise remain disconnected, as they belong to distant areas of the social graph.

Granovetter’s weak ties should play a central role in the spread of information. But how do our weak ties convey information?

We applied the Independent Cascade Model (ICM)¹¹ to simulate information propagation over a network, again taking the Facebook dataset

^b The definition of NMI is included in the supplementary material (see footnote a); if NMI, which ranges between 0 and 1, approaches 1, then the communities found by the two algorithms can be viewed as coinciding.

and Erdős-Rényi's random graphs for comparison. In ICM, a vertex v_0 is selected uniformly at random to forward a message to its neighbors, with probability equal to p_{inf} (infection probability).

Each "infected" vertex can, in turn, recursively propagate the message to its neighbors. Leskovec et al.¹⁶ reported reasonable values of p_{inf} are 0.01, 0.02, and 0.03.

To generate statistically significant results, we selected the vertex v_0 to start from 250 times; at each selection of v_0 , we simulated the propagation of a message, measuring the coverage σ defined as the ratio of the number of vertices receiving a message (infected vertices) to the total number of vertices. We repeated the experiment by progressively (and randomly) deleting a fraction τ of weak ties.

In our simulation, τ ranged from 0.1 to 0.9, and, for each value, we computed the corresponding coverage. The whole procedure was repeated separately for weak and strong ties. Figure 3 and in the supplementary material (<http://informatica.unime.it/weak-ties/>) show the values of σ obtained by applying the LM and Infomap on Erdős-Rényi's random graphs, respectively. Here, we consider a graph with $|V| = 512$ vertices, $p_{\text{link}} = 0.05$ and $p_{\text{link}} = 0.03$.

Note the gradual deletion of weak ties yields a decrease of σ ; on average, the greatest decrease of σ is 11.98%, with an average decrease equal to 5.71%, with standard deviation of 3.4%. We found the same behavior when removing strong ties, but the decrease was less marked. We observed that if $|V|$ increases, the coverage σ also increases. The values of σ associated with our Facebook dataset are reported in Figure 3b and in the supplementary material.

Again, for a fixed value of τ , removing weak ties yields a decrease of σ more marked than the removal of strong ties. The greatest decrease of σ we observed was approximately 14.31%, with an average decrease of 6.26%, with standard deviation of 3.1%.

These values are always greater than those for Erdős-Rényi's random graphs, proving the removal of weak ties is more significant in real OSNs than in networks not equipped with a meaningful community structure. We

thus conclude that our definition of weak ties captures Granovetter's idea, in that deleting weak ties decreases/obstructs the flow of information much more effectively than removing strong ties.

Conclusion

We have presented a novel definition of weak ties designed for OSNs like Facebook based on the community structure of the network itself. The experiments we conducted on a Facebook sample of 957,000 users and randomly generated graphs highlight the role and importance of weak ties. We characterized the overall statistical distribution of weak ties as a function of the size of a community and its density. We studied their role in information-diffusion processes, with results suggesting a connection between our definition of weak ties for OSNs and Mark Granovetter's original definition.

Even though several recent research projects have focused on Facebook's social graph,^{1,4} community structure,⁷ and weak ties per se,² our community-based definition of weak ties better fits Facebook and similarly large (and dense) OSNs.

As for future work, we hope to follow up with two more projects: The first concerns investigating the applicability of network-weighting strategies, so the strength of ties can be computed according to a given rationale (such as the ability of each link to spread information); we intend to adopt a novel method of weighting edges suited for OSNs we devised in De Meo et al.⁶ The second concerns analysis of geographical data related to Facebook users; due to the merging of different graphs (such as social and geographical), we expect to gain additional insight into the role of physical vs. virtual distances.

Acknowledgments

We thank Minas Gjoka for making his Facebook dataset available to us and to the anonymous reviewers for their constructive comments. C

References

1. Backstrom, L., Boldi, B., Rosa, M., Ugander, J., and Vigna, S. Four degrees of separation. In *Proceedings of the ACM Web Science Conference* (Evanston, IL, June). ACM Press, New York, 2012, 33–42.
2. Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. The role of social networks in information diffusion. In

- Proceedings of the World Wide Web Conference* (Lyon, France). ACM Press, New York, 2012, 519–528.
3. Blondel, V.D., Guillaume, J.L., Lambiotte, R., and Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* (Oct. 2008).
4. Catanese, S., De Meo, P., Ferrara, E., Fiumara, G., and Provetti, A. Crawling Facebook for social network analysis purposes. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics* (Sogndal, Norway, May 25–27). ACM Press, New York, 2011, 52:1–52:8.
5. Centola, D. The spread of behavior in an online social network experiment. *Science* 329 (Sept. 2010), 1194–1197.
6. De Meo, P., Ferrara, E., Fiumara, G., and Ricciardello, A. A novel measure of edge centrality in social networks. *Knowledge-Based Systems* 30 (June 2012), 136–150.
7. Ferrara, E. A Large-scale community structure analysis in Facebook. *EPJ Data Science* 1, 1 (Nov. 2012), 1–30.
8. Fortunato, S. Community detection in graphs. *Physics Reports* 486 (Feb. 2010), 75–174.
9. Gilbert, E. and Karahalios, K. Predicting tie strength with social media. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (Boston, Apr.). ACM Press, New York, 2009, 211–220.
10. Gjoka, M., Kurant, M., Butts, C.T., and Markopoulou, A. Practical recommendations on crawling online social networks. *IEEE Journal on Selected Areas in Communications* 29, 9 (Oct. 2011), 1872–1892.
11. Goldenberg, J., Libai, B., and Muller, E. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters* 12, 3 (2001), 211–223.
12. Grabowicz, P.A., Ramasco, J.J., Moro, E., Pujol, J.M., and Eguluz, V.M. Social features of online networks: The strength of intermediary ties in online social media. *PLoS ONE* 7, 1 (Jan. 2012).
13. Granovetter, M.S. The strength of weak ties. *American Journal of Sociology* 78, 6 (May 1973), 1360–1380.
14. Heider, F. *The Psychology of Interpersonal Relations*. Lawrence Erlbaum, New York, 1982.
15. Kleinberg, J. The convergence of social and technological networks. *Commun. ACM* 51, 11 (Nov. 2008), 66–72.
16. Leskovec, J., Adamic, L., and Huberman, B. The dynamics of viral marketing. *ACM Transactions on the Web* 1, 1 (May 2007).
17. Marsden, P.V. and Campbell, K.E. Measuring tie strength. *Social Forces* 63, 2 (Dec. 1984), 482–501.
18. McPherson, M., Smith-Lovin, L., and Cook, J.M. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27, 1 (Aug. 2001), 415–444.
19. Petroczi, A., Nepusz, T., and Bacs, F. Measuring tie-strength in virtual social networks. *Connections* 27, 2 (2006), 39–52.
20. Rosvall, M. and Bergstrom, C.T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* 105, 4 (Jan. 2008), 1118–1123.
21. Zhao, J., Wu, J., and Xu, K. Weak ties: Subtle role of information diffusion in online social networks. *Physical Review E* 82, 1 (July 2010).

Pasquale De Meo (pdmeo@unime.it) is an assistant professor in the Department of Ancient and Modern Civilizations at the University of Messina, Italy.

Emilio Ferrara (ferrarae@indiana.edu) is a postdoctoral researcher in the School of Informatics and Computing at Indiana University, Bloomington, IN.

Giacomo Fiumara (gfiumara@unime.it) is an assistant professor in the Department of Mathematics and Informatics at the University of Messina, Italy.

Alessandro Provetti (ale@unime.it) is an associate professor in the Department of Mathematics and Informatics at the University of Messina, Italy.

Computing Reviews

Connect with our Community of Reviewers

“There aren’t too many walk-in technical bookshops where I live; checking out reviewer comments on new books provides me with an equivalent experience.”

- Graham Jenkins



Association for
Computing Machinery

ThinkLoud

www.computingreviews.com

It's not just about being fair.

BY BERYL NELSON

The Data on Diversity

PEOPLE WORKING TOGETHER can achieve more than they can alone; this is a fundamental principle upon which organizations are founded. Social scientists have shown that teams and organizations whose members are heterogeneous in meaningful ways, for example, in skill set, education, work experiences, perspectives on a problem, cultural orientation, and so forth, have a higher potential for innovation than teams whose members are homogeneous. These findings are not without controversy, yet the implications for the computing industry are profound, given the relative homogeneity of the field along a few important dimensions. Take, for example, the composition of degrees awarded in computer science, computer engineering, and informatics in 2012 at research institutions in the U.S.

► 13.3% of BS degrees, 28.7% of MS degrees, and 19.2% of Ph.D.'s were awarded to female candidates, down from a high of 37% of BS degrees in computer science in 1986.

► 5.3% of BS degrees were awarded to African American candidates, as were 2.7% of MS degrees, and 2% of Ph.D.'s.

Among computing professionals, about 20% of CS faculty in U.S. universities are women, and 1.6% are African American.⁵⁰ Similar numbers exist in industry.

Diversity, bias, and stereotypes have traditionally been discussed in very relativistic terms: surveys of whether people thought there was bias, and so on. In recent years, imaginative researchers have developed ways to gather quantitative data about the benefits of, as well as the challenges to, having a diverse workforce. This article explores the benefits that diversity can bring to teams, and the cognitive factors—namely, stereotypes based on social group membership—that keep us from achieving optimal levels of diversity.

Benefits of Diversity

Diverse teams are more effective: they produce better financial results and better results in innovation. These results show that having a diverse organization is a business imperative.

» key insights

- Teams and organizations whose members are heterogeneous in meaningful ways have a higher potential for innovation than teams whose members are homogeneous.
- Social science experiments using quantitative methods show bias, stereotype threat, and methods to combat them.
- Effectiveness of diverse teams depends on trusting and supportive cultures. Data publication is one of the most important tools to identify and combat identity threat and biased decision making.
- There is hope! There are tools that have been shown to combat bias and identity threat effectively.



Financial results. Organizations that include a high percentage of women in senior positions show better financial results.

Companies in the top quartile for women in the executive committee from 2007–2009 had 41% greater return on equity, and 56% greater earnings before interest and taxes than companies with no women in the executive committee, for companies within the same industrial sector (see

Figure 1).¹⁵

Financial results for companies with at least three women serving on the board of directors are better: in 2007, return on equity was 16.7%, as opposed to an average 11.5%; return on sales was 16.8%, as opposed to an average 11.5%; return on invested capital was 10%, as opposed to an average 6.2% (see Figure 2).²⁶

This is also true in other geographies. Profits of Indian companies

headed by women grew 56% within five years, but grew even faster, at the rate of 64%, within three years.³³ The BSE-30 companies posted a growth rate of 27% and 23%, respectively, during the same period.

Similar results have been found for race: organizations with greater racial diversity were associated with greater sales revenue, a larger number of customers, greater market share, and greater profits. In the same study,

greater gender diversity was also found to be associated with better results in sales revenue, number of customers, and profitability.²³

These studies are correlational:

causation can be inferred, but not proven. Researchers have posited that diversity may actually be one causal determinant of firm performance due to the increased innova-

tion that occurs when organizations are diverse. The next section reviews this evidence.

Innovation. Diversity has been shown to create a cognitive and social environment that is a positive indicator for innovation and a negative indicator for routine tasks. These dynamics may have real-world consequences for scholarship in our field.

In a study of collective intelligence and creativity, researchers gave subjects aged 18 to 60 standard intelligence tests and assigned them randomly to teams. Each team of 3–5 people was asked to complete several tasks, including brainstorming, decision making, and visual puzzles, and to solve one complex problem that was too difficult for one brilliant individual to solve: that is, a team was required. Teams were given intelligence scores based on their performance.

The only predictor of team collective intelligence was whether there were women on the team. Note that all the high-scoring groups are close to 50% women; some of the low-scoring teams are also near 50%, but the groups with little gender mix did not score as highly. This was a surprise result to the researchers. With more investigation, it was found the difference was having the social skills that made it possible to use the contributions of all the team members, and these correlate more with women than with men. Figure 3 illustrates the relationship of team composition to success.^{48,49}

Patent authorship is another area that shows a benefit of diversity. More than 90% of all computer technology patents issued in the U.S. since 1980 have been granted to men only. Yet mixed gender patents are cited 26% to 42% more than any single gender patent.⁴ An update to this report in 2012 showed that mixed gender patents typically have a large number of authors. The higher citation rate (30% to 40% more in the 2012 update) is associated with higher numbers of authors. The reasons for this are not well understood.

These results show a correlation between diverse organizational composition, financial success, and innovation. While there is not a clear causal relationship shown between diversity and success, the results have been

Figure 1. Women at the top of corporations: Making it happen. Courtesy of McKinsey Report.¹⁵

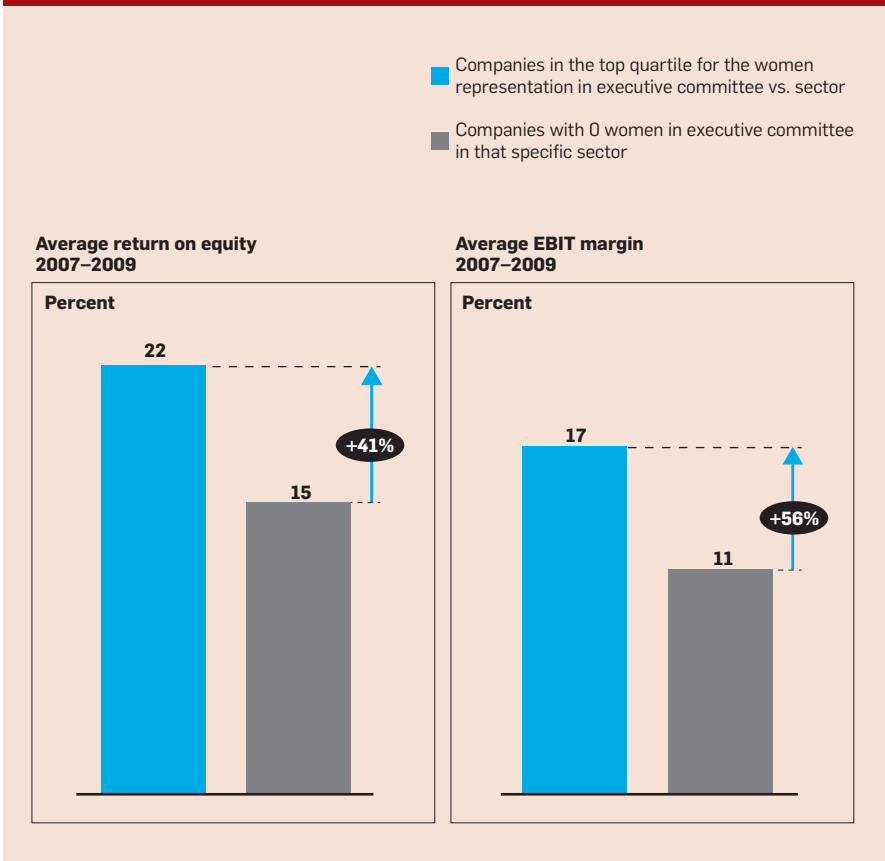
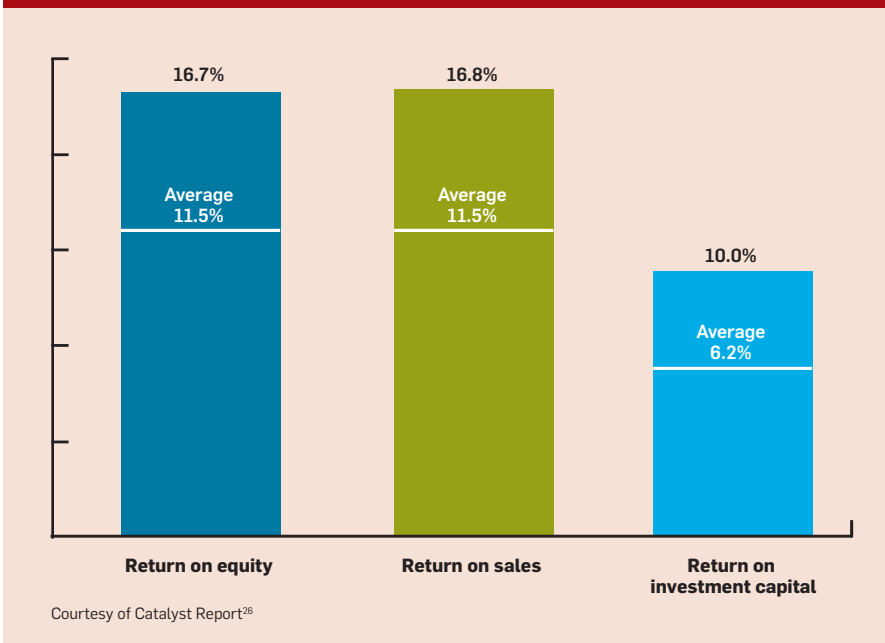


Figure 2. Financial performance at companies with three or more women board directors (WBD).



shown with varying methodologies and in varying geographies, to a degree that demands attention.

Challenges Faced by Diverse Teams

Unfortunately, it is not easy to make diverse teams effective.^a There are a number of forces that work against the desired effect: having the entire team productive. There can be potential negative effects of any of the following:

- unconscious bias,
- stereotype threat,
- exclusion from critical social networks,
- lack of role models, and
- unaware managers.

The following sections address primarily unconscious bias and stereotype threat.

Unconscious bias. One of the factors that both limits the diversity present in a team or organization, and that inhibits the potential success of diverse teams, is the unconscious bias, that is, stereotypes, that we hold toward people based on the social groups to which they belong. Stereotypes are, simply, a constellation of traits, characteristics, skills, and values that we ascribe to members of social groups, such as gender, race, age, religion, nationality, and others. These are learned through cultural messages and stories, comments from family and friends, portrayals in the media, and so forth. These stereotypes, despite our best intentions, can bias our impressions of, and affect our actions toward, others in our environment. The stereotypes especially relevant in work situations include those characteristics that are visible, such as sex, race, weight, and age; but also those not visible but relatively easy to discern, such as educational background and nationality.

Project Implicit at Harvard hosts an online test of implicit associa-

tions: the user's implicit association between two concepts is measured via user response time. There are reports for many associations, and the results are stunning.⁵

► Almost everyone has measurable biases (for example, 70%–80% have biases against women in technology, or preferring white to African American, or preferring young people).

► Almost no one reports such biases (for example, 15% report a preference for white people).

► Even the people who are the subject of a bias may have that bias. For example, I tested as moderately biased against women in science and technology, and this is totally against my self-interest.

The book *Blindspot*⁵ explains the development of the implicit association test and its results, exploring the reasons for differences between unconscious perception and our assumptions. For example, it has been found that age is one of the strongest biases tested; this is true even among the elderly, and in societies in Asia that traditionally have valued the wisdom

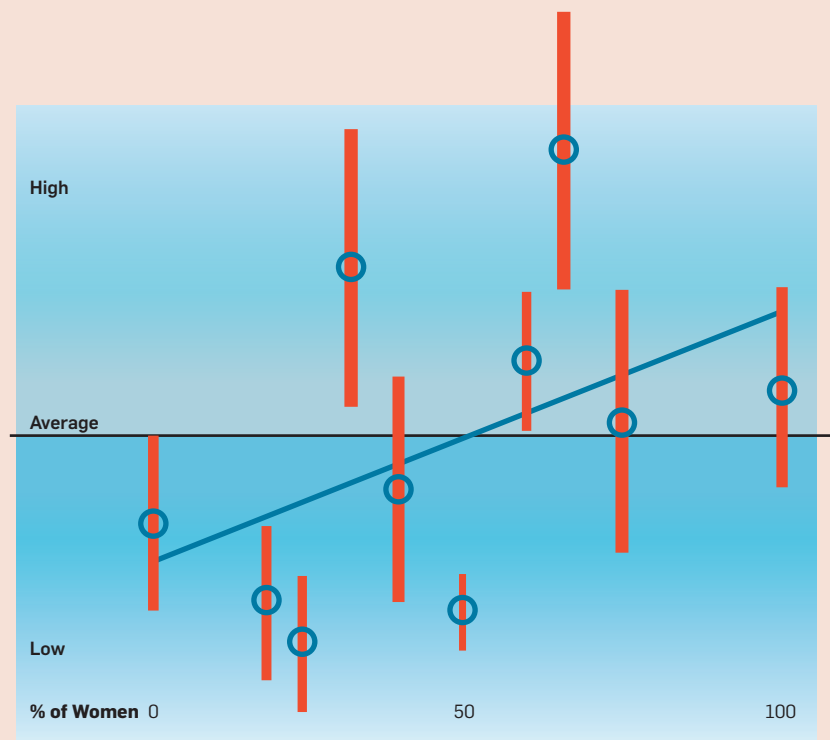
of age and experience.⁵ *Blindspot* also considers the evidence for any predictive link between measured biases and outcomes or behaviors; results are not yet conclusive in this area.

In individual cases, bias can be very difficult to verify. In aggregate, however, bias and other forms of unconscious decision-making are readily apparent:

► About 58% of CEOs of Fortune 500 companies are taller than 6ft. (about 183cm), and almost a third are taller than 6ft. 2in. (about 188cm). In the population in general, about 14.5% are taller than 6ft., and 3.9% taller than 6ft. 2in.¹⁷ Most people would not believe height predicts competence, and yet, these choices are frequently made. In fact, height is strongly correlated with career success.

► In a laboratory situation, applicants were seen in the waiting room either alone, or sitting next to another applicant. Applicants who were seen sitting next to an overweight applicant were less likely to be hired than an applicant sitting alone or next to an average weight applicant,

Figure 3. Collective Intelligence vs. Team Composition.



The horizontal axis indicates team composition. The blue circles represent averages for each percentage level; the red bars indicate standard deviation. Courtesy of Woolley and Malone.⁴⁹

a Many studies referenced here refer to women, but these results should largely be considered to apply across all axes of diversity; for example, gender, cultural and national origin, sexual orientation, age, educational background, religion, and other life experiences. It is more difficult to study differences that are not externally visible, such as differences in economic class, than visible differences like gender or race, but these less-visible differences are also important to consider in conversations about diversity.

and regardless of their own weight.⁴⁶

► In an examination of data on wages of a given population over the course of five years, it was observed that there is a correlation between weight and income: for women, this is a negative correlation; for men, a positive correlation, excluding obesity.²⁷

► A disturbing finding is that African Americans and other minorities receive less effective healthcare interventions and outcomes, even when controlling for socioeconomic class and insurance coverage.⁵

► At a coffee station at a company in Northern England, the sign for the amount to pay for milk (to put in tea or coffee, on an honor system) had a photo of flowers or eyes on alternate weeks. The amount of money received was about double in the weeks with eyes over the weeks with flowers. And if the eyes were looking directly out, the amount received was much higher; see Figure 4. Although their behavior was ap-

parently heavily influenced by the photos, at the end of the trial, no one remembered the photos changing.⁷ Clearly, there was unconscious decision making in each of these cases, but the data alone does not give tools to overcome bias.

Science faculty at respected U.S. research institutions show measurable bias toward male students: the faculty were given applications for lab manager that varied only in the gender of the name. Male students were rated significantly more competent, and were given higher average salaries and more mentoring opportunities than women students. Importantly, even the female faculty showed this bias.³² Male faculty offered \$30,520.83 to male students on average, and \$27,111.11 to female students. Female faculty offered \$29,333.33 to male students on average, and \$25,000.00 to female students.

Shankar Vedantam explains the bias illustrated in these examples by

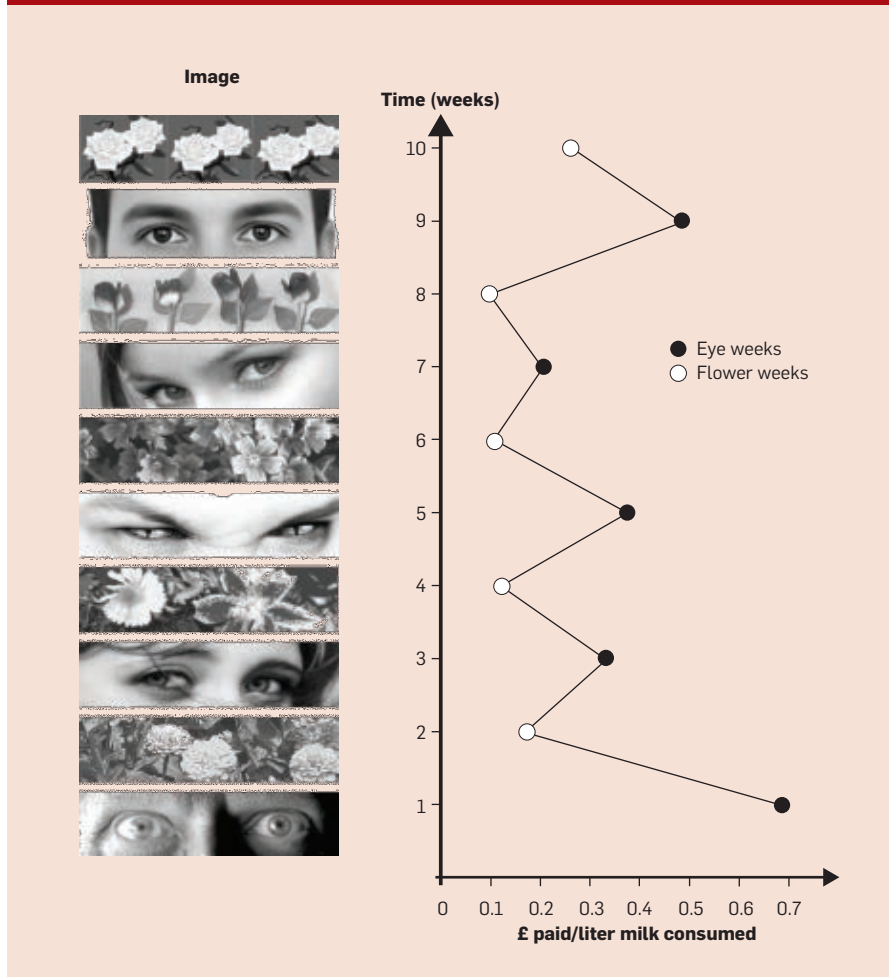
saying that we create a set of mental shortcuts, or heuristics, for the situations we experience every day, and that we actively work against those biases for the most part, to the extent we think the bias is not “good.” For example, you see very small children, as well as people with dementia, expressing strong bias; or politicians may make a slip of the tongue when they are tired.⁴⁶

Until this point, I have been describing the stereotyping process as a negative force for individual and team functioning. However, this process actually stems from an adaptive, often functional psychological process. Mental heuristics and cognitive shortcuts enable us to process information without conscious deliberation: they fill in the informational gaps we often experience when making decisions. In other words, habits of mind help us to save brain power for more difficult tasks. Joseph Pieper’s classic *Leisure, the Basic of Culture*³⁵ sets out this theory well. In the 1980s, people thought we could create expert systems by interviewing experts like brain surgeons or oil exploration specialists, and creating a rule chaining prolog environment that would recreate their decision-making ability. The problem was the experts did not know how they knew what they knew. That is, experts are creating associations between disparate experiences and pieces of knowledge, using the subconscious brain. In software development, we see this in our most skillful engineers: they can debug a complex problem, but they will probably not be able to specify a set of rules that would work as well as consulting the expert.

What it does say, however, is that we should have better knowledge of our own biases and unconscious decision-making.

Stereotype threat. In *Whistling Vivaldi*,⁴⁷ Claude Steele published a history of research about stereotype effects and identity threat. The original problem he tried to solve was, why were college entrance scores not predictive of college success for African Americans? Steele asserts we each have multiple identities (such as I am a woman; I am an MIT alumna; I am American); and that we respond

Figure 4. Coffee station pictures and money collected. Courtesy of Bateson et al.⁷



to the identity under threat. He says, “Being threatened because we have a given characteristic is what makes us most aware of being a particular type of person.” In Steele’s terminology, each of these identities can have what he calls “identity contingencies;” each of these contingencies can be positive, neutral, or negative in a given social situation.⁴⁷

Based on research into this and other areas, including the success of girls and women in science and math, it has been established that when someone is confronted with a situation that is consistent with a stereotype, and that stereotype places his or her identity with a negative contingency, and if the person cares about this, then performance suffers. For example:

- Girls perform as well as boys in math tests when there is no stereotype threat, but significantly worse when there is high stereotype threat.³⁹

- It starts young: Girls aged 5–7 show worse performance on a math test if they first color a picture of a girl holding a doll, compared to coloring a picture with an Asian child eating with chopsticks, or a landscape.²

- Women do well at math tasks if told that women do well at these sorts of exercises, or if they read about successful women before the test.³⁰

- White men perform worse in math when reminded of the Asian math stereotype.³

- Black men perform better than white in athletics ability tests.⁴²

- But worse if the problem is presented as a problem of “sports strategic intelligence.”⁴²

- Note that the stereotype does not have to be explicitly mentioned for the effect to be felt.⁴⁰

People in general do not report they are under stereotype threat; they say they do not feel any stress.⁴⁷ But there are physiological effects that can be measured: blood pressure, sweat; and which correlate with performance.¹⁴

Why is this important? We want to distinguish between people who can do work but are stressed, and people who cannot do the work. Moreover, these effects are continuous: it does not end at the job interview. An ideal environment allows everyone to perform at their maximal level, but stereotype threat interferes with this.

Unsung Heros

Many women¹ and minorities⁴⁷ have been essential contributors to the development of computing; but their contributions are not well known.²⁸ Here are just a few:

Augusta Ada Byron King (1815–1852). Countess of Lovelace, analyst, mathematician, the world’s first programmer Ada wrote a program for Babbage’s Analytical Engine that computed a sequence of Bernoulli numbers.

Rear Admiral Grace Hopper (1906–1991). Hopper was a mathematician who joined the Navy during the World War II. She was assigned to Harvard, where she initially programmed the Mark I to solve complex differential equations. She was a fantastic educator and spokesperson for computing and the sciences, known for her explanation of the nanosecond. She was one of the developers of Cobol and an early advocate of compiled languages, and believed strongly that the future was going to be in the speeding up of networks and computing speed.

The ENIAC Programmers: Betty Snyder Holberton, Jean Jennings Bartik, Kathleen McNulty Mauchly Antonelli, Marlyn Wescoff Meltzer, Ruth Lichterman Teitelbaum, and Frances Bilas Spence. During World War II, many women with a mathematics background were hired to compute ballistic trajectory tables at the University of Pennsylvania. In 1945, six of these women were assigned to the ENIAC, the first all-electronic digital computer. Equipped only with logic diagrams, they built the first programs, using the 3,000 switches and dozens of cables and digit trays. Betty Holberton invented the first Sort routine, and the first software application.²⁸

Frances Elizabeth Allen, recipient of the 2006 ACM A.M. Turing Award, for pioneering contributions to the theory and practice of optimizing compiler techniques that laid the foundation for modern optimizing compilers and automatic parallel execution.

Barbara Liskov, recipient of the 2009 ACM A.M. Turing Award, for contributions to practical and theoretical foundations of programming language and system design, especially related to data abstraction, fault tolerance, and distributed computing.

Clarence “Skip” Ellis was the first African American to receive a Ph.D. in computer science, in 1969. Among his many accomplishments, he was on the team that built the first icon-based GUI at Xerox PARC.

Mark Dean was the first African American to be named an IBM Fellow, in 1996.

He holds three of the original nine patents for the IBM PC, and was on the team that developed the ISA systems bus, that allows the PC to communicate to external devices.

reotype threat interferes with this.

Stereotype threat can be mitigated in a number of ways. Note that developing trust is essential, and several of these suggestions help:

- A credible statement that “this test has been shown to not be subject to the stereotype threat” or “this test is a study of problem solving, not diagnostic of individual ability.”³⁹

- Remind the subject of a positive stereotype: “You are a student at Stanford,” or “You are Asian.”³⁶

- Critical mass is defined by Steele as “the point at which there are enough minorities in a setting that other minorities no longer feel uncomfortable there because they are minorities.” An exact number needed is difficult to define, but in orchestras, 40% women was seen to be the critical mass: “the point at which men and women alike began to report more satisfying experiences.”⁴⁷ In the U.S. Supreme Court, adding a second woman made a difference.⁴⁴ In another experiment, a job brochure had photographs of employees; African American applicants

start to feel safe in applying once they saw 33% of the faces being minority, whereas white applicants felt safe at all levels of minorities (33% was the highest used in the experiment).³⁷ When women take a math test in a group of three women, they do better than women in groups with two women and one man; and women in those groups do better than women in groups with one woman and two men.²⁵

- Offering a credible narrative: It helps to hear the experiences of people who have been through similar experiences and who have overcome the situation.⁴⁷ This includes an expandable view of intelligence (that is, you can learn), and self-affirmation, as explained in the next point:

- Spending 15 minutes writing an affirmation of one’s positive values has been shown to have an enduring positive effect on classroom performance.¹²

- Improving critical feedback: neutral feedback, or encouragement, does not work to reduce the stress of identity threat. It works to say, “I have high standards, I think you can meet

them, and we are going to work together to make this happen.” The difference is engagement.¹³

► **Communities of support:** a key difference between African American students and Asian students at U.S. colleges turned out to be that the Asian students would get together, share their experiences, and study together. People who feel isolated sometimes refuse help, as they think that it will confirm the stereotype. The African American students did not tend to form study groups naturally; when these communities were introduced, their performance increased dramatically.⁴⁵

► **Fostering intergroup conversations as learning opportunities.** These conversations are difficult, as people are afraid of appearing to be biased or ill informed, so they often steer away from the situation. Subconsciously, they pull their chairs further apart. When a facilitator starts a conversation by saying that tensions are natural, but the conversation should be treated as a learning experience, that signals to people that differences can be learned. People pull their chairs closer together. This does not happen when the facilitator says the participants will not be judged, nor when the facilitator says all points of view are welcome.¹⁸

Consequences of stereotyping. As mentioned earlier, there are a number of other forces that can hinder the ability of diverse teams to function optimally.

Stereotypes make us feel as though we have useful information about people’s strengths, weaknesses, and personal characteristics. But they operate in a more prescriptive way as well: They shape our expectations of what people should be doing, especially at work. Thus, women are expected to be nurturing and collaborative at work, in accord with their stereotyped strengths. When they deviate from that, they incur penalties. Men are not expected to be altruistic, but if they are, then they are given credit for giving assistance. This was demonstrated in an experimental setting in which a person was asked for help with a technical task: men were given an increased performance rating if they gave assistance, whereas a woman who assisted was not; her performance rating remained at the base level. If a man refused to



Most people would not believe that height predicts competence, and yet, these choices are frequently made. In fact, height is strongly correlated with career success.



assist, then the rating of his work was not changed from the norm; if a woman refused to assist, her performance rating decreased.²¹

If a man and a woman perform a male-stereotyped task together, the majority of people will attribute the success of that task to the man, unless

- the contributions of the woman are specifically attributed, or
- there is information about the way the task was structured, or
- there is some clear example of prior competence by the woman.²²

There are differences in the communication styles considered acceptable in men and women leaders: women have a very narrow band of acceptable behaviors. In general, women are given less time to speak, are more likely to be interrupted, and if they do interrupt someone, that is most likely to be another woman.⁴³ As a consequence, women are less likely to be able to hold the floor in meetings: an important quality in a leader.

Career advancement strategies that work for men do not always work for women. When comparing only women who have not taken career breaks for family with men, there are still significant differences in achievement levels.¹⁰ The study on the Myth of the Ideal Worker says, “Women benefit most by making their achievements known. Men benefit most by scanning for external opportunities and blurring work-life boundaries. Both benefit by gaining access to powerful others.” The report also says, “changing jobs accelerated compensation growth for men, but slowed it for women.” It has been shown that sponsorship, in which a sponsor actively promotes and takes risks for the sponsee, is more effective for women than mentorship, in which the mentor merely gives advice.²⁴

There are other career differences observable between men and women as well: women are more than 10% less likely than men to change jobs for a raise in salary or for a promotion; but they are almost 10% more likely than men to change jobs because of a bad manager (see Figure 5).⁹ The numbers who leave jobs to take care of family are very small for both men and women. In Japan as well, significantly fewer women report they have off-

ramped because of childcare-related reasons (32%) than those who have off-ramped because they feel stymied and stalled at work (49%).¹¹ One interpretation: It is difficult to balance both working and parenting, and one is not likely to want to stay unless the work is rewarding. Moreover, career priorities differ: A one-size-fits-all management structure will not work well for a diverse organization.

This theory is supported by a study of why women stay in engineering: people who choose to stay are engaged and basically hopeful, and are proactive about the problems they see.⁸

Even a small bias can result in a large difference in the representation of minorities at the top levels of a company. A simulation of promotions was performed with only 1% of bias in ratings of women, and a posited eight levels for promotion. The simulation starts with 50% women at each level, and ends when the entire organization has been replaced with new employees. At the end of the simulation, at the lowest level of the organization there are 53% women, and the top levels of an organization go to only 35% women.²⁹

Transgender studies show some fundamental differences in the ways that society treats the same person as either a man or a woman. These studies are particularly interesting in that the same person experiences life as both genders. The studies consistently show that men who change to women have an average lower salary after the change, and women who change to men have a higher salary, as well as finding everyday life situations easier.³⁸ A very powerful example is the history of two biologists at Stanford University, Ben Barres (a man, formerly a woman) and Joan Roughgarden (a woman, formerly a man).⁴⁶

Ben Barres has said, “When it comes to bias, it seems that the desire to believe in a meritocracy is so powerful that until a person has experienced sufficient career-harming bias themselves they simply do not believe it exists ... By far, the main difference that I have noticed is that people who don’t know I am transgendered treat me with much more respect: I can even complete a whole sentence without being interrupted by a man.” Joan

Roughgarden, on the other hand, has said that as a woman, “You get interrupted when you are talking, you can’t command attention, but above all you can’t frame the issues.”

Ben Barres also wrote a strong commentary in *Nature*⁶ refuting Lawrence Summers’ speech, in which Summers implied that women inherently have less aptitude for science. He cited studies showing that many selection processes set an extremely high bar for competence for women and minorities: 2.5 times in the case of a research grant proposal. And yet, most people have a strong desire to believe the world is fair, so there is widespread belief that little discrimination exists. He gave a number of recommendations to reduce discrimination.

Case Studies: What Works

Many of the issues described so far are major societal issues, which lead to a diffuse pipeline of minority students into the workforce. Despite the challenges posed by a narrow pipeline, bold leaders have made significant strides in reducing the gender gap in computing, as we will show through a few examples.

At Harvey Mudd College, a liberal arts college of science and engineering, the college went from the average 12% of women in CS to about 40% in five years by taking these actions:¹⁹

- The introductory CS class was changed to be more holistic, and not assume the students already know how to program. The act of designing systems to solve problems that one

cares about is a big motivator. A CS class is required for all students in the first semester.

- First-year women students are encouraged to attend the Grace Hopper Celebration of Women in Computing, a large (about 4,000 attendees) conference, regardless of major; a large percentage of the students choose to attend. This provides role models and breaks the stereotype about what computing is about: hundreds of happy, successful, technical women attend, at all stages of their careers. Even the women who choose not to study computer science say this event has a huge impact on their career choices, and many do decide to study computer science as a result.

- A research program was offered after the first year for about 10 female students each year for four years. At this time, it is still true that some students (both male and female) have access to research after the first year.

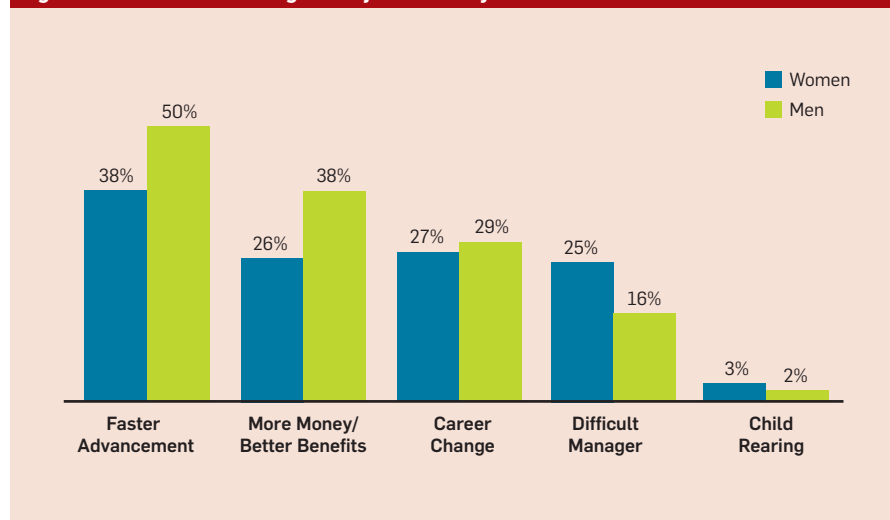
None of these suggestions offer a particularly new insight alone, but in combination they have had a dramatic and lasting effect. Consider these results:

- In the same period, enrollment by women at Harvey Mudd increased from 33% to 46%.

- These changes have attracted more computer science majors overall: Harvey Mudd graduated over 70 CS majors in 2014, compared formerly with 25 to 30.

- In 2014, 56% of those graduating with engineering degrees from Mudd were female. The proportion of grad-

Figure 5. Reasons for leaving a first job. Courtesy of Carter et al.⁹



uating CS majors who were female was 39.4%.

Taken together, they break the stereotypes that girls often have about CS, they provide confidence and role models, and they show all students, not just women, the impact they can have in computer science.

The importance of collecting and publishing data is illustrated by the example of the science faculty at MIT. Around 1993, a few of the senior women faculty in science at MIT felt there was bias in the way lab space was allocated, committee appointments were made, and so on. They were able to get support from the dean, and a comprehensive study was undertaken, starting in 1995, which showed that indeed, in some of the science departments, there were differences in salary, resources, awards, and allocated space between men and women with similar accomplishments. The committee then established goals to address these issues. An important factor in the success of the study was the use of data, and the inclusion of both senior women faculty and men who were or had been department chairs, in the committees. A report was published in 1999, and it is all the more remarkable for its openness.³¹ Importantly, part of the follow-up was to continue to collect and review data. A follow-up study in 2011 showed distinct progress: for example, a woman was hired as president of the institute. Hiring was increased: women represented about 8% of faculty for the 10, or even 20, years prior to the study. Women faculty increased from 30 to 52 in science, and 32 to 60 in engineering; and women now hold some senior administrative positions.

To give an example of the contribution of diversity to team success, the closed captioning feature of YouTube at Google was developed and its release driven by a deaf engineer, Ken Harrenstien.²⁰ Closed captioning turned out to have a huge business impact, beyond the community of those who cannot hear; well beyond what was anticipated. It happened, in large part, because of the efforts and advocacy of Harrenstien and the accessibility engineer who worked with him, Naomi Black.



Stereotypes make us feel as though we have useful information about people's strengths, weaknesses, and personal characteristics.



Conclusion

Diversity is important to organizations that innovate, but the culture of an organization determines whether minority members of the community can thrive. Computer science as a discipline has not done well in attracting and retaining women and minority practitioners in any proportional scale to population representation: the percentage of BS degrees conferred to women peaked in 1986 and is still on a downturn.³⁴ People from minority groups have contributed at the highest level to the CS discipline since its inception, despite obstacles they have faced, including limited visibility of their achievements. As noted earlier, as little as 1% of bias in ratings can result in reduced promotion rates and a skewed population of the top levels of organizations.²⁹ This makes awareness of the reproducible relevant research extremely important.^b

To summarize, things that can be done by leaders to make diverse teams effective:

Make data available. Note the stunning changes that were effected as a result of the MIT Science Faculty Study, because of the insights provided by the data collected, and the open publication of the results.³¹

Create an atmosphere of trust. Recall that women who stay in engineering are engaged and hopeful.⁸

Pay attention to critical mass.^{25,41,44} In order to attain critical mass, organizations should take actions to remove bias from the hiring and promotion processes.

Provide a credible narrative. Provide opportunities for everyone to see themselves as successful: to meet more experienced, successful people,

^b This research is particularly meaningful to me because of my experience of living in foreign countries from 1995–2014. I have come to understand that people see me through the lens of their own experiences and cultural attitudes. Similarly, I have come to better understand the biases that my own native culture has toward these other cultures, sometimes causing fundamental misunderstandings. Probably the most effective training that a person can have in terms of understanding what it is to be a minority is to become one: live in a place where you are the only one like you. However, these studies can help to understand some of what it means in day to day life to be Other.

similar to themselves, and who have faced the same barriers.^{30,41} Know the history of minority contributors to the field, and make sure these achievements are known.

Adopt an expandable view of intelligence: demonstrate you believe that skills can be learned.⁴¹

Know your own biases. Read some of the literature about unconscious bias and about the IAT, and then take the Implicit Attitude test⁵ at <https://implicit.harvard.edu/implicit/>.

Embrace differences: recall the women in the management study,⁹ and as a manager, pay attention to differences in needs by individuals.

Foster intergroup conversations as learning opportunities.¹⁸

Remember the subject of a bias is not always aware of the effects on him or her.⁴¹

For organizations, it is important for individuals in key positions to sponsor promising women and minorities.^{10,24}

An organization that says “we value diversity” is more trusted than one that says “we are color blind.”³⁷

Values affirmation has been shown to have long enduring positive effects on classroom performance.^{12,41}

Before important decisions, make sure you are well fed: in one study, subjects were tested for their bias against homosexuals. Half of the subjects drank lemonade with sugar before the test; the other half drank lemonade with a sugar substitute. The subjects who had sugar showed less bias than those who had a sugar substitute.¹⁶

Acknowledgments

I have conducted this research personally because of my interest in the subject, and any omissions are inadvertent. I would like to thank the many people who have sent me relevant literature, but most especially Caroline Simard, who has studied this area extensively (Note that the body of relevant research is larger than what is cited here.) My particular thanks to Robin Jeffries and Brian Welle for their critical reading of and comments on this article, to Bryant York for the reference to the computer scientists of the Black Diaspora, and to Alan Eustace for his instrumental role in making diversity a priority at Google. 

References

For a full reference list, please refer to *Supplemental Materials* in the ACM Digital Library or to my blog <http://rule-of-one.blogspot.com>.

1. The Ada Project. Pioneering Women in Computing Technology. Carnegie Mellon University; <http://www.women.cs.cmu.edu/ada/Resources/Women/>
2. Ambady, N., Shih, M., Kim, A., Pittinsky, T.L. Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science* 12, 5 (Sept. 2001), 385–390.
3. Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M. and Brown, J. When white men can't do math: Necessary and sufficient factors in stereotype threat. *J. Experimental Social Psychology* 35 (1999), 29–46.
4. Ashcraft, C. and Breitman, A. Who Invents IT? An Analysis of Women's Participation in Information Technology Patenting. NCWIT Report, Dec. 2006.
5. Banaji, M.R. and Greenwald, A.G. *Blindspot: The Hidden Biases of Good People*. Delacorte Press, 2013.
6. Barres, B. Commentary: Does gender matter? *Nature* 442 (Aug. 2006), 133–136.
7. Bateson, M., Nettle, D. and Roberts, G. Cues of being watched enhance cooperation in a real-world setting. *Biology Letters* 2 (June 27, 2006), 412–414.
8. Buse, K.R. Why They Stay: Individual Career Factors Predicting Career Commitment for Women Engineers. Social Science Research Network, May 2011.
9. Carter, N.M. and Silva, C. Pipeline's Broken Promise. Catalyst Report, 2010.
10. Carter, N.M. and Silva, C. Myth of the Ideal Worker. Catalyst Study, Oct. 2011.
11. Center for Work Life Policy. Off-Ramped Women May be the Answer to Japan's Demographic Crisis Finds New Study from the Center for Work-Life Policy, November 2011.
12. Cohen, G.L., Garcia, J., Apfel, N. and Master, A. Reducing the racial achievement gap: A social-psychological intervention. *Science* 313, 5791 (Sept. 2006), 1307–1310.
13. Cohen, G.L., Steele, C.M. and Ross, L.D. The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin* 25 (1999), 1302–1318.
14. Croizet, J.C., Depr s, G., Gauzins, M.E., Huguet, P., Leyens, J.P., M ot, A. Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *J. Personal and Social Psychology* 30, 6 (June 2004), 721–731.
15. Desvaux, G., Devillard-Hoellinger, S. and Sancier-Sultan, S. Women Matter: Women at the top of corporations: Making it happen. McKinsey Report, 2010.
16. Gailliot, M.T., Peruche, B.M., Plant, E.A., and Baumeister, R.F. Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *J. Experimental Psychology*, Jan. 2009.
17. Gladwell, M. *Blink*. Little Brown, 2005.
18. Goff, P.A., Steele, C.M. and Davies, P.G. The space between us: Stereotype threat and distance in interracial contexts. *J. Personality and Social Psychology* 94, 1 (2008), 91–107.
19. Haffner, K. Giving women the access code. *New York Times* (Apr. 2, 2012); <http://www.nytimes.com/2012/04/03/science/givingwomentheaccess-code.html>
20. Harrenstien, K. Finally, Caption Playback. Google Video Blog (Sept. 2006); <http://googlevideo.blogspot.com/2006/09/finallycaptionplayback.html>
21. Heilman, M.E. and Chen, J.J. Same behavior, different consequences: Reactions to men's and women's altruistic citizenship behavior. *J. Applied Psychology* 90, 3 (2005), 431–441.
22. Heilman, M.E., Chen, J.J. No credit where credit is due: Attributional rationalization of women's success in male-female teams. *J. Applied Psychology* 90, 5 (2005), 905–916.
23. Herring, C. Does diversity pay? Race, gender, and the business case for diversity. *American Sociological Review*, Apr. 2009.
24. Hewlett, S.A. Leader-Chiv e, L., Sumberg, K., Fredman, C. and Ho, C. Sponsor Effect. Center for Talent Innovation, U.K., 2012.
25. Inzlicht, M. and Ben-Zeev, T. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science* 11 (2000), 365–371.
26. Joy, L., Carter, N.M., Wagner, H.M., and Narayanan, S. The Bottom Line: Corporate Performance and Women's Representation on Boards. Catalyst Report, Oct. 2007.
27. Judge, T.A. and Cable, D.M. When it comes to pay, do the thin win? The effect of weight on pay for men and women. *Journal of Applied Psychology* 96, 1 (Jan. 2011), 95–112.
28. Kleiman, K. Eniac Programmers Project; <http://eniaprogrammers.org/>
29. Martell, R.F., Lane, D.M., Emrich, C. Male-female differences: A computer simulation. *American Psychologist* 51, 2 (Feb. 1996), 157–158.
30. McIntyre, R.B., Paulson, R.M. and Lord, C.G. Alleviating women's mathematics stereotype threat through salience of group achievements. *J. Experimental Social Psychology* 39 (2003), 83–90.
31. MIT Faculty Newsletter: A study on the status of women faculty in science at MIT, March 1999; <http://web.mit.edu/fnl/women/women.html>
32. Moss-Racusin, C.A., Dovidio, J.F., Brescoll, V.L., Graham, M.J. and Handelsman, J. Science faculty's subtle gender biases favor male students. In *Proceedings of the National Academy of Science* (Aug 2012).
33. NASSCOM-Mercer. *Gender Inclusivity in India: Building Empowered Organisations*, 2009.
34. Mitchell, R.L. Women computer science grads: The bump before the decline, Computerworld Blogs, (April 2013); <http://blogs.computerworld.com/itcareers/21993/womencomputersciencevisual-trendline>
35. Pieper, J. *Leisure, the Basis of Culture*. Ignatius Press, 2009.
36. Pittinsky, T.L., Shih, M., Ambady, N. Identity adaptiveness: Affect across multiple identities. *J. Social Issues* 55 (1999), 503–518.
37. Purdie-Vaughns, V., Steele, C.M., Davies, P.G., Dittmann, R. and Crosby, J.R. Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *J. Personality and Social Psychology* 94, 4 (2008), 615–630.
38. Schilt, K. and Wiswall, M. Before and after: Gender transitions, human capital, and workplace experiences. The B.E. *Economic Analysis and Policy* 8, 1 (2008), 1–28.
39. Spencer, S.J., Steele, C.M., Quin, D.M. Stereotype threat and women's math performance. *J. Experimental Social Psychology* 35 (1999), 4–28.
40. Steele, C.M. and Aronson, J. Stereotype threat and the intellectual test performance of African Americans. *J. Personality and Psychology* 69, 5 (Nov. 1995), 797–811.
41. Steele, C. *Whistling Vivaldi*. W. W. Norton, 2011.
42. Stone, J., Lynch, C.I., Sjomeling, M., Dartley, J.M. Stereotype threat effects on black and white athletic performance. *J. Personality and Social Psychology* 77, 6 (1999), 1213–1227.
43. Tannen, D. The power of talk: Who gets heard and why. *Harvard Business Review*, Sept. 1995.
44. Totenberg, N. Sandra Day O'Connor's Supreme Legacy: First Female High Court Justice Reflects on 22 Years on Bench. All Things Considered (May 14, 2003); <http://www.npr.org/templates/story/story.php?storyId=1261400>
45. Treisman, U. Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal* 23 (1992), 362–372.
46. Vedantam, S. *The Hidden Brain*. Random House Publishing Group, 2010.
47. Williams, S. Computer Scientists of the African Diaspora. SUNY Buffalo; <http://www.math.buffalo.edu/mad/computer-science/cs-peeps.html>
48. Woolley, A.W. Chabris, C.F., Pentland, A., Hashmi, N. and Malone, T.W. Evidence for a collective intelligence factor in the performance of human groups. *Science Magazine*, 330 (Oct. 29, 2010), 686–688; DOI: 10.1126/science.1193147.
49. Woolley, A. and Malone, T. Defend your research: What makes a team smarter? More women. *Harvard Business Review Report*, June 2011.
50. Zweben, S. and Bizot, B. 2012 Taulbee Survey: Strong increases in undergraduate CS enrollment and degree production; record degree production at doctoral level. *Computing Research News* 25, 5 (May 2013).

Beryl Nelson (beryl.nelson@gmail.com) is a software engineering manager at Google, Mountain View, CA.

Copyright held by author.

research highlights

P. 97

**Technical
Perspective**
**The Intricate Dance
of Fabric and Light**

By Szymon Rusinkiewicz

P. 98

**Building Volumetric
Appearance Models of Fabric
Using Micro CT Imaging**

By Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala

Technical Perspective

The Intricate Dance of Fabric and Light

By Szymon Rusinkiewicz

WHAT IS IT about different materials, beyond their color, that gives them their characteristic appearance? Over the decades this fundamental question has given rise to a branch of computer graphics that can truly be called a science: the search to understand how complex materials interact with light. And as with any science, this understanding comes from the combination of sophisticated mathematical modeling and detailed experimental work.

As light encounters any material, it may be reflected, scattered, or absorbed. Producing accurate and believable computer graphics renderings requires algorithms that reproduce these interactions. In the simplest systems, the models for how light reflects from a surface are far from realistic, giving all surfaces the plastic appearance that has become so commonly associated with computer graphics. Curing this problem, especially for fabrics, requires a series of conceptual leaps, leading to more complex models and requiring more complex algorithms.

First, it is necessary to abandon the notion that fabrics are surfaces. Though we may speak of the “surface” of a piece of cloth, a realistic simulation of light requires treating the cloth as a volume of small, but definitely non-zero, thickness. This volume is populated by fibers, all of which can scatter or absorb light. We thus must shift from a rendering system in which light is thought of as being reflected from surfaces, to one in which it is scattered inside volumes. Though volume rendering has been commonly used to model translucent objects such as clouds or marble or even human skin, it is only in recent years that researchers, such as the authors

of the following paper, have understood how important it is to model cloth in this way. A volumetric simulation can give rise to an appearance that “looks right” not only from far away, but also in extreme close-up.

Once we accept that fabric will be modeled as a collection of fibers, it is necessary to actually locate all of those fibers in space. In some cases the fibers are ordered, as in a weave or knit. In others, such as felt, they are arranged randomly. For years, our only option in graphics has been to write programs that simulate these arrangements, requiring tweaks to both the algorithms and their myriad parameters for each new fabric.

The following paper solves the problem in a creative way, by making use of actual fabric samples and micro-CT imaging. This is an instance of a trend in graphics that has been important since the mid-1990s: the use of new imaging and acquisition technologies (ranging from digital

cameras to 3D scanners and even industrial MRI machines) to capture the complexity of the real world that is difficult to model analytically or with simple algorithms.

Finally, we must model the way light is scattered and absorbed by each individual fiber. Here, the authors use a combination of theory and experiment. They propose an analytic model that is consistent with the underlying physics (for example, energy conservation) while having a number of unknown free parameters. Those parameters, in turn, are selected such that a rendering (including the full volumetric light scattering simulation and the measured fiber arrangement) matches the statistics of a photograph of the actual fabric. Of course, this process allows for proper experimental procedure, in which parameters that are fit from one photograph yield predictions that may be compared against another photograph.

The authors of the following paper present a state-of-the-art example of how algorithms (for volumetric light scattering simulation), math (dictating how scattering happens at a single point in the volume), and measurement (micro-CT to obtain fiber arrangement) all combine to help our understanding of how fabrics interact with light. The paper not only epitomizes the modern approach to understanding appearance, but also introduces new theoretical and experimental tools that will translate to many other kinds of materials. 

The authors present a state-of-the-art example of how algorithms, math, and measurement all combine to help our understanding of how fabrics interact with light.

Szymon Rusinkiewicz (smr@princeton.edu) is a professor in the computer science department of Princeton University, Princeton, NJ, where he also a member of the Princeton Computer Graphics Group.

Copyright held by author.

Building Volumetric Appearance Models of Fabric Using Micro CT Imaging

By Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala

Abstract

Cloth is essential to our everyday lives; consequently, visualizing and rendering cloth has been an important area of research in graphics for decades. One important aspect contributing to the rich appearance of cloth is its complex 3D structure. Volumetric algorithms that model this 3D structure can correctly simulate the interaction of light with cloth to produce highly realistic images of cloth. But creating volumetric models of cloth is difficult: writing specialized procedures for each type of material is onerous, and requires significant programmer effort and intuition. Further, the resulting models look unrealistically “perfect” because they lack visually important features like naturally occurring irregularities.

This paper proposes a new approach to acquiring volume models, based on density data from X-ray computed tomography (CT) scans and appearance data from photographs under uncontrolled illumination. To model a material, a CT scan is made, yielding a scalar density volume. This 3D data has micron resolution details about the structure of cloth but lacks all optical information. So we combine this density data with a reference photograph of the cloth sample to infer its optical properties. We show that this approach can easily produce volume appearance models with extreme detail, and at larger scales the distinctive textures and highlights of a range of very different fabrics such as satin and velvet emerge automatically—all based simply on having accurate mesoscale geometry.

1. INTRODUCTION

Cloth is a fundamental material in our day-to-day lives, and creating photo-realistic renderings of cloth has been an active research topic in computer graphics for decades, with applications in areas including virtual prototyping, entertainment (movies and games), and retail.

One important aspect contributing to the appearance of cloth is its complex 3D structure, yielding complicated textures and reflectance. Further, the structure is irregular, causing difficult-to-model, but visually important randomness. Volume rendering techniques, which model such structure correctly and simulate the interaction of light with cloth explicitly, have been explored since the 1990s. These approaches address the limitations of usual surface-based models, which are visually unsatisfactory because they treat cloth as infinitely thin sheets.^{8, 12, 18} Cloth exhibits a wide range of appearance, but shares a common basic structure of long, shiny fibers. The thick, fuzzy nature of cloth makes

volume models a good fit. Further, recent developments⁷ have brought enough generality to volume scattering that we can begin to render fully physically based *volumetric appearance models* for cloth, fur, and other thick, non-surface-like materials. However, a fundamental problem remains: creating these volumetric models themselves. Previous work has primarily relied on procedural methods (special-purpose programs) for modeling these volumes, but this approach has limited generality: significant creative effort is needed to design these programs for each new material. Further, the resulting models look unrealistically “perfect” since they lack subtle irregularities that appear in real cloth.

This paper explores an entirely different approach to building volume appearance models, focusing particularly on cloth. Since cloth’s detailed geometric structure is so difficult to model well, we use volume imaging to measure structure directly, and then fill in optical properties using a reference photograph. We do the latter by solving an inverse problem that statistically matches photographs and physically based renderings.

Many volume imaging technologies have been developed, including computed tomography (CT), magnetic resonance, and ultrasound, but unlike photographs, the resulting data does not directly relate to the optical appearance of the material: only to its structure. As a result, volume renderings of these images are useful for illustrating hidden internal geometry, but not directly for rendering realistic images. For instance, a micro CT scan of woven cotton cloth gives a detailed view of the interlaced yarns and their component fibers, showing exactly how the fibers are oriented and how the yarns are positioned, but no information about how they interact with light: there is no way to tell whether the fabric is black or white or any color in between.

We show in this paper that remarkably little additional information is required to extend CT data to a realistic appearance model. The value of knowing 3D structure is obvious for rendering close-up views where these details are visible. But equally importantly, the shape and arrangement of fibers in the material also determines the overall appearance of the material—the shape and quality of specular highlights, and how the visual texture varies with illumination and view. When coupled with the right rendering

The original version of this paper was published in *ACM Trans. Graph.* 30, 4 (2011).

technology, a simple local model of reflection from fibers automatically predicts the characteristic appearance of very different materials such as velvet and satin, simply by knowing the 3D structure of the material.

The contribution of this paper is to show how to enhance the structural information from a CT scan of a small sample of fabric by combining it with appearance information from a photograph of the material to construct plausible and consistent optical properties; this volumetric appearance model produces realistic appearance when rendered using a physically based volume renderer. We describe our end-to-end volume appearance modeling pipeline and demonstrate it by acquiring models of cloth with very different appearance, ranging from matte to shiny and textured to smooth, capturing their characteristic highlights, textures, and fuzziness.

2. RELATED WORK

We categorize realistic volumetric rendering and modeling research in the related areas of surface appearance modeling, cloth reflectance modeling, and cloth structure modeling.

Appearance modeling. Because standard surface-oriented models are inadequate for complex thick materials, researchers and practitioners have had to fall back on image-based rendering methods such as Bidirectional Texture Functions (BTF), which essentially consist of an exhaustive set of photographs of the surface under all possible illumination and viewing directions.^{4,5} Although BTFs produce realistic results for many otherwise difficult materials, the image-based approach requires a significant amount of storage, and is often not of enough resolution for capturing high glossiness, and generally fails to capture or predict grazing angles, making silhouettes and edges unrealistic.

Two prominent early volume appearance models are Kajiya and Kay's⁸ fur rendering and Perlin and Hoffert's¹² "hypertexture." Although it has since become more common to render hair and fur using discrete curves, their results demonstrate the value of volumetric models for complex, barely resolved detail. A similar approach is the Lumislice representation^{3,18} which focused on modeling and rendering knitwear. Magda and Kriegman¹¹ describe a method for acquiring *volumetric textures* that combine a volumetric normal field, local reflectance functions, and occupancy

information. All these approaches need significant modeling effort. Recently, Jakob et al.⁷ introduced a principled formulation for rendering anisotropic, oriented volumetric media, which opens possibilities for more physically based volume appearance models.

Cloth reflectance models. Cloth has perennially appeared in graphics as a source of complicated optical behavior. Westin et al.¹⁷ modeled cloth's reflectance profile by ray-tracing mesostructure models, which is related to the way cloth highlights emerge in our system. Ashikhmin et al.² rendered velvet and satin using hand-designed microfacet distributions. Adabala et al.¹ proposed a rendering method for woven cloth based on microfacet models, and Irawan and Marschner⁶ presented an elaborate model, based on the analysis of fiber tangent directions in a range of woven fabrics, and validated it against reflectance measurements. Each of these methods achieved good appearance relative to the then-current state of the art, but they are all specially hand-designed models for individual materials or specific classes.

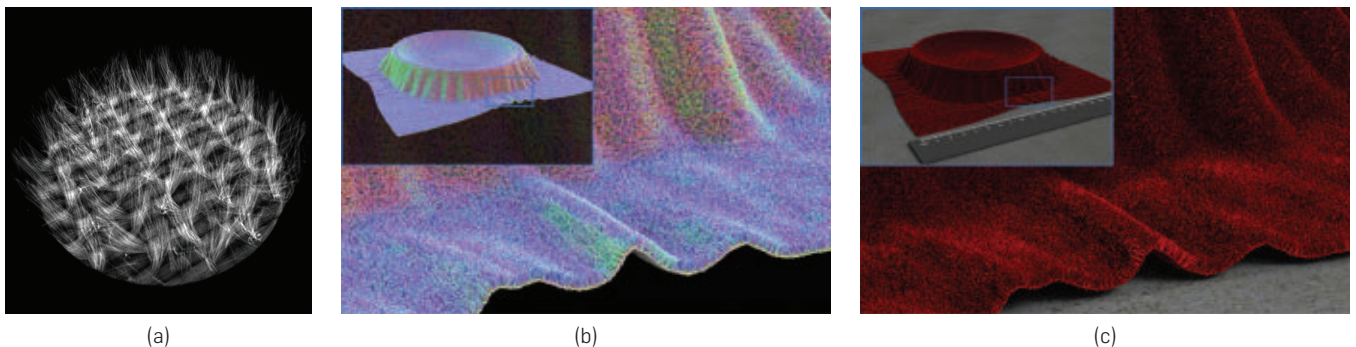
Since our approach is based on a completely general system that only has a volume with fibers as its underlying assumption, we have few fundamental limitations on what textile or textile-like materials can be handled. Further, by importing volumetric detail from the real world, we can achieve good appearance in close-ups, and at silhouettes, edges, and corners, where surface models appear unrealistically smooth and flat.

Cloth structure. The geometry of cloth structure has been studied for decades.^{9,13} More recently, X-ray tomography, using synchrotron facilities¹⁶ or the rapidly improving micro-CT scanners,^{10,15} has been used to examine the structure of textiles in several applications. These studies focus on extracting geometric information related to a material's mechanical properties, but have also produced some analysis tools¹⁵ that we use.

3. OVERVIEW

The goal of our system is to create realistic volumetric appearance models of cloth. We need to generate a sampled 3D volume that describes the optical properties of the material at each voxel so that, when rendered with a physically based rendering system, it realistically reproduces the appearance of real cloth (Figure 1).

Figure 1. We build volumetric appearance models of complex materials such as velvet using CT imaging: (a) CT data gives scalar density over a small volume; (b) we extract fiber orientation (shown in false color) and tile larger surfaces; and (c) we match appearance parameters to photographs to create a complete appearance model. Both fine detail and the characteristic highlights of velvet are reproduced.



Because cloth is made of fibers, we need a volume scattering model that can handle the anisotropy of fibers; we chose a modified version of the model proposed by Jakob et al.⁷ (detailed in Section 4) for this purpose. This model requires an optical density, an albedo, and two phase function parameters: an orientation vector and a specular lobe width. Intuitively, the optical density describes how often light scatters within the cloth; the albedo and the phase function respectively capture the fraction of light being absorbed and how light changes its direction at each scattering location.

Our technique begins with a micro CT scan of a small area of material, showing detail at the level of individual fibers over a fraction of a square centimeter. Such scans can readily be ordered at moderate cost (a few hundred US dollars) from a number of facilities, and suitable desktop CT scanners are becoming available. In a sequence of three stages (Figure 2), we process and augment this data, ending with a volume that defines the required scattering model parameters using density and orientation fields derived from the CT data, plus three global parameters: the albedo, the lobe width, and a density multiplier that scales the density field.

The first stage (Section 5) processes the density volume to augment it with orientation information and to remove noise by convolving the data with 3D oriented filters to detect oriented structures, and thresholding to separate meaningful structure from noise. This stage produces the density and orientation fields.

This volume can be rendered only after the global optical parameters are determined. The second stage (Section 6) makes use of a single photograph of the material under known (but not controlled) lighting, and associates optical properties with the oriented volume from the first stage by matching the texture of the rendered volume to the texture of the photograph.

The resulting volume model is good for rendering small samples; the third stage takes this small patch and maps it over a large surface of cloth, using randomized tiling to replicate the material and shell mapping¹⁴ to warp it.

The resulting renderings (Section 7) show that this unique approach to appearance modeling, leveraging direct information about mesoscale geometry, produces excellent appearance from the small scale, where the

geometry itself is visible, to the large scale, where the directional scattering properties naturally emerge from the measured 3D structure. The characteristic appearance of difficult materials such as velvet and satin is predicted by our rather minimal volume scattering model, even though we use no light scattering measurements that could tell these materials apart, because accurate geometric information is available.

4. FIBER SCATTERING MODEL

We model light transport using the anisotropic radiative transfer equation (RTE) from Jakob et al.⁷ which states that within participating media,

$$(\omega \cdot \nabla)L(\omega) + \sigma_t(\omega)L(\omega) = \sigma_s(\omega) \int_{S^2} f_p(\omega' \rightarrow \omega)L(\omega') d\omega'$$

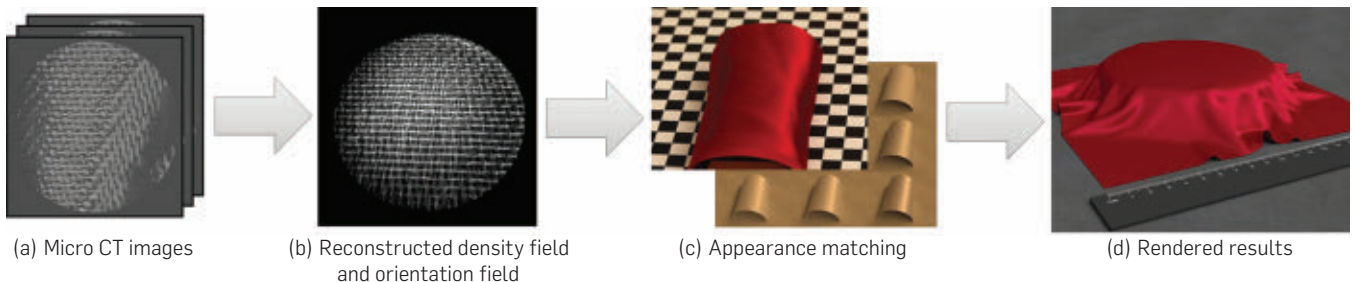
where σ_s and $\sigma_t : S^2 \rightarrow \mathbb{R}$ are the anisotropic scattering and extinction coefficients, and f_p is the phase function. Spatial dependence has been omitted for readability.

This equation can be understood as a generalization of the isotropic RTE that adds support for a directionally varying amount of “interaction” with a medium. For instance, the directional dependence of $\sigma_t(\omega)$ is necessary to model the effect that light traveling parallel to coherently aligned fibers faces less obstruction than light traveling perpendicular to the fibers.

To specify the problem to be solved, we must choose a compatible scattering model that will supply internally consistent definitions of σ_t , σ_s , and f_p . For this purpose, we use the *micro-flake* model proposed in the same work. This volume analogue of microfacet models represents different kinds of volume scattering interactions using a directional *flake distribution* $D(\mathbf{m})$ that describes the orientation $\mathbf{m}$ of (unresolved) idealized mirror flakes at every point in space. Similar to microfacet models, the phase function then involves evaluating $D(\mathbf{m})$ at the half-way direction between the incident and outgoing direction. For completeness, we reproduce the model's definition as follows:

$$\begin{aligned}\sigma_t(\omega) &= a \rho \int_{S^2} |\omega \cdot \mathbf{m}| D(\mathbf{m}) d\mathbf{m}, \\ \sigma_s(\omega) &= \alpha \sigma_t(\omega), \\ f_p(\omega' \rightarrow \omega) &= \frac{a \rho \alpha}{4 \sigma_s(\omega)} (D(h(\omega, -\omega')) + D(-h(\omega, -\omega')))\end{aligned}$$

Figure 2. Our volume appearance modeling pipeline: (a) CT images are acquired; (b) the density field and orientation field of the volume are created; and (c) optical parameters of the volumetric model are assigned by matching statistics of photographs with rendered images. (d) Larger models are rendered using our acquired volumetric appearance and geometry models.



Here, ρ denotes the particle density, a is the area of a single flake, α is the associated albedo, and $h(\omega, \omega') := (\omega + \omega')/(\omega + \omega')$. Note that the above expressions are simplified by assuming the flakes have albedo independent of the scattering angle. This reduces our search space considerably and still leads to a model that can represent scattering interactions with a variety of fibrous materials reasonably well.

4.1. Flake distribution

We propose a flake distribution that is convenient to integrate while capturing the same key feature as the one proposed by Jakob et al.⁷ We use the following density function, which specifies a truncated Gaussian centered around the great circle perpendicular to the local fiber orientation $\bar{\omega}$:

$$D(\omega) = \frac{1}{(2\pi)^{3/2} \gamma \operatorname{erf}\left(\frac{1}{\sqrt{2}\gamma}\right)} \exp\left(-\frac{(\bar{\omega} \cdot \omega)^2}{2\gamma^2}\right)$$

where the standard deviation γ determines the roughness of the fiber. More precisely, the parameters required to create renderings are:

- $\bar{\omega}$, the local fiber orientation,
- γ , the standard deviation of the flake distribution,
- α , the single scattering albedo of the flakes,
- a and ρ , the area and density of micro-flakes. Their product roughly corresponds to the interaction coefficient σ_i in traditional isotropic volume rendering, and we therefore set them to a multiple of the processed CT densities, that is, $a\rho(\mathbf{x}) := d \cdot \text{CT}(\mathbf{x})$, where d is a constant of proportionality.

Section 5 discusses the steps needed to obtain $\text{CT}(\mathbf{x})$ and $\bar{\omega}(\mathbf{x})$; in section 6, we describe how to find α , γ , and d .

5. CT IMAGE PROCESSING

Micro CT (computed tomography) devices, which use X-ray CT methods to examine small to microscopic structures, are increasing in availability, and this imaging modality is suited to a wide range of materials from which a small sample can be extracted for scanning.

In this section we describe the process of extracting fiber orientation from the CT density volume using a special fiber-detecting filter. Following this, we explain the processing steps needed to obtain orientation and density fields suitable for rendering.

5.1. Recovering the orientation field

CT images provide a voxelized density field with no direction information. Since our optical model requires an orientation for the phase function, it is necessary to reconstruct an orientation for every nonempty voxel. Our approach uses oriented filters to detect fibers, based on similar filters used by Shinohara et al.,¹⁵ to locate fibers in CT data.

To detect a fiber with orientation $\mathbf{d}$ at location $\mathbf{p}$, Shinohara proposes a cylindrically symmetric filter oriented

with the axis $\mathbf{d}$, consisting of a difference of Gaussians in distance from the axis: $q(\mathbf{d}; \mathbf{p}) := -2 \exp(-ur^2) + \exp(-wr^2)$, where $r = \|\mathbf{p} - (\mathbf{p} \cdot \mathbf{d})\mathbf{d}\|$ is the distance from the filter's axis and the parameters u and w (normally $u < w$) are empirically adjusted based on the size of the fibers present in the sample (see Figure 3).

The raw CT volume is thresholded at a value ϵ_d , resulting in a binary volume b : for any $\mathbf{x}$, set $b(\mathbf{x})$ to 0 if $\text{CT}_{\text{raw}}(\mathbf{x}) \geq \epsilon_d$ and to 1 otherwise. Then b is convolved with the filter q for each of a fixed set of orientations: $J(\mathbf{x}, \mathbf{d}) := \sum_{\mathbf{p} \in V} q(\mathbf{d}; \mathbf{p}) b(\mathbf{x} + \mathbf{p})$, where V is a cubic volume of edge length h .

As shown in Figure 3, the function J reaches a maximum value when $\mathbf{d}$ equals the fiber's orientation. So the orientation field is computed by finding, for each voxel $\mathbf{x}$, the $\mathbf{d}'$ that maximizes $J(\mathbf{x}, \mathbf{d}')$ and setting $\bar{\omega}(\mathbf{x}) = \mathbf{d}'$. In our implementation, we precompute q on a set of directions $\{\mathbf{d}_i\}$ picked from a $32 \times 32 \times 6$ cubemap. Then for each nonempty voxel $\mathbf{x}$, we set $\bar{\omega}(\mathbf{x}) = \mathbf{d}_j$, where $j = \arg \max_i J(\mathbf{x}, \mathbf{d}_i)$.

5.2. Denoising CT images

The CT images usually contain considerable amounts of noise, particularly for low-density materials like our cloth samples, and removing the noise is critical for obtaining good quality data for rendering. Since cloth structure is always oriented, and the noise is generally fairly isotropic, the value of J is useful in noise removal.

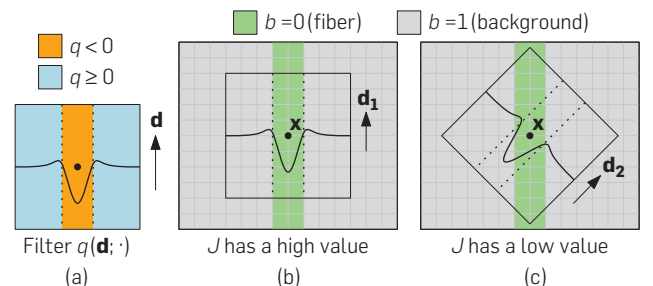
In our system we use two thresholds to remove noise. The first threshold ϵ_d is on the voxel values themselves and is used to remove faint background noise that would otherwise cloud the model. This thresholding creates the binary volume b . The second threshold ϵ_j is on the value of J and is used to remove isotropic noise that has density values that are too high to remove by the first threshold. We set

$$\text{CT}(\mathbf{x}) := \begin{cases} \text{CT}_{\text{raw}}(\mathbf{x}) & \text{CT}_{\text{raw}}(\mathbf{x}) \geq \epsilon_d \wedge J(\mathbf{x}, \bar{\omega}(\mathbf{x})) \geq \epsilon_j; \\ 0 & \text{otherwise.} \end{cases}$$

5.3. Data replication

The volume data needs to be replicated for rendering since our samples are very small. We explored example-based synthesis in Zhao et al.¹⁹ which provides sophisticated tools to do this, but it is orthogonal to this paper. Here we consider two simple randomized tiling methods to cover the surfaces with tiles of volume data drawn from our models without introducing distracting regular structures. In both methods the

Figure 3. Computing function J in 2D: (a) shape of the filter q ; (b) when q is aligned to the fiber; (c) when q is unaligned.



surface is simply covered by a rectangular array of tiles copied from the volume, without continuity at the tile boundaries.

For materials without visible regularity, such as velvet and felt, each tile on the surface is copied from a rectangular region centered in the volume. To provide variation in local structure, for each tile this source rectangle is rotated by a different random angle. For materials with woven structure, such as silk and gabardine, we use a similar approach, but use random translations of the source tile instead of rotations. The weave pattern in each sample is manually identified and a rectangular area is marked that contains an integer number of repeats. Then each (smaller) surface tile is chosen from a subrectangle that contains a matching section of the weave. The result is a tiling that reproduces the correct weave pattern and avoids obvious repeating of texture. We then map the tiled data to arbitrary surfaces using shell mapping.¹⁴

6. APPEARANCE MATCHING

Processing the CT data yields the spatially varying density and orientation for the volume. But the optical appearance parameters of the model remain to be determined. Since the CT scan does not give us the material's optical properties, we make use of a photograph of the material to compute the appearance parameters.

To make the problem tractable, we assume that the volume contains the same material, with differences only in density and orientation. This is appropriate for fabrics made from a single type of fiber, which encompasses many important examples. Fabrics containing yarns of different materials are future work. Thus, the appearance parameters that must be determined are the same across the whole volume. They are: the standard deviation of the flake distribution γ (corresponding to fiber roughness), the scattering albedo α (corresponding to material color), and the density scale d (corresponding to opacity). Figure 4 illustrates the effects of these parameters.

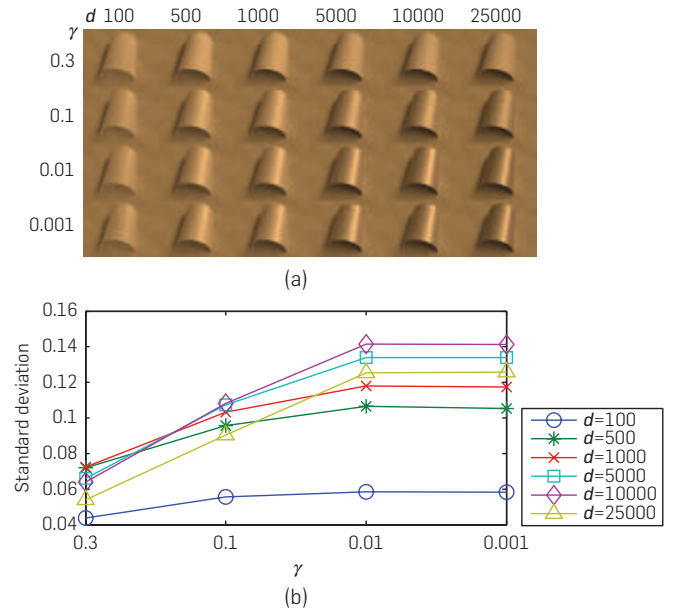
To match the material's optical properties, we must use photographs of the sample. One approach is to photograph the same sample that was scanned, calibrating the camera to the scan and associating pixels in the image with rays in the volume. This calibration and acquisition is nontrivial; the fine resolution of the scans poses practical difficulties. Further, we found that this level of detail is not required to determine the small number of parameter values we seek. Instead, we assume that the fabric is statistically similar across different patches. Thus, our approach is to statistically match the texture of rendered images with a photograph of a different section of the same cloth under uncontrolled but known lighting.

We now describe the metrics we use to match the optical parameters to the photograph, and then describe our matching algorithm.

6.1. Metrics for matching

Appearance matching is not a straightforward process of mapping colors from the photos into the volume, because the volume model describes local scattering properties, but the appearance is defined by a global volumetric multiple scattering process. Our approach is to repeatedly render the

Figure 4. (a) Renderings of a cylinder tiled with the satin volume, with fixed albedo, and varying lobe width γ and density multiplier d . (b) The corresponding standard deviation of pixel values for the satin sample: sharper lobes provide shinier appearance and result in greater standard deviation. The role of d is more complicated.



volume using our physically based renderer and to adjust the optical parameters to match certain texture statistics of the rendered images to statistics of the photograph.

We match two simple statistical measures: the mean pixel value and the standard deviation of pixel values, computed over corresponding regions of a photograph and a rendering of approximately similar geometry. This approach effectively matches the image brightness and texture contrast in the matching region. We tried other measures, but found that the mean and standard deviation measures were simpler and robust. Thus, the only information that flows from the photograph to the volume model is the mean and standard deviation of pixels in a single rectangle.

The appearance matching process involves choosing the geometry, camera position, lighting, and matching region. These are inherently manual choices, and we used the principle of choosing a setup that shows the distinctive features of the cloth's appearance. For instance, we made sure to use a configuration where the highlight was visible on the satin. Beyond this we did not take any special care in arranging the appearance matching inputs, and the results do not appear to be sensitive to the details.

6.2. Optimization procedure

As shown in Figure 4, the density multiplier plays a fairly complicated role with respect to both measures. Given that our forward process, which is essentially Monte Carlo path tracing, is quite expensive, we chose to predetermine the density multiplier in our implementation by rendering such a matrix. Fixing the density multiplier simplifies the inverse problem and leads to a practical solution. We found that the algorithm is not particularly sensitive to the choice of

density multiplier; our results use two main settings which differ by an order of magnitude (see Table 1).

With a fixed density multiplier, we solve for the values of albedo (α , estimated separately in red, green, and blue) and lobe width (γ , a single scalar value) using an iterative algorithm. Note that the mean and standard deviation of pixel values change monotonically with changes in α and γ , respectively. Thus, a binary search can be used to significantly improve performance as follows: first, an initial guess of γ is assumed, and we search for the α to match the mean pixel value. Then, fixing α , we perform a search for the γ to match the standard deviation. These iterations are repeated until a match is found. In practice, this approach converges quickly, usually in two or three iterations.

Finally, we take another photo under a different setup and render a corresponding image as a qualitative validation (see Section 7). Figure 5 shows the appearance matching results for two different materials.

7. RESULTS

Our results are based on samples of silk satin, velvet, felt, and wool gabardine, which were sent to the High-Resolution X-ray Computed Tomography Facility at The University of Texas at Austin. All fabrics were scanned in an XRadia MicroXCT scanner using 1024^3 volumes with a $5\mu\text{m}$ voxel size, which observed circular areas of approximately 5 mm diameter. Our rendering implementation is based on the

open source rendering system Mitsuba, which was extended to handle the new micro-flake distribution (Section 4). All the rendered results are generated using Monte Carlo volume path tracing.

Figure 6 shows the resulting models shell-mapped onto draped fabric geometry and rendered under environment lighting. The corresponding high-resolution renderings are available on the project Web site at <http://www.cs.cornell.edu/projects/ctcloth-sg11>. Fabrics are usually rendered with surface-based models, making them thin 2D sheets so that cut edges look artificial; our volumetric model, on the other hand, explicitly captures 3D structures and provides a proper impression of the thickness and weight of the fabrics. Furthermore, surface-based models show smooth edges at silhouettes while our model is able to produce fuzzy silhouettes with rich details, bringing a new level of realism to fabric rendering.

The silk satin (charmeuse) has a structure of mainly parallel fibers on the surface, resulting in a strong anisotropic highlight. In Figure 5(1), the appearance matching pair uses a cylindrically curved piece of material, and the matching region was chosen to include a highlight to allow the matching process to tune γ appropriately. Good results are obtained despite the mismatch between the ideal cylinder in the rendering and the flatter shape of the real material, illustrating that a casual setup suffices. Using the parameters obtained from this view, the validation pair shows the fabric rotated 90 degrees and draped over the same cylinder. At this angle the fabric exhibits almost no highlight; this anisotropic appearance is correctly predicted by our model.

The satin is shown in a draped configuration in Figure 6(a). No reflectance model such as BTF or other multi-view image data is used for these renderings—the orientation information in the volume automatically causes the characteristic appearance of this fabric to emerge when the model is rendered.

For gabardine, a wool twill fabric, the variation in texture with illumination direction is an important appearance characteristic. In Figure 5(2), the appearance matching pair is lit

Table 1. Fiber scattering model parameter values for our material samples: d , the density multiplier; γ , the standard deviation of the flake distribution; α , the single scattering albedo

Material	Data size	d	γ	α
Gabardine	$992 \times 1012 \times 181$	5000	0.1	(0.892, 0.063, 0.048)
Silk	$992 \times 1013 \times 46$	5000	0.01	(0.699, 0.030, 0.080)
Velvet	$992 \times 1012 \times 311$	500	0.1	(0.555, 0.040, 0.074)
Felt	$992 \times 1012 \times 485$	500	0.125	(0.518, 0.915, 0.365)

Figure 5. Appearance matching results: (top) silk, (bottom) gabardine. Columns (a) and (c) show photographs of the materials, and (b) and (d) show rendered images. The left two columns form the appearance matching pair, in which the blue boxes indicate manually selected regions for performing our matching algorithm. The right two columns, the validation pair, validate our matches qualitatively under different configurations.

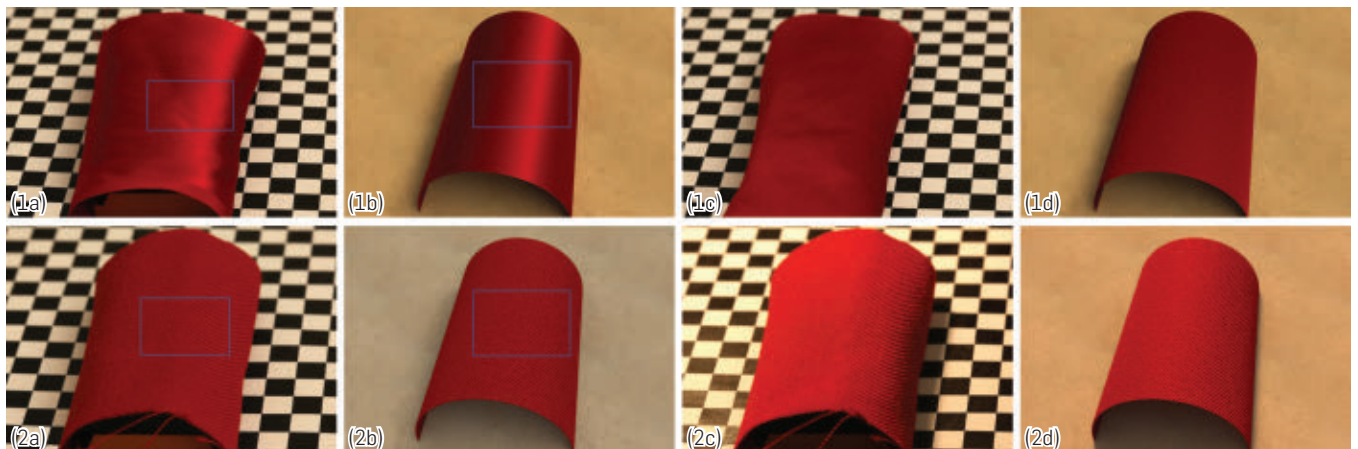
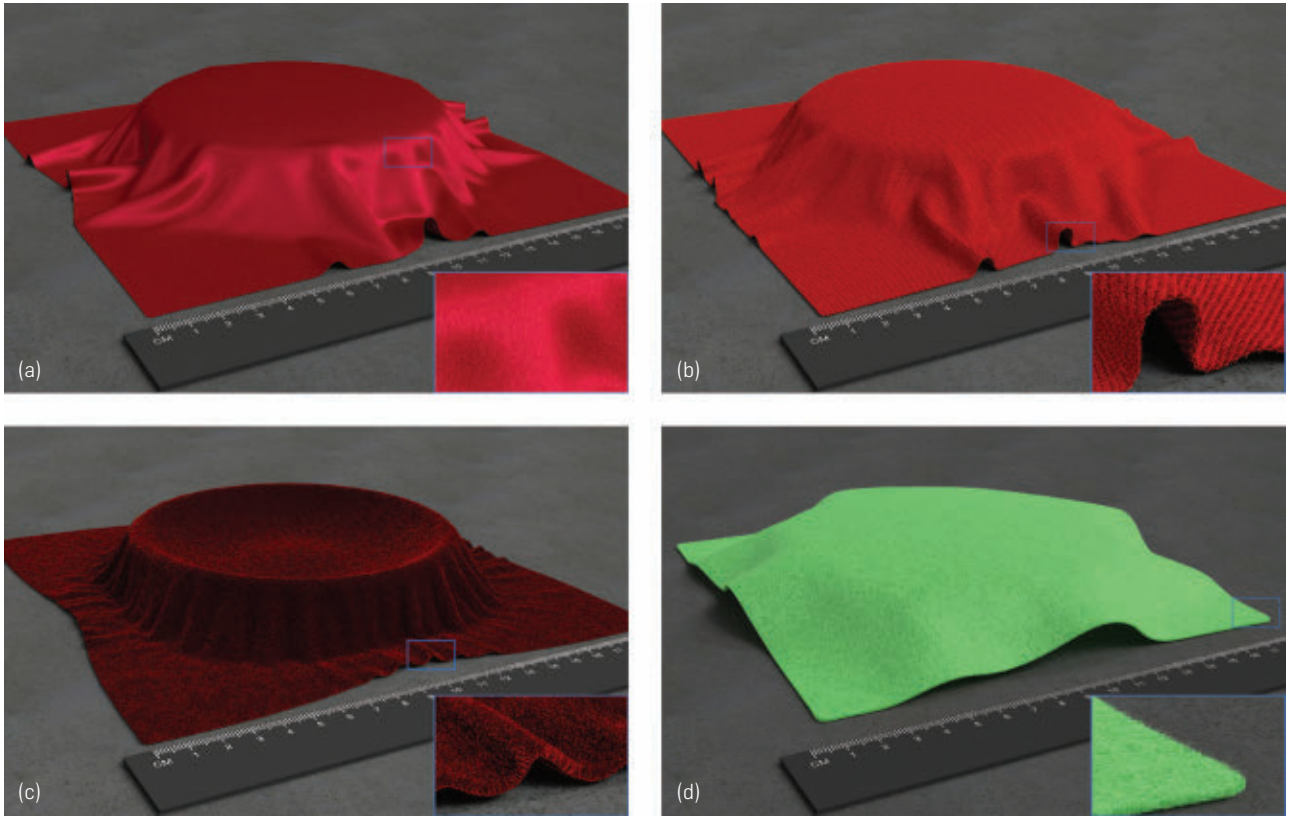


Figure 6. Fabrics in draped configurations with our volumetric appearance model: (a) silk satin, (b) gabardine, (c) velvet, and (d) felt.



with a low-frequency environment map. The validation pair accurately predicts the texture under a different lighting condition, which involves a strong luminaire at the top. In the draped configuration in Figure 6(b), the volume model captures subtle foreshortening effects and the silhouette appearance, as well as the subtle variations in texture across the surface. The appearance at the cut edge gives the proper impression of the thickness of the fabric (compare to the very thin satin material), which is a perennial difficulty with surface models.

Velvet, a material with a cut pile (like a carpet), has a visible surface composed of fibers that stick up from the base material. It has a very distinctive appearance, with a characteristic grazing-angle highlight. The appearance of velvet depends on how the fibers are brushed, and our random tile rotation method produces randomly brushed velvet. In Figure 6(c), we demonstrate how our model reproduces the characteristic velvet highlights. Further, the edges and silhouettes convey the considerable thickness and weight of this material.

Felt is a nonwoven textile consisting of a disorganized layer of matted fibers. The thickness and fuzziness of this material are important appearance attributes that are generally difficult to model and render. Since felt does not exhibit an overall specular highlight, we used a flat patch for appearance matching; because of limited depth of field, we limited the matching region to a thin rectangle where the photograph is in good focus. The illumination conditions for the appearance matching and the validation are the same as those for the gabardine. The color and the contrast due to

self-shadowing attributes are matched nicely and generalize well to the second illumination condition. One limitation for this material is that it has substantial low-frequency content in its texture, which our small sample area did not capture in the CT imaging, leading to a slightly more uniform appearance in our tiled material. Figure 6(d) demonstrates the ability of our volumetric appearance model to capture the material's thick, fuzzy appearance.

A 3D, physically based model also allows more meaningful editing than image-based methods. In Figure 7, we extend the gabardine model with a spatially varying albedo value. The albedo is computed as a function of orientation, so that fibers in the warp and weft are assigned different colors. With blue warp and white weft a fabric similar to denim is produced, though made of wool rather than cotton.

8. CONCLUSION

We have demonstrated a new, multimodal approach to making realistic volume models of cloth that capture both the 3D structure evident in close-up renderings and the reflectance evident in farther-away views. Unlike previous methods for capturing cloth appearance using BTFs, our method explicitly models the 3D structure of the material and, interestingly, is able to capture the directional reflectance of the material automatically because of this structure.

Our modeling approach uses CT imaging where it is strongest, in measuring 3D structure, and it uses photographs where they are strongest, in measuring color and

Figure 7. Renderings obtained by editing the volumetric representation: the gabardine sample is rendered with a blue hue (left); we then detect weft fibers based on their orientation and color them white, which produces a material resembling denim (right).



texture. By matching texture statistics we merge these two sources of information, resulting in a volume model that can produce both close-up views with rich detail of fuzz and fiber structure *and* the characteristic highlights of these materials that emerge naturally from rendering the measured structure. No reflectance measurements are made, and only a few parameters are adjusted in the optical model. The appearance of the cloth is created by a simple anisotropic phase function model together with the occlusion and orientation information extracted from the volume. This paper shows that since geometric structure is what creates the complex appearance of textiles, once we acquire the structure, we are most of the way to modeling the appearance.

Aside from its implications regarding how material appearance can be modeled from structure, this is also quite a practical method for appearance modeling. All that is required to model a material is a CT scan, which can be obtained at reasonable cost from a number of facilities (or in the future from the rapidly improving technology of desktop CT scanning) and a few photographs under known illumination, which takes only a few minutes with a camera and a mirror sphere. In addition, it is possible to CT scan a few samples with elementary weave structures and assemble the resulting volumes to form fabrics with many different designs,¹⁹ further reducing per-design cost. The resulting models are volumetric in nature, and physically based, which makes them easier to edit than image-based data. It is easy to adjust color, glossiness, opacity, and material thickness by scaling parameters of the volume geometry; and a range of more fundamental changes to the material's structure can be made by editing the volume data.

This paper has demonstrated the usefulness of the CT modeling approach for textiles, but the approach does have some limitations. Particularly, it requires that changes in optical properties correlate with changes in density, and this requirement could limit the kinds of materials that can be captured using this imaging modality. Further, the scanner can only image small samples, less than a centimeter across, at the resolution needed to produce clear fiber orientation maps. Thick materials that do not fit fully in the volume (e.g., materials with very long flyaway fibers) cannot be handled well. Some unusual materials, such as metallic fibers, may be problematic for CT because of limited dynamic range. Also, texture content at larger scales will be missed. These

problems will decrease as CT scanners improve in resolution and dynamic range. CT is very well suited to textiles, and it remains to be seen what other materials it performs well for, and how other volume imaging methods work in this technique. Further, materials with differently colored yarns cannot be currently captured by our method.

There are many areas of future work. This work was done using extremely small samples, and with larger samples, which should be possible as CT technology improves and becomes cheaper, better texture could be produced. To extend the range of materials that can be handled, new parameter estimation methods are needed that can identify and fit multiple materials within a single volume. To improve accuracy, more photographs under varying conditions can be used, allowing more parameters (e.g., more complex phase functions) to be fit. Ultimately, this method can be extended to work for a wide range of types of materials whose appearance is difficult to capture using surface models. C

References

- Adabala, N., Magnenat-Thalmann, N., Fei, G. Visualization of woven cloth. In *14th Eurographics Workshop on Rendering* (2003), 180–185.
- Ashikhmin, M., Premoze, S., Shirley, P.S. A microfacet-based BRDF generator. In *Proceedings of ACM SIGGRAPH 2000* (2000), 65–74.
- Chen, Y., an Hua Zhong, S.L., Xu, Y.Q., Guo, B., Shum, H.Y. Realistic rendering and animation of knitwear. *IEEE Trans. Visual. Comput. Graph.* 9, 1 (2003), 43–55.
- Dana, K.J., van Ginneken, B., Nayar, S.K., Koenderink, J.J. Reflectance and texture of real-world surfaces. *ACM Trans. Graph.* 18, 1 (1999), 1–34.
- Furukawa, R., Kawasaki, H., Ikeuchi, K., Sakauchi, M. Appearance based object modeling using texture database: Acquisition, compression and rendering. In *Eurographics Workshop on Rendering* (2002), 257–266.
- Irawan, P., Marschner, S. Specular reflection from woven cloth. *ACM Trans. Graph.* 31, 1 (2012), 11:1–11:20.
- Jakob, W., Arbree, A., Moon, J.T., Bala, K., Marschner, S. A radiative transfer framework for rendering materials with anisotropic structure. *ACM Trans. Graph.* 29, 4 (2010), 53:1–53:13.
- Kajiya, J.T., Kay, T.L. Rendering fur with three dimensional textures. *SIGGRAPH Comput. Graph.* 23, 3 (1989), 271–280.
- Kawabata, S., Niwa, M., Kawai, H. The finite deformation theory of plain weave fabrics. Part I: The biaxial deformation theory. *J. Textile Instit.* 64, 1 (1973), 21–46.
- Lomov, S., Parnas, R., Ghosh, S.B., Verpoest, I., Nakai, A. Experimental and theoretical characterization of the geometry of two-dimensional braided fabrics. *Textile Res. J.* 72, 8 (2002), 706–712.
- Magda, S., Kriegman, D. Reconstruction of volumetric surface textures for real-time rendering. In *Proceedings of the Eurographics Symposium on Rendering (EGSR)* (2006), 19–29.
- Perlin, K., Hoffert, E.M. Hypertexture. *SIGGRAPH Comput. Graph.* 23, 3 (1989), 253–262.
- Pierce, F.T. The geometry of cloth structure. *J. Textile Instit.* 28, 3 (1937), 45–96.
- Porumbescu, S., Budge, B., Feng, L., Joy, K. Shell maps. *ACM Trans. Graph.* 24, 3 (2005), 626–633.
- Shinohara, T., Takayama, J., Ohyama, S., Kobayashi, A. Extraction of yarn positional information from a three-dimensional CT image of textile fabric using yarn tracing with a filament model for structure analysis. *Textile Res. J.* 80, 7 (2010), 623–630.
- Thibault, X., Bloch, J. Structural analysis by X-ray microtomography of a strained nonwoven papermaker felt. *Textile Res. J.* 72, 6 (2002), 480–485.
- Westin, S.H., Arvo, J.R., Torrance, K.E. Predicting reflectance functions from complex surfaces. *SIGGRAPH Comput. Graph.* 26, 2 (1992), 255–264.
- Xu, Y.Q., Chen, Y., Lin, S., Zhong, H., Wu, E., Guo, B., Shum, H.Y. Photorealistic rendering of knitwear using the Lumislice. In *Proceedings of ACM SIGGRAPH 2001* (2001), 391–398.
- Zhao, S., Jakob, W., Marschner, S., Bala, K. Structure-aware synthesis for predictive woven fabric appearance. *ACM Trans. Graph.* 31, 4 (2012), 75:1–75:10.

Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala (szhao, wenzel, srm, kbj@cs.cornell.edu), Cornell University, Ithaca, NY.

CAREERS

Air Force Office of Scientific Research (AFOSR),

**Program Officer (PO), Complex Networks/
Principle Computer Scientist**

The **Air Force Office of Scientific Research (AFOSR)**, Arlington VA, who manages all basic research for the **United States Air Force (USAF)**, has a **Program Officer (PO)** position open in **Complex Networks/Principle Computer Scientist** with a strong background in applied theoretical and experimental foundations of computer and network science, including the representation and transformation of networked information structures and the theoretical models for such representations and transformations. Evaluates scientific opportunities and proposes resource distribution in this area. Evaluates and selects proposals for support and controls expenditures; advocates programs through oral briefings and written reports within the Air Force and to external review panels. Actively coordinates the research program within the Air Force, Department of Defense, Army, Navy, Defense Advanced Research Projects Agency and other national and international organizations. In addition, Joint Service Reliance representative to the Strategic Planning Group coordinating tri-service planning in basic research and conducting a world-wide international program for Air Force Research Laboratory. The present annual salary ranges from \$124,995 to \$152,697, depending upon the individual qualifications and salary history.

Applicant should have at least one year of specialized experience as a principle scientist; AND complex network science experience managing or leading a multi-person program in basic or applied research in those R&D areas. A PhD is highly desired.

Applicants must be U.S. citizens, be able to obtain a security clearance and to travel by military and civil aircraft, and comply with provisions of the Ethics in Government Act.

Apply online:

http://www.teamafrl.afciviliancareers.com/sites/default/files/documents/AFRLS_EProfessionalsAdvancedDegrees_HowToApply.pdf

Cal Poly, San Luis Obispo

Tenure Track Position – Big Data/Data Science

The Computer Science Department at **Cal Poly, San Luis Obispo, CA** is seeking a full-time, academic year, assistant or associate professor with a projected start date 9-14-2015.

For details, qualifications, and application instructions (online faculty application required), visit WWW.CALPOLYJOBS.ORG and apply to **recruitment #103399**. Review Begin Date: 12-01-2014. EEO.

Cal Poly, San Luis Obispo

Tenure Track Position - Cybersecurity

The Computer Science Department at **Cal Poly, San Luis Obispo, CA** is seeking a full-time, academic year, assistant or associate professor with a projected start date 9-14-2015.

For details, qualifications, and application instructions (online faculty application required), visit WWW.CALPOLYJOBS.ORG and apply to **recruitment #103378**. Review Begin Date: 12-01-2014. EEO.

California State University, East Bay (Hayward, CA)

Department of Mathematics and Computer Science

Faculty Position in Computer Science

POSITION (OAA Position No. 15-16 MCS-DATA/CLOUD-TT): The Department invites applications for a tenure-track appointment as Assistant Professor in Computer Science (preference to emerging areas of computer science such as big data, cloud computing, mobile app development, data mining) starting Fall 2015.

Teaching includes day and evening courses at B.S. and M.S. levels, with a typical teaching responsibility of three classes/Quarter. Required: potential for excellent teaching, research and curriculum development; ability to teach, advise and mentor students from diverse backgrounds; potential to serve Department and University. Applicants must have a Ph.D. by September 2015.

Please submit an application letter and a current and complete vita through the following URL: https://my.csueastbay.edu/psp/pspdb1/EMPLOYEE/HRMS/c/HRS_HRAM.HRS_CE.GBL

Additionally, please email graduate transcripts, 3 letters of recommendation, 3 references, a statement of teaching philosophy, and evidence of teaching and research abilities to the Computer Science Search Committee at the email address below.

A detailed position announcement is available at:

<http://www20.csueastbay.edu/about/career-opportunities/>

For questions, email:

CSSearch@mcs.CSUEastBay.edu

APPLICATION DEADLINE The deadline for applications is October 31, 2014; review of applications will begin November 1, 2014. CSUEB, situated in the hills overlooking San Francisco Bay, is an EOE, committed to "educational excellence for a diverse society". If you are considered as a finalist for the position, you may be subject to a background check.

California State University, Monterey Bay

**Computer Science and Information Technology
Tenure-Track Faculty**

California State University, Monterey Bay (CSUMB) is searching for three tenure-track faculty; two in computer science and one in computer engineering. CSUMB's computer science and information technology bachelor's degree is growing rapidly, with over \$5M in extramural funding to support innovative education initiatives, research and development projects, student scholarships and more. To support this growth and launch a new computer engineering degree to meet growing student demand, CSUMB is adding these three tenure-track lines. Please visit <http://itcd.csUMB.edu/jobs> to learn about these positions and the program. Apply URL: <http://hr.csUMB.edu/jobs/db/fac/>

Caltech

Tenure-track faculty position

The Computing and Mathematical Sciences (CMS) Department at Caltech invites applications for a tenure-track faculty position. Our department is a unique environment where innovative, interdisciplinary, and foundational research is conducted in a collegial atmosphere. We are looking for candidates who have demonstrated exceptional promise through novel research with strong potential connections to natural, information, and engineering sciences. Research areas of particular interest include applied mathematics, computational science, as well as computing. A commitment to high-quality teaching and mentoring is expected.

The initial appointment at the assistant-professor level is for four years and is contingent upon the completion of a Ph.D. degree in Applied Mathematics, Computer Science, or related field. Exceptionally well-qualified applicants may also be considered at the full professor level.

To ensure the fullest consideration, applicants are encouraged to have all their application materials on file by December 28, 2014. For a list of documents required and full instructions on how to apply on-line, please visit <http://www.cms.caltech.edu/search>. Questions about the application process may be directed to: search@cms.caltech.edu.

Caltech is an Equal Opportunity/Affirmative Action Employer. Women, minorities, veterans, and disabled persons are encouraged to apply.

Carleton College

Tenure-Track Position in Computer Science

Carleton College invites applications for a **tenure-track position in computer science, in any area of specialization**, beginning September 1, 2015. A Ph.D. or its imminent completion is required. We

expect to make an appointment at the assistant professor level, but highly qualified candidates at higher ranks will be considered.

Carleton is a highly selective liberal arts college with outstanding, enthusiastic students. We seek an equally enthusiastic computer scientist committed to excellence in teaching, curriculum design, ongoing research, and undergraduate research advising. We are particularly interested in candidates committed to teaching a diverse student body. The appointee would be expected to teach courses at a range of levels, from the introductory level to advanced undergraduate electives.

Carleton's Department of Computer Science (cs.carleton.edu) includes eight faculty and an experienced full-time system administrator, and typically graduates between 25 and 60 majors annually.

To apply, please visit the Carleton College website at jobs.carleton.edu and complete the online application, including a cover letter, CV, brief statements describing teaching philosophy and research agenda, a graduate transcript, and contact information for writers of three letters of reference (at least one of which should address your teaching). Questions about the position or the application process may be directed to csjobs@carleton.edu.

Applications completed by December 15, 2014, including reference letters, will receive full consideration.

Carleton is committed to developing its faculty to better reflect the diversity of our student body and American society. Women and members of minority groups are strongly encouraged to apply. Carleton College does not discriminate on the basis of race, color, creed, ethnicity, reli-

gion, sex, national origin, marital status, veteran status, actual or perceived sexual orientation, gender identity and expression, status with regard to public assistance, disability, or age in providing employment or access to its educational facilities and activities.

Carnegie Mellon University Computer Science Department Teaching Track Position

Applications are invited for a teaching-track position in Computer Science, beginning Fall 2015. This is a renewable, career-oriented position with an initial appointment for three years. We seek highly qualified applicants with a strong commitment to excellence in teaching and the ability to teach at all levels in the undergraduate curriculum. Preference will be given to candidates with a strong systems background.

Applicants for the position must have a Ph.D. in Computer Science or a related field, and demonstrated excellence in teaching Computer Science courses. Teaching-track appointments are typically at the rank of Assistant Teaching Professor, with the possibility of promotion to the ranks of Associate Teaching Professor and Teaching Professor. None of these ranks are tenured; applicants seeking a tenure-track position at a research university are therefore not a good match for these positions.

In order to receive full consideration, applicants should submit a letter of application, curriculum vitae, a statement of teaching phi-

losophy, and the names and email addresses of three or more individuals whom the applicant has asked to provide letters of reference. Applicants should arrange for reference letters to be sent directly to the contact below. This information should be sent by December 31, 2014, to the contact listed below.

Additionally, applicants are encouraged to submit a video sample of their teaching, demonstrating their excellence.

Please send your applications and accompanying materials to

Dr. Klaus Sutner
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
email: sutner@cs.cmu.edu

Carnegie Mellon is an affirmative action/equal opportunity employer and we invite and encourage applications from women and under-represented minorities.

Fordham University Department of Computer and Information Science, Open Position: Department Chair (Full Professor)

Fordham University invites applications for the position of Chair in the Department of Computer and Information Science. The department is currently in a process of growth with administrative



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Faculty Positions in Computer Science at Ecole polytechnique fédérale de Lausanne (EPFL)

The School of Computer and Communication Sciences at EPFL invites applications for faculty positions in computer and communication sciences. We are seeking candidates for tenure-track assistant professor as well as for senior positions.

Successful candidates will develop an independent and creative research program, participate in both undergraduate and graduate teaching, and supervise PhD students.

Candidates from all areas of computer science and communication systems will be considered, but preference will be given to candidates in the fields of data analytics, machine learning, security and privacy, and visual computing.

EPFL offers internationally competitive salaries, significant start-up resources, and outstanding research infrastructure. To apply, please follow the application procedure at

<https://academicjobsonline.org/ajo/jobs/4591>

The following documents are requested in PDF format: curriculum vitae including publication list, brief statements of research and teaching interests, names and addresses (including e-mail) of 3 references for junior positions and 6 for senior positions.

Screening will start on **December 1, 2014**.

Further questions can be addressed to:

Prof. Willy Zwaenepoel
Chairman of the recruiting committee
recruiting.ic@epfl.ch

For additional information on EPFL, please consult:
<http://www.epfl.ch> or <http://ic.epfl.ch>

EPFL is an equal opportunity employer.

support for the expansion of undergraduate and graduate programs. We are seeking a dynamic and innovative leader who will guide the department to greater prominence within the University and nationally. Candidates must have a Ph.D. in computer science, information science, or a related area and an active research program. The position requires distinguished research, significant administrative experience, and strengths in program development, collaborative academic leadership, and teaching. Outstanding candidates in all areas of computer and information science are strongly encouraged to apply. The department is situated on Fordham's campuses in the center of Manhattan, the Bronx, and Westchester, where faculty and students take advantage of New York City's vast resources for research, education, and service. For complete position descriptions, see <http://www.cis.fordham.edu/openings.html>.

The department offers undergraduate majors in computer science, information science, and interdisciplinary programs in Fordham's undergraduate colleges as well as MS programs in Computer Science and in Cybersecurity. The department is currently developing new undergraduate and graduate programs in computer science, information science, and interdisciplinary areas. Fordham's administration is committed to increasing the department's local and national distinction in research and science education. Additional new members will join the faculty in the coming years to build upon the faculty's wide range of expertise and to serve new programs. With administrative support and new resources, the successful applicant will lead the Department's trajectory toward

becoming exemplary within Fordham, in New York City, and nationally as a center of education, research, and professional engagement.

An application can be submitted electronically to Interfolio Scholar Services: apply.interfolio.com/25960 with a cover letter; brief statements of administrative philosophy, teaching pedagogy, and research agenda; c.v.; samples of scholarship; and names with contact information of three references who may be contacted for letters of recommendations. Application review begins on November 15 and will continue until the position is filled. For information, contact Dr. David Wei (dsl.wei01@gmail.com; wei@cis.fordham.edu).

Fordham, an independent, Catholic University in the Jesuit tradition, is committed to excellence through diversity and welcomes candidates of all backgrounds. Fordham is an Equal Opportunity Employer.

Georgia Institute of Technology Tenure-Track Faculty Positions at the School of Computer Science

The School of Computer Science at the Georgia Institute of Technology invites applications for tenure-track faculty positions from all areas of computer science. We seek candidates that can add to or enhance the current research areas of the school, and we put more emphasis on excellence than on candidates' specific area of expertise. In order to receive full consideration, please submit your application online by December 8, 2014. Full details at: <http://www.scs.gatech.edu/about/facultypositions>.

Georgia Tech is an Affirmative Action/Equal Opportunity Employer. Applications from women and under-represented minorities are strongly encouraged.

Georgia State University Department of Computer Science Lecturer Position

The Department of Computer Science of Georgia State University invites applications for a lecturer position. Position will begin in January 2015 and an earned Ph.D. in Computer Science is required (candidates who are expected to receive their Ph.D. by Summer of 2015 will be considered). Preference will be given to candidates with prior experience in teaching undergraduate computer science courses.

The lecturer will be responsible for teaching a wide variety of undergraduate courses, including principles of computer science and programming, theoretical foundations of computer science, data structures, systems-level programming, and computer organization and assembly language. In addition, the lecturer will participate in advising, courses development, and lab supervision.

Georgia State University (GSU) is an urban public research university serving 32,000 students and is centered in the historic financial hub of downtown Atlanta. The Computer Science Department offers programs leading to the B.S., M.S., and Ph.D. degrees in computer science.

Women and minorities are especially encouraged to apply. The application should consist of a

ETH zürich

Assistant Professorships (Tenure Track) in Computer Science

→ The Department of Computer Science (www.inf.ethz.ch) at ETH Zurich invites applications for Assistant Professorships (tenure track) with focus on the following areas:

- Software Engineering and Programming Languages
- Information Systems (with emphasis on Big Data)
- Theory

→ For candidates with exceptional research accomplishments also applications for a tenured full professorship will be considered.

→ Applicants should have internationally recognized expertise in their field and pursue research at the forefront of Computer Science. Successful candidates should establish and lead a strong research program. They will be expected to supervise PhD students and teach both undergraduate level courses (in German or English) and graduate level courses (in English). Collaboration in research and teaching is expected both within the department and with other groups of ETH Zurich and related institutions.

Assistant professorships have been established to promote the careers of younger scientists. The initial appointment is for four years with the possibility of renewal and promotion to a permanent position.

→ Please apply online at www.facultyaffairs.ethz.ch

→ Your application should include your curriculum vitae, a list of publications, a statement of research and teaching interests and the names of at least three referees. The letter of application should be addressed to the **President of ETH Zurich, Prof. Dr. Ralph Eichler**. The closing date for applications is **15 December 2014**. ETH Zurich is an equal opportunity and family friendly employer and responsive to the need of dual career couples. We specifically encourage women to apply.

cover letter, CV, a brief statement describing your teaching interests and philosophy, and a list of references. Letters of reference will be requested from the candidates who are invited for a campus interview. Electronic submissions in PDF format are required and should be sent to search@cs.gsu.edu.

Georgia State University, a Research University of the University System of Georgia, is an AA/EEO employer.

Grand View University **Assistant, Associate, or** **Full Professor of Computer Science**

We have a full-time tenure-track opening for Assistant, Associate, or Full Professor of Computer Science beginning Fall 2015. For details on this position and to apply, please visit our website: www.grandview.edu. Apply URL: <http://grandview.hrmdirect.com/employment/job-opening.php?req=203850&&nohd#job>

Indiana University - Bloomington **School of Informatics and Computing** **Chair of Computer Science** **Faculty Positions in Computer Science** **and Informatics**

The School of Informatics and Computing (SoIC) at Indiana University Bloomington invites applications for chair of computer science, a full professor with a 3-year initial term as chair, and

for two asst/assoc/full professors, in complex networks and systems (CX) and in interactive and intelligent systems (IIS).

Chair applicants should have a world-class research record, a strategic vision for excellence in Computer Science research and education, and the academic and research leadership skills to advance that vision. IU Computer Science is strong and rapidly growing with 39 tenure-line faculty, many hired in the last 5 years.

CX applications are especially encouraged from established leaders, who will have opportunities for leadership roles in the Center for Complex Networks and Systems and in a new and ambitious university-wide network science initiative to be announced.

IIS applications are welcome from a range of related areas such as computer vision, smart imaging, pervasive computing, social robotics, and internet of things.

Applicants should have an established record (senior level) or demonstrable potential for excellence (junior) in research and teaching, and a Ph.D. in a relevant area, or (junior) expected by 8/15.

The SoIC is the first of its kind and among the largest in the country, with unsurpassed breadth. Its mission is to excel and lead in education, research, and outreach spanning and integrating the full breadth of computing and information technology. It includes Computer Science, Informatics, and Library and Information Science, with over 85 tenure-line faculty, 900 graduate students, and 1100 undergraduate majors. It offers Ph.D.s in Computer Science, Informatics, and Information Science.

Bloomington is a culturally thriving college town with moderate cost of living and amenities for an active lifestyle. IU is renowned for its top-ranked music school, HPC and networking facilities, and performing and fine arts.

Applicants should submit a CV, a statement of research and teaching, and names of 6 references (3 for junior level) using links below (preferred) or to Faculty Search, SoIC, 919 E 10th St, Bloomington, IN 47408. Chair candidates should also submit a statement of leadership philosophy and vision for IU CS. Questions may be sent to hir-ing@soic.indiana.edu. For full consideration applications are due by 12/1/14.

<http://indiana.peopleadmin.com/postings/968> (Chair, CS)

<http://indiana.peopleadmin.com/postings/969> (CX)

<http://indiana.peopleadmin.com/postings/970> (IIS)

Indiana University is an equal employment and affirmative action employer and a provider of ADA services. All qualified applicants will receive consideration for employment without regard to age, ethnicity, color, race, religion, sex, sexual orientation or identity, national origin, disability status or protected veteran status.

Indiana University – Purdue University **Fort Wayne, Indiana** **Department of Computer Science** **Assistant Professor of Computer Science**

The Department of Computer Science invites applications for Assistant Professor (Tenure-Track) beginning August 17, 2015. Outstanding candidates must have strong expertise in Software Engineering. Candidates specializing in Mobile and Embedded Systems, Cyber Security, Cloud Computing, Graphics & Visualization, or Bioinformatics in addition to some Software Engineering experiences will be considered.

Requirements include a Ph.D. in Computer Science or closely related field; and exceptional record of research that supports the teaching and research missions of the Department; excellent communication and interpersonal skills; and an interest in working with students and collaborating with the business community.

Required duties include teaching undergraduate and graduate level courses in Computer Science, academic advising, and a strong pursuit of scholarly endeavors. Service to and engagement with the University, Department, and community is also required.

Review of applications will begin November 1, 2014 and will continue until the position is filled. Please submit a letter of application addressing qualifications, and your teaching and research statements. For teaching, we require an example of your teaching experience or effectiveness and a 1-2 page teaching philosophy. Also required for submission are your CV, with the names and contact information of three references. All materials must be submitted electronically in .pdf format to Lisa Davenport, Secretary for the Department of Computer Science at davenport@ipfw.edu. All candidates who are interviewed should prepare a 45-60 minute instructional student presentation.

Please see an extended ad at our website IPFW.edu/cs, click on "Jobs" in the sidebar. IPFW



THE CHINESE UNIVERSITY OF HONG KONG

Applications are invited for:-

Faculty of Engineering **Professors / Associate Professors / Assistant Professors** (Ref. 1415/039(255)/2)

The Faculty invites applications for several faculty posts at Professor / Associate Professor / Assistant Professor levels, with prospect for substantiation, in the interdisciplinary area of 'Big Data Analytics', which is a new strategic research initiative supported by the University's Focused Innovations Scheme and will complement current/planned strengths in different Departments under the Faculty. To lead the big data research initiative, senior professors in this area are particularly welcome. Currently, the Faculty is seeking candidates in the following areas:

- Theoretical, mathematical and algorithmic aspects in large data analytics;
- Large scale software systems and architecture in large data analytics;
- Combining big data analytics with statistical modelling and operations research methods for optimal decision making;
- Application areas in large data analytics (including information systems and the Web, bioinformatics, financial engineering, logistics and supply chain management, public health, social networks, etc.)

Applicants should have (i) a PhD degree; and (ii) a strong scholarly record demonstrating potential for teaching and research excellence. The appointees will be expected to (a) teach undergraduate and postgraduate courses; (b) develop a significant independent research programme with external funding; and (c) supervise postgraduate students. Appointments will normally be made on contract basis for three years initially, which, subject to performance and mutual agreement, may lead to longer-term appointment or substantiation later. Applications will be accepted until the posts are filled. Further information about the Faculty is available at <http://www.erg.cuhk.edu.hk>.

Salary and Fringe Benefits

Salary will be highly competitive, commensurate with qualifications and experience. The University offers a comprehensive fringe benefit package, including medical care, plus a contract-end gratuity for appointments of two years or longer, and housing benefits for eligible appointees. Further information about the University and the general terms of service for appointments is available at <http://www.per.cuhk.edu.hk>. The terms mentioned herein are for reference only and are subject to revision by the University.

Application Procedure

Please send full resume, copies of academic credentials, a publication list with abstracts of selected published papers, details of courses taught and evaluation results (if any), a research plan, a teaching statement, together with names of three to five referees, to the Dean, Faculty of Engineering by e-mail to recruit-bda@erg.cuhk.edu.hk. For enquiries, please contact Professor John C.S. Lui, the leader of the strategic initiative (e-mail: cslui@cse.cuhk.edu.hk). Applicants are requested to clearly indicate that they are applying for the posts under 'Big Data Analytics Initiative'. The Personal Information Collection Statement will be provided upon request. Please quote the reference number and mark 'Application - Confidential' on cover.

is an Equal Opportunity/Equal Access/Affirmative Action University fully committed to achieving a diverse workforce.

The Johns Hopkins University Department of Computer Science Faculty Search

With the opening of our new building, Malone Hall, the **Department of Computer Science at The Johns Hopkins University** is planning to continue a substantial multi-year growth that has already added 11 new faculty members to the department since 2010. Our faculty search this year is open to both tenure track and teaching track positions at all levels and from all areas of Computer Science. For tenure track positions, applicants in systems, networking, and related areas are strongly desired. We also encourage applicants with an interest in systems and network security, cloud and mobile security, and health-related applications security to apply. For teaching track positions, applicants with the ability to teach courses related to software engineering at both undergraduate and graduate levels are particularly encouraged.

All applicants must have a Ph.D. in Computer Science or a related field. Tenure track faculty will be expected to establish a strong, independent, multidisciplinary, internationally recognized research program. Commitment to quality teaching at the undergraduate and graduate levels is required. The department is committed to building a diverse educational environment; women and minorities are especially encouraged to apply.

A more extensive description of our search and additional supporting information can be found at <http://www.cs.jhu.edu/Search2015>. More information on the department is available at <http://www.cs.jhu.edu>.

Applicants should apply using the online applications which can be accessed from <http://www.cs.jhu.edu/apply>. Applications should be received by December 15, 2014 for full consideration. Questions should be directed to fsearch@cs.jhu.edu. The Johns Hopkins University is an EEO/AA employer.

Faculty Search
Johns Hopkins University
Department of Computer Science
Malone Hall Rm 160
Baltimore, MD 21218-2694
Fax: 410-516-6134
Phone: 410-516-8775
fsearch@cs.jhu.edu
<http://www.cs.jhu.edu/apply>

Kansas State University Faculty Position Department of Computing and Information Sciences

The department of Computing and Information Sciences (CIS) at Kansas State University invites applications for multiple tenure-track positions beginning in Fall 2015 at all levels (Assistant, Associate, or Full Professor). Applicants should have a background in cybersecurity or high-assurance software. In addition, applicants whose research interfaces with one or more of the following areas are highly encouraged to apply: bio-security, cyber-physical systems, mobile computing, cloud-

based systems, big data, or human-computer interfaces. Successful applicants will demonstrate commitment to both teaching and research. A doctorate degree in computer science or related disciplines is required; salary will be commensurate with qualifications.

Kansas State University is committed to the growth and excellence of the department of Computing and Information Sciences. CIS department offers a stimulating environment for research and teaching, and has several ongoing collaborative projects involving researchers in different areas of computer science as well as physics, psychology, anthropology, and other engineering and science departments. The department offers BS, MS, and Ph.D. degrees, has a faculty of 19, and over 100 graduate and 400 undergraduate students. The department offers modern research and teaching facilities, including Beocat, a Beowulf cluster with over 2500 cores. The department has modern media-enabled classrooms and several laboratories for cybersecurity, high-assurance systems, software analysis, embedded systems, adaptive systems, computational engineering and science, and data-mining. In Fall 2015, CIS will move to a new state-of-the-art facility in the College of Engineering complex (<http://www.engg.ksu.edu/phaseiv/>). Kansas State is also home to the Biosecurity Research Institute and is the site of the new National Bio and Agro-Defense Facility. Details of the CIS Department can be found at the URL <http://www.cis.ksu.edu/>.

Please send applications to Scott DeLoach, Recruitment Chair, Department of Computing and Information Sciences, 234 Nichols Hall, Kan-

sas State University, Manhattan, KS 66506 (email: recruiting@cis.ksu.edu). Applicants should submit a letter of interest, curriculum vitae, research and teaching statements, copies of representative publications and contact information of at least three references. Review of applications will commence December 01, 2014 and continue until the positions are filled.

Kansas State University actively seeks diversity among its employees. Kansas State University is an EOE of individuals with disabilities and protected veterans. Background check required.

Lehigh University Teaching Faculty Position in Computer Science

Applications are invited for a full-time, non-tenure track, Teaching Faculty position (Professor of Practice) in the Computer Science and Engineering Department (<http://www.cse.lehigh.edu>) of Lehigh University, to start in January or August 2015.

The successful applicant should hold a graduate degree in Computer Science, Computer Engineering, or a closely related field. The candidate must demonstrate a strong commitment to quality and innovation in undergraduate education. Applicants should be comfortable teaching core courses in computer science as well as elective courses in their areas of interest.

The faculty of the Computer Science and Engineering department maintains an outstanding international reputation in a variety of research areas, and includes ACM and IEEE fellows as well as six NSF CAREER award winners. The depart-



University at Buffalo *The State University of New York*

STATE UNIVERSITY OF NEW YORK AT BUFFALO Department of Computer Science and Engineering

The State University of New York at Buffalo Department of Computer Science and Engineering invites candidates to apply for **multiple tenured and tenure-track faculty positions** beginning in the 2015-2016 academic year. We invite candidates from all areas of computer science and engineering, especially software engineering and programming languages, cyber-security, and computer vision and pattern recognition. Applicants should have a Ph.D. in computer science or a related area and demonstrate potential for excellence in research, teaching, and mentoring.

Applications should be submitted by **December 31, 2014**, electronically via <http://www.ubjobs.buffalo.edu/>. Please use posting number 1400598 to apply. The University at Buffalo is an Equal Opportunity Employer.

Computer Science and Engineering Department

The Department of Computer Science and Engineering has 34 tenured and tenure-track faculty and four teaching faculty. We have hired fifteen faculty members in the last five years and continuing to expand rapidly. Eight members of our faculty are NSF CAREER award recipients. Our faculty members are actively involved in successful interdisciplinary programs and centers devoted to biometrics; bioinformatics; biomedical computing; cognitive science; document analysis and recognition; high performance computing; information assurance and cyber security; and computational and data science and engineering.

The State University of New York at Buffalo (UB)

UB 2020 (<http://www.buffalo.edu/UB2020>) is a strategic development plan initiated by University President Satish Tripathi for enhancing educational opportunities for our students, advancing research discoveries that improve life for people throughout the world, and increasing our university's economic impact on our region and New York State. To achieve these ambitious goals, the University at Buffalo will hire over 250 faculty during the next five years. UB 2020 is part of NYSUNY 2020 plan initiated by Governor Andrew Cuomo.

Buffalo

Buffalo is the second largest city in New York state, and was rated the 10th best place to raise a family in America by Forbes magazine in 2010 due to its short commutes and affordability. Located in scenic Western New York, Buffalo is located near world-famous Niagara Falls, the Finger Lakes, and the Niagara Wine Trail. The city is renowned for its architecture and features excellent museums, dining, cultural attractions, and several professional sports teams.

ment currently has 15 full-time faculty members, and more than 300 undergraduate and 60 graduate students. The department offers B.A., B.S., M.S., and Ph.D. degrees in Computer Science and jointly oversees B.S., M.S., and Ph.D. degree programs in Computer Engineering with the department of Electrical and Computer Engineering. The department also offers a B.S. in Computer Science and Business in cooperation with the College of Business and Economics.

Lehigh University is a private, highly selective institution that is consistently ranked among the top 40 national research universities by U.S. News & World Report. Located in Bethlehem, Pennsylvania, Lehigh is 80 miles west of New York City and 50 miles north of Philadelphia, providing an accessible and convenient location that offers an appealing mix of urban and rural lifestyles.

Applications can be submitted online at <https://academicjobsonline.org/ajo/jobs/4559> and should include a cover letter, curriculum vita, teaching statement, and contact information for at least three references. Review of applications will begin November 1, 2014 and will continue until the position is filled.

Lehigh University is an affirmative action/equal opportunity employer and does not discriminate on the basis of age, color, disability, gender, gender identity, genetic information, marital status, national or ethnic origin, race, religion, sexual orientation, or veteran status. Lehigh University is a 2010 recipient of an NSF ADVANCE Institutional Transformation Grant for promoting the careers of women in academic science and engineering. Lehigh University provides

comprehensive benefits including domestic partner benefits (see also <http://www.lehigh.edu/worklifebalance/>). Lehigh Valley Inter-regional Networking & Connecting (LINC) is a newly created regional network of diverse organizations designed to assist new hires with dual career, community and cultural transition needs. Please contact infdcap@lehigh.edu for more information. Questions concerning this search may be sent to pop-search@cse.lehigh.edu.

Marist College Assistant/Associate Professor, Computing Technology

The Marist College Department of Computing Technology is seeking applications for three tenure-track positions in Information Technology & Systems and Computer Science.

We will consider only candidates who have teaching and research interests in the following areas: cloud computing, networking, security, business intelligence, software development, systems programming, data management, and business analytics. Applicants with the ability to teach in multiple areas across several disciplines will receive preference. Applicants should be willing to teach graduate and undergraduate courses in both regular settings (in a classroom) and online environments. Other required duties include scholarly activities that result in peer-reviewed publications as well as engagement in college and professional services including advising and mentoring students as well as serving on depart-

ment, school, and college committees. We are also looking for a candidate with an outstanding record in research and teaching for a senior level position holding the William Cadden Chair.

Candidates must have a doctoral degree in Computer Science, Information Systems, or a closely related field. ABDs in an appropriate field who will complete their dissertation within one year will be considered. Evidence of excellence in teaching and scholarly work is required. Excellent written and oral communication skills are required. Industry and/or consulting experience is highly desirable. As our programs host a diverse population, the proven ability to work effectively in a multicultural environment is highly regarded.

Marist College is an independent and comprehensive liberal arts institution located in New York's historic Hudson River Valley. Situated on 210 acres overlooking the Hudson River, it enrolls 4,787 traditional undergraduate, 799 full and part-time graduate and 543 continuing education students. Marist also has a branch campus in Florence, Italy, and extension sites throughout New York. Marist has been recognized for excellence by U.S. News & World Report, TIME Magazine, The Princeton Review's The Best 376 Colleges, and Barron's Best Buys in College Education and is noted for being a pioneer in the area of online degree programs.

To learn more or to apply, please visit <http://jobs.marist.edu>. Only online applications are accepted. The review of applications will begin immediately and will continue until the position is filled.

Marist College is strongly committed to the principle of diversity and is especially interested in receiving applications from members of ethnic and racial minority groups, women, persons with disabilities, and persons from other under-represented groups.

AN EQUAL OPPORTUNITY/AFFIRMATIVE ACTION EMPLOYER

Max Planck Institute for Software Systems Tenure Track Openings

Applications are invited for tenure-track faculty positions in all areas related to the theory and practice of software systems. These areas include, but are not limited to, security and privacy, embedded and mobile systems, social computing, data science, programming languages and systems, software verification and analysis, parallel, distributed, and networked systems.

A doctoral degree in computer science or related areas and an outstanding research record are required. Successful candidates are expected to build a team and pursue a highly visible research agenda, both independently and in collaboration with other groups.

MPI-SWS, founded in 2005, is part of a network of over 80 Max Planck Institutes, Germany's premier basic research facilities. MPIs have an established record of world-class, foundational research in the fields of medicine, biology, chemistry, physics, technology and humanities. Since 1948, MPI researchers have won 17 Nobel prizes. MPI-SWS aspires to meet the highest standards of excellence and international recognition with its research in software systems.

To this end, the institute offers a unique environment that combines the best aspects of a



ShanghaiTech University

ShanghaiTech University invites highly qualified candidates to fill multiple tenure-track/tenured faculty positions as its core team in the School of Information Science and Technology (SIST). Candidates should have exceptional academic records by international standards or demonstrate strong potential in cutting-edge research areas of information science and technology. English fluency is required and overseas academic experience is highly desired.

ShanghaiTech aims to become a world-class research university for training future scientists, entrepreneurs, and technological leaders. Located in Zhangjiang High-Tech Park in the cosmopolitan Shanghai, we shall trail-blaze a new education system in China. Besides establishing and maintaining a world-class research profile, faculty candidates must also contribute substantially to graduate and undergraduate education.

Academic Disciplines: We welcome candidates in all cutting edge areas of information science and technology. Our recruitment focus includes, but is not limited to: computer architecture and software, cloud and high performance computing, computational foundations, data mining and analysis, visualization, computer vision, machine learning, data sciences and statistics, IC designs, solid-state electronics, high speed and RF circuits, embedded systems, intelligent and signal processing systems, smart energy/power devices and systems, next-generation networking, control systems, robotics, sensor networks as well as inter-disciplinary areas involving information science and technology.

Compensation and Benefits: Salary and startup funds are highly competitive, commensurate with experience and academic accomplishment. We also offer a comprehensive benefit package to employees and eligible dependents, including housing benefits. All regular faculty members will join our new tenure-track system commensurate with international practice for tenure evaluation and promotions.

Qualifications:

- A well articulated research plan and demonstrated record/potentials;
- Ph.D. (Electrical Engineering, Computer Engineering, Computer Science, Statistics, or related field);
- A minimum relevant research experience of 4 years.

Applications: Submit (in English, PDF) a cover letter, a 2-page research plan, a CV plus copies of 3 most significant publications, and names of three referees to: sist@shanghaitech.edu.cn by **December 31st, 2014** (or until positions are filled). More information is at <http://www.shanghaitech.edu.cn>.

university department and a research laboratory:

a) Faculty enjoy academic freedom, receive institutional funding and attract additional third-party funds to build and lead a team of graduate students and post-docs.

b) Faculty supervise doctoral theses, and have the opportunity to teach graduate and undergraduate courses.

c) Faculty enjoy outstanding technical and administrative support, as well as internationally competitive compensation.

MPI-SWS currently has 10 tenured and tenure-track faculty and about 50 doctoral and post-doctoral researchers. The institute will grow to at least 17 faculty over the coming years. We maintain an open, international and diverse work environment and seek applications from outstanding researchers regardless of national origin or citizenship. The working language is English; knowledge of the German language is not required for a successful career at the institute.

The institute is located in Kaiserslautern and Saarbrücken, in the tri-border area of Germany, France and Luxembourg. In immediate proximity are the MPI for Informatics, Saarland University, the Technical University of Kaiserslautern, the German Center for Artificial Intelligence (DFKI), and the Fraunhofer Institutes for Experimental Software Engineering and for Industrial Mathematics.

Qualified candidates should apply online via the secure application form. The review of applications will begin on December 1, 2014, and applicants are strongly encouraged to apply by that date; however, applications will continue to be accepted through December 2014.

The institute is committed to increasing the representation of minorities, women and individuals with physical disabilities in Computer Science. We particularly encourage such individuals to apply.

The initial tenure-track appointment is for five years; it can be extended to seven years based on a midterm evaluation in the fourth year. A permanent contract can be awarded after a successful tenure evaluation in the sixth year.

Apply URL:

<http://www.mpi-sws.org/index.php?n=careers/tenure-track>

Milwaukee School of Engineering Software Engineering Faculty

MSOE invites applications for a full-time faculty position in our Software Engineering program beginning in the fall of 2015. Rank will depend on qualifications and experience of the candidate.

Applicants must have an earned doctorate degree in software engineering, computer science, computer engineering, or closely related field by the beginning date of appointment. This full-time position is open to applicants in all areas of software engineering.

MSOE expects and rewards a strong primary commitment to excellence in teaching at the undergraduate level. Excellent written and oral communication skills are required. Leadership in the growth of undergraduate engineering education and continued professional development are expected.

Our ABET-accredited undergraduate software engineering program had its first graduates in

spring 2002, among the first in the US. Founded in 1903, MSOE is a private, application-oriented university with programs in engineering, business, and nursing. The review of applications will begin as they are received and continue until the position is filled.

Please visit our website at www.msos.edu in order to apply. EEO Employer F/M/Vet/Disabled

Mississippi State University Faculty Positions in Computer Science and Engineering

The Department of Computer Science and Engineering (<http://www.cse.msstate.edu>) is seeking to fill **two open positions** for tenure-track faculty members at the **Assistant/Associate Professor levels**. Evidence of strong potential for excellence in research (including the ability to attract external funding) and teaching at the graduate and undergraduate levels is required. The department has six core research areas: Artificial Intelligence, Computational Science, Graphics and Visualization, Human-Centered Computing, Software Engineering, and Systems. The primary areas of interest for these positions are software engineering and computer security, however exceptional candidates in all areas will be considered.

Mississippi State University is the largest university in the State of Mississippi with approximately 1300 faculty and 20,000 students. The Department of Computer Science and Engineering has 16 tenure-track faculty positions and offers academic programs leading to the bachelor's, mas-

ter's and doctoral degrees in computer science and bachelor's degrees in software engineering and computer engineering. Faculty members and graduate students work with a number of on-campus research centers. Department research expenditures totaled approximately \$10M in FY12.

Candidates for this position are expected to hold a PhD in computer science or closely related field (ABDs may be considered). Level of appointment is commensurate with qualifications and experience.

Applicants must apply on-line <http://www.jobs.msstate.edu/> and complete a Personal Data Information Form (Administrators/Faculty). A letter of application, curriculum vita, teaching statement, research statement, and names and contact information of at least three references must also be submitted. Review of applications will begin as early as November 2014 and continue until the position is filled. We are an equal opportunity employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status or any other characteristic protected by law. Women and minorities are encouraged to apply.

New York University/Courant Institute of Mathematical Sciences Department of Computer Science Faculty

The department expects to have several regular faculty positions beginning in September 2015

IOWA STATE UNIVERSITY

Lecturer or Senior Lecturer The Department of Electrical and Computer Engineering

The Department of Electrical and Computer Engineering at Iowa State University invites applications for the position of Lecturer or Senior Lecturer to begin in 2015 (www.ece.iastate.edu). All interested, qualified persons must apply for this position by visiting <http://www.iastatejobs.com/postings/7548> and completing the Employment Application for vacancy #400010, with applications to be reviewed beginning immediately. For full consideration, applications must be received by December 1, 2014.

Women and members of underrepresented minorities are strongly encouraged to apply. Iowa State University is an Equal Opportunity/Affirmative Action employer. All qualified applicants will receive consideration for employment without regard to race, color, age, religion, sex, sexual orientation, gender identity, genetic information, national origin, marital status, disability, or protected veteran status, and will not be discriminated against. Inquiries can be directed to the Director of Equal Opportunity, 3350 Beardshear Hall, 515 294-7612, email eooffice@iastate.edu.

and invites candidates at all levels. We will consider outstanding candidates in any area of computer science.

Faculty members are expected to be outstanding scholars and to participate in teaching at all levels from undergraduate to doctoral. New appointees will be offered competitive salaries and startup packages, with affordable housing within a short walking distance of the department. New York University is located in Greenwich Village, one of the most attractive residential areas of Manhattan.

The department has 35 regular faculty members and several clinical, research, adjunct, and visiting faculty members. The department's current research interests include algorithms, cryptography and theory; computational biology; distributed computing and networking; graphics, vision and multimedia; machine learning and data science; natural language processing; scientific computing; and verification and programming languages.

Collaborative research with industry is facilitated by geographic proximity to computer science activities at AT&T, Facebook, Google, IBM, Bell Labs, NEC, and Siemens.

Please apply at <https://cs.nyu.edu/webapps/facapp/register>

To guarantee full consideration, applications should be submitted no later than December 1, 2014; however, this is not a hard deadline, as all candidates will be considered to the full extent feasible, until all positions are filled. Visiting positions may also be available.

New York University is an equal opportunity/affirmative action employer.

Northeastern Illinois University **Computer Science Department** **Assistant Professor**

The **Computer Science Department of Northeastern Illinois University** in Chicago invites individuals to apply for a **tenure-track, assistant professor position**, starting in September 2015. A Ph.D. in Computer Science or closely related field is required. We will consider applicants from all areas of computer science, but we are especially looking for candidates from the following areas: Software Engineering, Business Intelligence, Analysis and Design of Systems, Data Analytics, Informatics, and Emerging Technologies. Responsibilities include teaching workload of three courses per semester (fall/spring), advising Master's students, continuing with research, and service. Review of applications will begin on November 15, 2014, and continue until the position is filled. AA/EOE.

See:

<http://www.neiu.edu/academics/college-of-arts-and-sciences/departments/computer-science>.

Pacific Lutheran University **Assistant Professor of CSCE**

To teach, mentor and advise students at undergraduate level. Must hold Ph.D. in Computer Science or related field by September 1, 2015. To apply and for complete details visit <https://employment.plu.edu/postings/3171>

The Pennsylvania State University, **University Park** **College of Information Sciences and Technology** **Tenured and Tenure-Track Faculty Position in Information Security and Privacy**

The College of Information Sciences and Technology (IST) at the Pennsylvania State University invites applications for a faculty position in information security and privacy. We seek exceptional candidates with a high quality research and publication record to strengthen and complement our current research strengths in systems and software security, usability considerations in privacy and security, economics of information security, and privacy. We welcome applications from scholars at all ranks and with a variety of disciplinary backgrounds including computer and information sciences, decision sciences, and economics. Successful candidates will be expected to develop an internationally competitive, externally funded research program, contribute to graduate and undergraduate education and training. The Pennsylvania State University is the land grant institution of Pennsylvania. The College of IST offers a highly collaborative interdisciplinary research environment, strong research programs (the cyber-security program is currently funded by over \$4 million of prestigious external grants), a strong PhD program, and a very successful undergraduate program. To apply, please visit <http://apptkr.com/522425>. Review of applications will begin on Nov. 15, 2014, and continue until the position is filled. The Pennsylvania State University is an equal opportunity employer and is committed to increasing the diversity of its faculty. Enquiries about the position may be directed to recruit@ist.psu.edu.

CAMPUS SECURITY CRIME STATISTICS: For more about safety at Penn State, and to review the Annual Security Report which contains information about crime statistics and other safety and security matters, please go to <http://www.psu.edu/clery/>, which will also provide you with detail on how to request a hard copy of the Annual Security Report.

Penn State is an equal opportunity, affirmative action employer, and is committed to providing employment opportunities to minorities, women, veterans, disabled individuals, and other protected groups.

Purdue University **Tenure-Track/Tenured Faculty Positions**

The Department of Computer Science at Purdue University is entering a phase of significant growth, as part of a university-wide Purdue Moves initiative. Applications are solicited for tenure-track and tenured positions at the Assistant, Associate and Full Professor levels. Outstanding candidates in all areas of computer science will be considered. Review of applications and candidate interviews will begin early in October 2014, and will continue until the positions are filled.

The Department of Computer Science offers a stimulating and nurturing academic environment with active research programs in most areas of computer science. Information about the department and a description of open positions are available at <http://www.cs.purdue.edu>.

Applicants should hold a PhD in Computer Science, or related discipline, be committed to excellence in teaching, and have demonstrated excellence in research. Successful candidates will be expected to conduct research in their fields of expertise, teach courses in computer science, and participate in other department and university activities. Salary and benefits are competitive, and Purdue is a dual career friendly employer. **Applicants are strongly encouraged to apply online at <https://hiring.science.purdue.edu>.** Alternatively, hardcopy applications can be sent to: Faculty Search Chair, Department of Computer Science, 305 N. University Street, Purdue University, West Lafayette, IN 47907. A background check will be required for employment. Purdue University is an EEO/AA employer fully committed to achieving a diverse workforce. All individuals, including minorities, women, individuals with disabilities, and protected veterans are encouraged to apply.

San Diego State University **Department of Computer Science** **Assistant Professor of Computer Science**

The Department of Computer Science at SDSU seeks an Assistant Professor with a PhD in Computer Science or a closely related field and with expertise in either high-end computing or computer vision/graphics/animation. For more details and application procedures, please see <http://apply.interfolio.com/26228>. **SDSU is an equal opportunity/Title IX employer.**

Swarthmore College **Assistant Professor of Engineering**

Swarthmore College invites applications for a tenure-track position in the area of Computer/Electrical Engineering at the Assistant Professor level, to start during the Fall semester of 2015. A doctorate in Computer or Electrical Engineering is required. For more details on the position and application instructions, visit <https://academicjobsonline.org/ajo/jobs/4467>.

Swarthmore has a strong institutional commitment to excellence through diversity in its educational program and employment practices and actively seeks and welcomes applications from candidates with exceptional qualifications, particularly those with demonstrable commitments to a more inclusive society and world.

Swarthmore College **Visiting Assistant Professor**

Applications are invited for multiple 2-year visiting positions at the assistant professor level beginning Fall 2015.

Swarthmore College has a strong institutional commitment to excellence through diversity and inclusivity in its educational program and employment practices. The College actively seeks and welcomes applications from candidates with exceptional qualifications, particularly those with demonstrable commitments to a more inclusive society.

Swarthmore College is a small, selective, liberal arts college located 10 miles outside of Philadelphia.

phia. The Computer Science Department offers majors and minors at the undergraduate level.

We seek applicants who are enthusiastic about teaching undergraduates and already have teaching experience. Applicants should be comfortable teaching a wide range of courses at the introductory and intermediate level of CS. A Ph.D. in CS by or near the time of appointment is required.

We are particularly interested in applicants whose areas will add to the breadth of our program. However, strong applicants in all areas will be considered.

Applications will be accepted until the positions are filled. Applications should include a vita, teaching statement, research statement, and three letters of reference, at least one that speaks to the candidate's teaching ability. Applicants must apply online at Academic Jobs Online <http://academicjobsonline.org/ajob/jobs/4689>

Texas A&M University - Central Texas **Assistant Professor - Computer Information Systems**

TERM: 9 months/Tenure Track **SALARY:** Negotiable

Teaches a variety of undergraduate and/or graduate courses in CIS and CS.

MINIMUM REQUIREMENTS: Earned doctorate in IS or related areas such as CS. Recent Graduates and candidates nearing completion of the doctorate are encouraged to apply.

Apply URL: https://www.tamuctjobs.com/applicants/jsp/shared/Welcome_css.jsp

Tufts University **Computer Science Faculty Machine Learning/ Data Mining**

The Department of Computer Science in the School of Engineering at Tufts University invites applications for two tenure-stream faculty appointments to begin in September 2015. We are seeking candidates at the rank of Associate Professor or full Professor, although exceptional candidates at the rank of Assistant Professor will also be considered.

We welcome outstanding applicants with a strong vision and research programs in the areas of machine learning and data mining broadly interpreted to include related fields where data and its analysis are central. These include, for example, adaptive systems in robotics, computational linguistics, computational sustainability, and visual analytics.

Application materials should be submitted online through Academic Jobs Online (AJO) at <https://academicjobsonline.org/ajob/jobs/4533>.

For more information please visit <http://www.cs.tufts.edu/>.

Inquiries should be emailed to cssearch@cs.tufts.edu.

Review of applications will begin December 8, 2014.

Tufts University is an Affirmative Action/Equal Opportunity employer. We are committed to increasing the diversity of our faculty. Members of underrepresented groups are strongly encouraged to apply.

University College London (UCL) **Department of Computer Science** **Faculty Position**

The Department of Computer Science at University College London (UCL) invites applications for a faculty position in the areas of **Computer Systems and Networking**. We seek world-class talent; candidates must have an outstanding research track record.

Areas of interest for this position include operating systems, systems security, distributed systems, networking, and their intersection, with an emphasis on experimental system-building. Appointments will be made at the rank of Lecturer, Senior Lecturer, or Reader (equivalent to Assistant Professor, junior Associate Professor, and senior Associate Professor, respectively, in the US system), commensurate with qualifications.

Candidates must hold an earned Ph.D. in Computer Science or a closely related field by the time they begin their appointment. They will be evaluated chiefly on the significance and novelty of their research to date, and their promise for leading a group in a fruitful program of research. They must also demonstrate a zest for innovative and challenging teaching at the graduate and undergraduate levels. A proven record of ability to manage time and evidence of ability to teach and to supervise academic work by undergraduates, masters, and doctoral students are desirable. Our department is a highly collaborative environment, and we seek future colleagues who enjoy working collaboratively. Candidates should further be committed to public communication, and to UCL's policy of equal opportunity, including working harmoniously with colleagues and students of all cultures and backgrounds.

Since 1973, when UCL CS became the first ARPAnet node outside the United States, the department has been a leading centre for networking research, as demonstrated by its long-standing strong presence in the SIGCOMM conference proceedings.

UCL, the third oldest university in England, and the first to admit students without regard for their race or creed, sits a few blocks from the British Museum in central London, a vibrant metropolis of 8.4 million that offers an ever-renewing panoply of cultural and culinary choices, and is a budget airfare away from all of continental Europe.

Further details about UCL CS, the posts, and how to apply may be found at <http://www.cs.ucl.ac.uk/vacancies>

All application materials must reach UCL by the 16th of January, 2015.

We particularly welcome female applicants and those from an ethnic minority, as they are under-represented within UCL at these levels.

UCL Taking Action for Equality.

University of Alabama **Tenured/Tenure-Track Faculty Positions,** **Computer Science** **Software Engineering Focus: CyberSecurity and Big Data**

The Department of Computer Science at the University of Alabama invites applications for two faculty positions at the Assistant/Associate level, to begin August 2015. The focus is software engi-

neering, specifically applied to cybersecurity and big data. However, outstanding candidates in all areas will be considered. At the time of appointment, candidates must have earned a Ph.D. in Computer Science or a related field.

Located in Tuscaloosa, AL, near Birmingham, the University of Alabama is the largest institution in the state with roughly 35,000 students. The Computer Science Department has approximately 450 undergraduate and 50 graduate students. For more information on the software engineering group, visit: <http://se.cs.ua.edu>.

For additional details, visit <http://cs.ua.edu/> or contact Dr. Jeffrey Carver (faculty.search@cs.ua.edu) To apply, visit: <https://facultyjobs.ua.edu/postings/35920>. Review of applications will begin immediately. The University of Alabama is an equal opportunity/affirmative action employer. Women and minority applicants are particularly encouraged to apply.

University of California, Santa Barbara **Department of Computer Science** **Tenure-track Appointment**

The **University of California, Santa Barbara** invites applicants for a tenure-track appointment in the Department of Computer Science, effective July 1, 2015. Applicants must have a Ph.D. in computer science or a closely related field. While exceptional individuals in all areas will be considered, we are especially interested in the broad area of Big Data, with emphasis on machine learning, statistical methods, data mining, and scalable algorithms. Qualified candidates will be considered for appointment to the Eugene Aas Endowed Chair.

The department seeks individuals with promise and a proven track record of research excellence, a deep commitment to educating undergraduate and graduate students, and professional leadership abilities. We are particularly interested in candidates who value cross-disciplinary and collaborative research.

UC Santa Barbara is a premier public research university. The Department of Computer Science is part of the College of Engineering, in which all five departments are ranked among the top ten by the National Research Council (NRC). With 34 full-time faculty and more than 120 doctoral students currently, the Department of Computer Science has witnessed a rapid rise over the past ten years, both in size and stature. We are looking for candidates who will add to the department's dynamic culture and stimulating intellectual environment.

Additional information about the department and our academic programs can be found at <http://www.cs.ucsb.edu>.

Primary consideration will be given to applications received by December 15, 2014, however the position will remain open until filled. Applications should be submitted electronically as PDF documents to: <https://recruit.ap.ucsb.edu/apply/JPF00348>

Applications must include a detailed resume, research and teaching statements, and 3 reference letters.

The department is especially interested in candidates who can contribute to the diversity and excellence of the academic community through research, teaching and service. The Uni-

versity of California is an Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, or any other characteristic protected by law including protected veterans and individuals with disabilities.

University of California, Santa Barbara
Media Arts & Technology Graduate Program
Associate Professor in Media Arts & Technology

The Media Arts & Technology Program at UC Santa Barbara invites applications at the **Associate Professor** level for a joint tenured position in **Media Arts & Technology and Electrical and Computer Engineering**. Information about MAT can be found at <http://www.mat.ucsb.edu>.

The position will be available July 1, 2015.

The department seeks exceptional candidates who will establish a vigorous research and teaching program in Physical Computing or related fields that engage embedded sensing systems, distributed or autonomous systems, robotics, mobile/flying devices, immersive, interactive, distributed environments, and high-dimensional data generated in scientific, artistic, and architectural domains. The successful candidate will be expected to actively engage in interdisciplinary research, creative work and teaching in art or architecture and engineering or science.

Applicants are expected to hold a doctoral degree in Media Arts and Sciences, Electrical

Engineering, Computer Science, Architecture or a closely related field, have demonstrated excellence in research, and have a strong commitment to creative activity, teaching and interdisciplinary scholarship to be given serious consideration for this position.

Applicants should submit a curriculum vitae, research and teaching statements, contact information for three letters of reference, and links to projects and publications. Review of applications will begin on December 20, 2014 and continue until the position is filled.

The program is especially interested in candidates who can contribute to the diversity and excellence of the academic community through research, teaching and service. Please submit your application at: <https://recruit.ap.ucsb.edu/apply>. Inquiries about your application may be directed to: recruit@mat.ucsb.edu

The University of California is an Equal Opportunity / Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, or any other characteristic protected by law including protected Veterans and individuals with disabilities.

University of Central Florida CRCV
UCF Center for Research in Computer Vision
Assistant Professor

Multiple Assistant Professor Positions

CRCV is looking for multiple tenure-track faculty members in the Computer Vision area. Of particular interest are candidates with a strong track record of publications. CRCV will offer competitive salaries and start-up packages, along with a generous benefits package offered to employees at UCF.

Faculty hired at CRCV will be tenured in the Electrical Engineering & Computer Science department and will be required to teach a maximum of two courses per academic year and are expected to bring in substantial external research funding. In addition, Center faculty are expected to have a vigorous program of graduate student mentoring and are encouraged to involve undergraduates in their research.

Applicants must have a Ph.D. in an area appropriate to Computer Vision by the start of the appointment and a strong commitment to academic activities, including teaching, scholarly publications and sponsored research. Preferred applicants should have an exceptional record of scholarly research. In addition, successful candidates must be strongly effective teachers.

Applicants must submit all required documents at the time of application which includes the following: Research Statement; Teaching Statement; Curriculum Vitae; and a list of at least three references with address, phone numbers and email address.

Applicants for this position will also be considered for position numbers 38406 and 37361.

Apply URL:

<http://www.jobswithucf.com/postings/34681>

UCF is an Equal Opportunity/Affirmative Action employer. Women and minorities are particularly encouraged to apply.

University of Central Missouri
Assistant Professor of Computer Science - Multiple Positions

The Department of Mathematics and Computer Science at the University of Central Missouri is accepting applications for three tenure-track and several non-tenure track positions in Computer Science beginning August 2015 at the rank of Assistant Professor. The UCM Computer Science program has 18 full time faculty and about 1400 majors in both undergraduate and graduate programs. We are looking for faculty excited by the prospect of shaping our department's future and contributing to its sustained excellence.

Positions #997516 and #997517: Ph.D. in Computer Science by August 2015 is required. All areas in computer science will be considered with preference given to candidates with expertise in software engineering, information retrieval, data mining or cybersecurity.

Position #998458: M.S. in Computer Science is required. A doctorate must be completed before tenure and promotion consideration. The preferred area is in computer science education.

Non-Tenure Track Positions: Ph.D. in Computer Science or a closely related area is preferred. ABD will be considered. Previous college/university teaching experience is highly desirable.

Initial screening of applications begins December 1, 2014, and continues until position is filled. For more information about the positions and the application process, visit <http://www.ucmo.edu/math-cs/openings.cfm>. Apply URL: <https://jobs.ucmo.edu>. Contact: Dr. Songlin Tian; Email: tian@ucmo.edu.

University of Chicago
Department of Computer Science
Assistant Professor

The **Department of Computer Science** at the University of Chicago invites applications from exceptionally qualified candidates in the areas of (a) systems, and (b) theory of computing for faculty positions at the rank of **Assistant Professor**.

Systems is a broad, synergistic collection of research areas spanning systems and networking, programming languages and software engineering, software and hardware architecture, data-intensive computing and databases, graphics and visualization, security, systems biology, and a number of other areas. We encourage applicants working within our strategic focus of data-intensive computing, but also in all areas of systems.

The Theory of Computing ("Theory" for short) strives to understand the fundamental principles underlying computation and explores the power and limitations of efficient computation. While mathematical at its core, it also has strong connections with physics (quantum computing), machine learning, computer vision, natural language processing, network science, cryptography, bioinformatics, and economics, to name just a few areas. We encourage applications from researchers in core areas of Theory such as complexity theory and algorithms as well as in any area with a significant Theory component.

The University of Chicago has the highest standards for scholarship and faculty quality, is dedicated to fundamental research, and encourages collaboration across disciplines. We encourage connections with researchers across campus in such areas as bioinformatics, mathematics, molecular engineering, natural language processing, and statistics, to mention just a few.

The Department of Computer Science (cs.uchicago.edu) is the hub of a large, diverse computing community of two hundred researchers focused on advancing foundations of computing and driving its most advanced applications. Long distinguished in theoretical computer science and artificial intelligence, the Department is now building strong systems and machine learning groups. The larger community in these areas at the University of Chicago includes the Department of Statistics, the Computation Institute, the Toyota Technological Institute at Chicago (TTIC), and the Mathematics and Computer Science Division of Argonne National Laboratory.

The Chicago metropolitan area provides a diverse and exciting environment. The local economy is vigorous, with international stature in banking, trade, commerce, manufacturing, and transportation, while the cultural scene includes diverse cultures, vibrant theater, world-renowned symphony, opera, jazz, and blues. The University is located in Hyde Park, a Chicago neighborhood on the Lake Michigan shore just a few minutes from downtown.

Applicants must have completed all requirements for the PhD at the time of appointment. The PhD should be in Computer Science or a related field such as Mathematics, or Statistics, etc.

Applications must be submitted through the University's Academic Jobs website.

To apply for the Assistant Professor - Systems, go to: <http://tinyurl.com/k3wgaqv>

To apply for the Assistant Professor - Theory, go to: <http://tinyurl.com/khyc74d>

To be considered as an applicant, the following materials are required:

- ▶ cover letter,
- ▶ curriculum vitae including a list of publications,
- ▶ statement describing past and current research accomplishments and outlining future research plans, and
- ▶ description of teaching philosophy, and
- ▶ three reference letters, one of which must address the candidate's teaching ability.

Reference letter submission information will be provided during the application process.

Review of application materials will begin on January 15, 2015 and continue until all available positions are filled.

All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, age, protected veteran status or status as an individual with disability.

The University of Chicago is an Affirmative Action / Equal Opportunity / Disabled / Veterans Employer.

University of Chicago **Department of Computer Science** **Associate Professor**

The **Department of Computer Science** at the **University of Chicago** invites applications from exceptionally qualified candidates in the areas of (a) systems, and (b) theory of computing for faculty positions at the rank of **Associate Professor**.

Systems is a broad, synergistic collection of research areas spanning systems and networking, programming languages and software engineering, software and hardware architecture, data-intensive computing and databases, graphics and visualization, security, systems biology, and a number of other areas. We encourage applicants working within our strategic focus of data-intensive computing, but also in all areas of systems.

The Theory of Computing ("Theory" for short) strives to understand the fundamental principles underlying computation and explores the power and limitations of efficient computation. While mathematical at its core, it also has strong connections with physics (quantum computing), machine learning, computer vision, natural language processing, network science, cryptography, bioinformatics, and economics, to name just a few areas. We encourage applications from researchers in core areas of Theory such as complexity theory and algorithms as well as in any area with a significant Theory component.

The University of Chicago has the highest standards for scholarship and faculty quality, is dedicated to fundamental research, and encourages collaboration across disciplines. We encourage connections with researchers across campus in such areas as bioinformatics, mathematics, molecular engineering, natural language processing, and statistics to mention just a few.

The Department of Computer Science (cs.uchicago.edu) is the hub of a large, diverse computing community of two hundred researchers focused on advancing foundations of computing and driving its most advanced applications. Long distinguished in theoretical computer science and artificial intelligence, the Department is now building strong systems and machine learning groups. The larger community in these areas at the University of Chicago includes the Depart-

ment of Statistics, the Computation Institute, the Toyota Technological Institute at Chicago (TTIC), and the Mathematics and Computer Science Division of Argonne National Laboratory.

The Chicago metropolitan area provides a diverse and exciting environment. The local economy is vigorous, with international stature in banking, trade, commerce, manufacturing, and transportation, while the cultural scene includes diverse cultures, vibrant theater, world-renowned symphony, opera, jazz, and blues. The University is located in Hyde Park, a Chicago neighborhood on the Lake Michigan shore just a few minutes from downtown.

Applicants must have a doctoral degree in Computer Science or a related field such as Mathematics, Statistics, etc. Applicants are expected to have established an outstanding research program and will be expected to contribute to the department's undergraduate and graduate teaching programs.

Applications must be submitted through the University's Academic Jobs website.

To apply for the position of Associate Professor-Systems, go to: <http://tinyurl.com/pkzpcy5>

To apply for the position of Associate Professor-Theory, go to: <http://tinyurl.com/kwzb9zu>

To be considered as an applicant, the following materials are required:

- ▶ cover letter,
- ▶ curriculum vitae including a list of publications,
- ▶ statement describing past and current research accomplishments and outlining future research plans, and
- ▶ description of teaching philosophy, and
- ▶ a reference contact list consisting of three people

Review of complete applications will begin January 15, 2015 and will continue until all available positions are filled.

All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, age, protected veteran status or status as an individual with disability.

The University of Chicago is an Affirmative Action / Equal Opportunity / Disabled / Veterans Employer.

University of Idaho **Computer Science Dept** **Assistant Professor**

The Computer Science Dept, CoEngr, Univ. of Idaho invite applications for a tenure-track faculty position. Specialization in big data analytics and data interoperability. Areas of expertise of interest include data analytics for scientific applications in ecology, biology, and social sciences, exploiting data from disparate sources and heterogeneous formats. Responsibilities include undergraduate & graduate teaching, scholarship & research, and professional and university service. Apply URL: <https://uidaho.peopleadmin.com/postings/5466>. Contact Person: Rhonda Zenner, Email Address: rzenner@uidaho.edu, Phone: 208-885-6589, Fax: 208-885-9052

University of Illinois at Chicago **Department of Computer Science** **Multiple Faculty Positions**

The Computer Science Department at the University of Illinois at Chicago invites applications in

all areas of Computer Science for multiple tenure-track positions at the rank of Assistant Professor (exceptional candidates at other ranks will also be considered). We are looking to fill:

(a) One position in Big Data, where our focus ranges from data management and analytics to visualization and applications involving large volumes of data.

(b) Two positions in Computer Systems, where we are looking for candidates whose work is experimental and related to one or more of the following topics: operating systems, networking, distributed computing, mobile systems, programming languages and compilers, security, software engineering, and other broadly related areas.

(c) One position for which candidates from all other areas will be considered.

The University of Illinois at Chicago (UIC) ranks among the nation's top 50 universities in federal research funding and is ranked 4th best U.S. University under 50 years old by Times Higher Education. The Computer Science department has 24 tenure-track faculty representing major areas of computer science, and offers BS, MS and PhD degrees. Our faculty includes ten NSF CAREER award recipients. We have annual research expenditures of \$8.4M, primarily federally funded. UIC is an excellent place for interdisciplinary work—UIC houses the largest medical school in the country, and our faculty are engaged with several cross-departmental collaborations with faculty from health sciences, social sciences and humanities, urban planning, and the business school. UIC has an advanced networking infrastructure in place for data-intensive scientific research that is well-connected regionally, nationally and internationally. UIC also has strong collaborations with Argonne National Laboratory and with the National Center for Supercomputing Applications, with programs in place for faculty members to apply for time on their high-performance supercomputing systems.

Chicago epitomizes the modern, livable, vibrant city. Located on the shore of Lake Michigan, it offers an outstanding array of cultural and culinary experiences. As the birthplace of the modern skyscraper, Chicago boasts one of the world's tallest and densest skylines, combined with a 8100-acre park system and extensive public transit and biking networks. Its airport is the second busiest in the world, with frequent non-stop flights to virtually anywhere. Yet the cost of living, whether in an 88th floor condominium downtown or on a tree-lined street in one of the nation's finest school districts, is surprisingly low.

Applications must be submitted at <https://jobs.uic.edu/>. Please include a curriculum vitae, teaching and research statements, and names and addresses of at least three references in the online application. Applicants needing additional information may contact the Faculty Search Chair at search@cs.uic.edu.

For fullest consideration, please apply by December 1, 2014. We will continue to accept and process applications until all the positions are filled. The University of Illinois is an Equal Opportunity, Affirmative Action employer. Minorities, women, veterans and individuals with disabilities are encouraged to apply.

University of Illinois at Urbana-Champaign
Department of Electrical and Computer Engineering (ECE)
Positions in Computing

The Department of Electrical and Computer Engineering (ECE) at the University of Illinois at Urbana-Champaign invites applications for faculty positions at all areas and levels in computing, broadly defined, with particular emphasis on big data and its applications, including complex data analysis and decision science, scalable hardware and software systems; parallel, high-performance, and energy-efficient computing; reliable and secure computing; wired/wireless/social networking & distributed computing; mobile, wearable sensing & applications; and bioinformatics & systems biology, among other areas. From the transistor and the first computer implementation based on von Neumann's architecture to the Blue Waters petascale computer – the fastest computer on any university campus, ECE Illinois faculty have always been at the forefront of computing research and innovation. Applications are encouraged from candidates whose research programs specialize in core as well as interdisciplinary areas of electrical and computer engineering. The department is engaged in exciting new and expanding programs for research, education, and professional development, with strong ties to industry. This year the ECE Department has moved into its new 235,000 sq. ft. net-zero energy design building, which is a major campus addition with maximum space and minimal carbon footprint.

Qualified senior candidates may also be considered for tenured full Professor positions as part of the Grainger Engineering Breakthroughs Initiative (<http://graingerinitiative.engineering.illinois.edu>), which is backed by a \$100-million gift from the Grainger Foundation to support research in big data and bioengineering, broadly defined. In addition, the University of Illinois is home to Blue Waters – one of the most powerful supercomputers in the world, supported by the National Science Foundation and developed and operated by the University of Illinois' National Center for Supercomputing Applications. Qualified candidates may be hired as Blue Waters Professors who will be provided substantial allocations on and expedited access to the supercomputer. To be considered as a Blue Waters Professor, candidates need to mention Blue Waters as one of their preferred research areas in their online application, and include a reference to Blue Waters in their cover letter.

Please visit <http://jobs.illinois.edu> to view the complete position announcement and application instructions. Full consideration will be given to applications received by November 15, 2014, but applications will continue to be accepted until all positions are filled.

Illinois is an EEO Employer/Vet/Disabled www.inclusiveillinois.illinois.edu.

University of Kansas
Department of Electrical Engineering and Computer Science
Electrical Engineering/Computer Engineering / Computer Science Faculty

The University of Kansas (KU) Department of Electrical Engineering and Computer Science

(EECS) seeks outstanding individuals for **four tenure track positions** in the disciplines of electrical engineering, computer engineering, or computer science. Successful candidates are expected to contribute to the development of academic and research programs and to contribute to the research community. Applicants are expected to have an earned doctorate, ABD, or equivalent in electrical engineering, computer science, computer engineering, or related fields at the time of appointment. The successful candidate must be eligible to work in the U.S. by the time of appointment. The following positions are open at the assistant professor (tenure-track) level, with the possibility of higher rank for the exceptional applicant, with experience in:

Data Science/Health Informatics — Data science and applications; large-scale data analytics; smart health; decision supporting systems in health; text mining; information retrieval; data mining; machine learning. (<http://employment.ku.edu/academic/1781BR>)

Computer Architecture and Reconfigurable Computing — Multiple-core architectures; Systems-on-Chip (SoC), and Networks-on-Chip (NOC); embedded systems; programmable and adaptive architectures; digital system design and synthesis; high-assurance, secure, resilient, and cyber-physical computer architectures; exascale computing architecture and systems; special purpose computing and memory architectures for data analytics and retrieval. (<http://employment.ku.edu/academic/1779BR>)

Bioinformatics — New computationally intensive technologies and software tools for efficiently; storing biological data, retrieving biological data, organizing biological data, analyzing biological data. Other areas of computer science focused on computational life sciences may be considered for exceptional applicants. (<http://employment.ku.edu/academic/1780BR>)

Wireless Devices, Communications and Applications — Wireless and sensing devices integrated into distributed systems to enable new mobile applications; sensors, e.g., radar or biomedical, with integrated wireless communications; wireless devices, sensors and systems applied to challenged environments, e.g., civilian disasters, or military, or covert communications; "Networks of things" and distributed sensing; telemetry and applications; smart grid communications for power distribution and monitoring. (<http://employment.ku.edu/academic/1766BR>)

Exceptional applicants in other closely related areas to the above topics may be considered.

The KU School of Engineering is rapidly expanding and plans to add thirty new faculty lines, over a 5 year period, with expanded research and teaching facilities. The University of Kansas is focused on four key campus-wide strategic initiatives: (1) Sustaining the Planet, Powering the World; (2) Promoting Well-Being, Finding Cures; (3) Building Communities, Expanding Opportunities; and (4) Harnessing Information, Multiplying Knowledge. For more information, see <http://www.provost.ku.edu/planning/themes/>. Successful candidates should describe in their application materials how their work addresses one or more of KU's strategic initiatives.

Applications should be submitted at <https://employment.ku.edu> and click on Faculty link. Applications should include a letter of application, curriculum vitae, a statement of research interests and plans, a statement of teaching interests and plans, and contact information for three references. Applications will be reviewed beginning December 1, 2014 and will be accepted until the position is filled. The appointment will be effective as negotiated. Questions can be sent to: EECS_Search@eeecs.ku.edu. KU is an EO/AEE. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex (including pregnancy), age, national origin, disability, genetic information or protected Veteran status.

The University of Michigan, Ann Arbor
Department of Electrical Engineering and Computer Science
Computer Science and Engineering Division
Faculty Positions
EECS/CSE Faculty

Applications and nominations are solicited for multiple faculty positions in the Computer Science and Engineering (CSE) Division. Highly qualified candidates from all areas of computer science and computer engineering will be considered for positions at all levels and are encouraged to apply.

Qualifications include an outstanding academic record, a doctorate or equivalent in computer science or computer engineering, and a strong commitment to teaching and research. Applications must be received by January 1, 2015.

To apply, please complete the form at: <https://www.eecs.umich.edu/eeecs/etc/jobs/csejobs.html>

Electronic applications are strongly preferred, but you may alternatively send resume, teaching statement, research statement and names of three references to:

Professor Satinder Singh Baveja
Chair, CSE Faculty Search
Department of Electrical Engineering and Computer Science
University of Michigan
2260 Hayward Street
Ann Arbor, MI 48109-2121

The University of Michigan is a Non-Discriminatory/Affirmative Action Employer with an Active Dual-Career Assistance Program. The college is especially interested in candidates who can contribute, through their research, teaching, and/or service, to the diversity and excellence of the academic community.

University of Nebraska-Lincoln
Department of Computer Science and Engineering
Assistant Professor

The Department of Computer Science and Engineering at the invites applications for a tenure track faculty position at the rank of Assistant Professor to begin in August, 2015.

We are seeking a faculty member who can establish a strong scholarly research and teaching

program in the area of **Robotics**, complementing the work being performed in the robotics, cyber-physical networking, software engineering, and informatics labs in the department. We are particularly interested in the subareas of controls, perception, coordination, collaboration, and human-robot interaction.

To apply go to <http://employment.unl.edu> and complete the Faculty/Administration application for requisition #F_140123 and attach 1) a single-page cover letter explaining your interest in UNL, 2) a curriculum vitae, 3) teaching and research statements, 4) 1-2 papers that best represent research contributions and scholarship, and 5) a list of at least three references.

Review of applications will begin on December 1st, 2014, and will continue until the position has been filled. The position is contingent on the availability of funds.

The official advertisement can be viewed at: <http://cse.unl.edu/facultysearch>.

The University of Nebraska-Lincoln is committed to a pluralistic campus community through affirmative action, equal opportunity, work-life balance, and dual careers.

University of Nebraska-Lincoln
Department of Computer Science and Engineering
Assistant, Associate or Full Professor

The Department of Computer Science and Engineering at the University of Nebraska – Lincoln invites applications for a tenure track faculty position at the rank of Assistant, Associate or Full Professor to begin in August, 2015.

We are seeking a faculty member who can establish a strong scholarly research and teaching program in the area of **software security**, complementing the Department's strengths in informatics, software engineering, and systems. Areas of interest include data security, system security, trustworthy computing, and usable security and privacy.

To apply, go to: <http://employment.unl.edu> and complete the Faculty/Administrative application for requisition F_140125. Attach 1) cover letter explaining why you are interested in UNL, 2) a curriculum vitae, 3) research and teaching statements, 4) 1-2 papers that best represent research contributions and scholarship, and 5) a list of at least three references.

Review of applications will begin on December 1st, 2014, and will continue until the position has been filled. The position is contingent on the availability of funds.

The official advertisement can be viewed at: <http://cse.unl.edu/facultysearch>.

The University of Nebraska-Lincoln is committed to a pluralistic campus community through affirmative action, equal opportunity, work-life balance, and dual careers.

University of Nebraska-Lincoln
Department of Computer Science and Engineering in the College of Engineering
Assistant Professor

The **Department of Computer Science and Engineering in the College of Engineering** at the **University of Nebraska-Lincoln (UNL)** invites ap-

plications for a tenure-track position at the rank of **Assistant Professor**. We are particularly seeking candidates who specialize in **Autonomous Systems**.

The UNL College of Engineering is undergoing an exciting period of significant growth, related in part to the University joining the Big Ten Conference in 2011. The College anticipates hiring 100 new faculty in the next five years. An allocation of \$2.5 million has been authorized for new faculty hires to start in Fall 2015, and this level of funding is expected to continue for several years. A major focus of the college's teaching, research and outreach enterprise in the coming decade will be on the three major industries of the state of Nebraska: manufacture of food, manufacture of civil infrastructure, and manufacture of equipment, devices and components. Consequently the College is seeking to hire a multidisciplinary group of faculty that have complementary expertise related to these manufacturing foci, including but not limited to: dynamics and controls; robotics and automation; materials science; sensors and sensor development; telecommunications to interface between sensors and computers; big data for real time monitoring and control; computational modeling; and systems and industrial engineering.

Applicants should have a demonstrated potential for outstanding research and a promise of effective teaching at undergraduate through graduate levels including the advising of M.S. and Ph.D. students. Candidates will hold a PhD in Robotics, Computer Science, Computer Engineering, or a closely related discipline. Applicants should have excellent communication skills and a strong desire to work in a team/collaborative environment.

To be considered for this position, please go <http://employment.unl.edu> (requisition #F_140149), complete the Faculty/Academic Administrative Form, and attach required documents. Complete applications will include 1) a single-page cover letter explaining your interest in UNL, 2) a curriculum vitae, 3) teaching and research statements, 4) 1-2 papers that best represent research contributions and scholarship, and 5) a list of at least three references. Additional information about the Department of Computer Science and Engineering can be found at <http://www.cse.unl.edu>. Inquiries regarding this search may be addressed to: Dr. Elbaum, Chair of the Faculty Search Committee, elbaum@unl.edu, (402) 472-6748.

Review of applications will begin December 1, 2014, and continue until the position is filled. The University of Nebraska is committed to a pluralistic campus community through affirmative action and equal opportunity, work-life balance, and dual careers.

University of Notre Dame
Department of Computer Science and Engineering
Assistant or Associate Professor

The Department of Computer Science and Engineering at the University of Notre Dame invites applications for Assistant or Associate Professor. Excellent candidates in all areas will be considered.

The Department offers the PhD degree and undergraduate Computer Science and Computer

Engineering degrees. Faculty are expected to excel in classroom teaching and to build and lead highly-visible research projects that attract substantial external funding.

The University of Notre Dame is a private, Catholic university with a doctoral research extensive Carnegie classification, and consistently ranks in USNWR as a top-twenty national university. The South Bend area has a vibrant and diverse economy with affordable housing and excellent school systems, and is within easy driving distance of Chicago and Lake Michigan.

Applicants should send (pdf format preferred) a CV, statement of teaching and research interests, and contact information for three professional references to: cse-search-2015@nd.edu.

The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

University of Texas at Austin
Department of Computer Science
Tenured/Tenure-Track Faculty

The Department of Computer Science of the University of Texas at Austin invites applications for tenure-track positions at all levels. Outstanding candidates in all areas of Computer Science will be considered, particularly in experimental systems. All tenured and tenure-track positions require a Ph.D. or equivalent degree in computer science or a related area at the time of employment.

Successful candidates are expected to pursue an active research program, to teach both graduate and undergraduate courses, and to supervise graduate students. The department is ranked among the top ten computer science departments in the country.

It has 41 tenured and tenure-track faculty members across all areas of computer science. Many of these faculty participate in interdisciplinary programs and centers in the University, including the Texas Advanced Computing Center (TACC), and those in Computational and Applied Mathematics, Computational Biology, and Neuroscience.

Austin, the capital of Texas, is a center for high-technology industry, including companies such as IBM, Dell, Freescale Semiconductor, Advanced Micro Devices, National Instruments, AT&T, Intel and Samsung. For more information please see the department web page <http://www.cs.utexas.edu/>

The department prefers to receive applications online, beginning September 2, 2014. To submit yours, please visit <http://www.cs.utexas.edu/faculty/recruiting>

Applicants for an assistant professor position must have at least three (3) referees send letters of reference directly to the address provided. Applicants for a tenured position (associate or full professor) must have at least six (6) referees send letters of reference directly. Inquiries about your application may be directed to faculty-search@cs.utexas.edu.

For full consideration of your application, please apply by January 31, 2015. Complete applications (including all reference letters) will begin being reviewed on December 16th. Women and minority candidates are especially encouraged to apply. The University of Texas is an Equal Opportunity Employer.

The University of Texas at San Antonio Faculty Positions in Computer Science

The Department of Computer Science at **The University of Texas at San Antonio** invites applications for multiple tenure/tenure-track positions, starting Fall 2015. Interested candidates may apply to one or more positions indicated below.

► **Endowed tenure-track faculty positions in Cloud Computing** focused on high performance computing (HPC) in cloud, big data and data science, cloud architecture, virtualization, risk assessment, cloud security and/or system management. There are multiple positions at the assistant or associate professor level. Each of these positions includes a mini-endowment established as part of the university-wide cloud computing initiative.

► **A tenure-track faculty position in Computer Science** focused on system areas, including HPC, architecture and system software. This is a tenure-track assistant or associate professor position.

► **A tenured, senior faculty position in Cyber Security** focused on systems security, including cloud security, mobile security, malware defense, and/or networking security. This is a tenured associate or full professor position. Excellent candidates in all areas of cyber security will be considered.

See <http://www.cs.utsa.edu/fsearch> for information on the Department and application instructions. Screening of applications will begin immediately. Full consideration will be given to applications received by January 5, 2015, and the search will continue until the positions are filled or the search is closed. The University of Texas at San Antonio is an Affirmative Action/Equal Opportunity Employer.

Department of Computer Science
RE: Faculty Search
The University of Texas at San Antonio
One UTSA Circle
San Antonio, TX 78249-0667
Phone: 210-458-4436

University of Texas at Tyler Chair, Computer Science and CIS Department

Applications are invited for Chair of the Department of Computer Science which is housed within the College of Business and Technology at the **University of Texas at Tyler**. The Department offers the BS and MS in Computer Science and the BS in Computer Information System.

Requirements include a Doctorate in CS or CIS or closely related field, and a record sufficient for appointment at the rank of Professor. There is a preference for experience in the CIS area. This is a 12 month position.

Send a letter of interest, summary of qualifications and a resume/CV to Dean James Lumpkin at kgaribaldi@uttyler.edu. Evaluation will begin immediately and continue until the position is filled. A January, 2015 start may be possible. The University of Texas at Tyler is an Equal Opportunity Employer.

Vanderbilt University Department of Electrical Engineering and Computer Science (EECS) Tenured/Tenure-Track (T/TK) Faculty Position

THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE (EECS)

AT VANDERBILT UNIVERSITY is seeking candidates for a tenured/tenure-track (T/TK) faculty position in Computer Science. Appointments at all ranks will be considered, with a preference for candidates at the assistant professor level. The areas of primary emphasis are Big Data, Data Visualization, and/or Cyber-physical Systems. The Vanderbilt CS program provides a unique, collaborative, and interdisciplinary research environment that spans a wide range of topics in Computer Science. New trans-institutional programs are creating opportunities for research on issues of broad significance that create and extend collaborations across multiple fields.

Average T/TK faculty funding is ~\$1M per year from DARPA, DoD, NASA, NIH, NSF, and many industry sponsors. All junior faculty members hired during the past decade have received prestigious young investigator awards, such as NSF CAREER and DARPA CSSG. Successful candidates are expected to (1) teach at the undergraduate and graduate levels, (2) develop and grow vigorous programs of externally funded research, and (3) strengthen the CS program's research areas and enhance the School of Engineering's strategic directions in health care and medicine, security, energy and natural resources, and entertainment.

Vanderbilt University is a private, internationally renowned research university located in vibrant Nashville, Tennessee. Its 10 schools share a single cohesive campus that nurtures interdisciplinary activities. The School of Engineering currently comprises 90 tenured and tenure-track faculty, operates with an annual budget of \$100 million including over \$70 million of externally funded research, and teaches 1,300 undergraduate and over 400 graduate students. The School of Engineering over the past decade has been on a strong upward trajectory in national and international stature and prominence, and is in the process of building infrastructure to support a significant expansion in faculty size over the next five years. In the 2014 rankings of graduate engineering programs by *U.S. News & World Report*, the School ranks third among programs with fewer than 100 faculty. With a metro population of approximately 1.5 million people, Nashville has been named one of the 15 best U.S. cities for work and family by *Fortune* magazine, was ranked as the #1 most popular U.S. city for corporate relocations by *Expansion Management* magazine, and was named by *Forbes* magazine as one of the 25 cities most likely to have the country's highest job growth over the coming five years. Major industries include tourism, printing and publishing, manufacturing technology, music production, higher education, finance, insurance, automobile production and health care management.

Vanderbilt University is an equal-opportunity, affirmative-action employer. We strongly encourage applications from women and members of under-represented minority groups. **Applications should be submitted on-line at** <https://academicjobsonline.org/ajob/jobs/4547>. For more information, please visit our web site: <http://engineering.vanderbilt.edu/eeecs/>. Applications will be reviewed on a rolling basis, with an initial deadline of December 15, 2014, but will be accepted until the position is filled.

Westmont College Computer Science Faculty

Westmont College in Santa Barbara, CA, seeks applicants for a tenure-track opening in CS. We are a selective Christian liberal arts college in the broad evangelical tradition. Review of applications begins Oct. 1.

Visit http://www.westmont.edu/_offices/provost/OpenPositions.html for more details.

Worcester Polytechnic Institute Computer Science Department Multiple Tenure-track Faculty

Looking for faculty colleagues who engage deeply in both teaching and research within a curriculum that embraces student projects and independent learning? Consider joining the faculty at **Worcester Polytechnic Institute**.

The Computer Science Department anticipates hiring **multiple tenure-track faculty** for the Fall of 2015 whose expertise is in the following areas: Cybersecurity including network, system, or software security, software reliability and usable security; Algorithms, particularly individuals with collaboration possibilities with other faculty in Computer Science; Visualization and User Interfaces; Robotics working in a cross-disciplinary program with faculty from Electrical & Computer Engineering and Mechanical Engineering; Data Science including in big data analytics, statistical learning, large-scale data management, and data/text mining, working in a cross-disciplinary program; and Serious Games working with faculty in our Interactive Media & Game Development as well as our Learning Sciences & Technologies interdisciplinary programs.

In addition to these specific areas, outstanding candidates in any area will receive full consideration. Candidates should have a PhD in Computer Science or a closely related field, and the potential for excellence in research and teaching.

Founded in 1865, WPI is one of the nation's first technological universities. A highly selective private university located within an hour of Boston, WPI is consistently ranked among the top 60 research institutions by *US News & World Report*. The university is home to an innovative and intensive project-based curriculum that empowers students with the knowledge and skills to address real world problems around the globe, an approach repeatedly cited for excellence by *The Fiske Guide to Colleges* and *The Princeton Review*.

Located in the heart of New England, WPI is surrounded by cultural and recreational opportunities. The UMass Medical Center, a large number of technology companies and many colleges and universities are located in the immediate area making it ideal for two-career families.

Questions about the hiring process should be sent to recruit@cs.wpi.edu.

More information about the positions and instructions for applying are available at <http://apptrkr.com/521088>. You will need to include detailed research and teaching statements, vitae and contact information for at least three references.

Review of applications will begin December 15, 2014 and continue until the positions are filled.

To enrich education through diversity, WPI is an affirmative action, equal opportunity employer.



DOI:10.1145/2670917

Dennis Shasha

Upstart Puzzles

Proving without Teaching/ Teaching without Proving

PETER WINKLER'S MATHEMATICALLY elegant, often whimsical puzzles have been a joy to read and wrestle with. My columns now embark on a different path. Each will come in two parts: one quite doable, for which I will provide a solution, the other an "upstart," or insolent variant, I do not know how to solve and that will be an open problem for the community. Here we go.

"Zero knowledge" proofs entail interaction between someone (the Prover) asserting a statement and someone else (the Verifier) determining whether it is true. The proofs give zero knowledge to the Verifier, as the Verifier gains no knowledge other than that the statement is true, either absolutely or beyond a reasonable doubt.¹ The statement in question is normally of a mathematical nature, but, in this case, it resembles, well, child's play.

Imagine the charming gentleman in the figure calls himself the Amazing Sand Counter. He says if he sees a bucket of sand, he will know how many grains are in it, though he will not tell you the number. Your task is to determine whether he is telling the truth beyond a reasonable doubt. Here are the rules of engagement: You may send him out of the room, put a tent over yourself and the bucket, and add or remove a few grains from/to the bucket. He will answer any question that will not give you information about the total number of grains in the bucket. This is the easy part. Try solving it before reading on.

Solution. Send him out of the room, cover yourself with a tent, take a few grains you count, put them in your pocket, remove the tent, invite him to



Amazing Sand Counter claims to know the number of grains in the bucket just by looking at it. Do you believe him?

return to the room, and ask him how many grains you have removed. Do this as many times as you like. If Amazing passes many such tests, then you know beyond a reasonable doubt he knows how many grains are in the bucket (assuming no X-ray vision), because the only information he has is that he is able to look at the bucket. In the meantime, you still have not learned any more than you might by simply counting the number of grains you added or removed.

Upstart, or open-problem, variant. Suppose Amazing says there are N grains. You do not necessarily believe him. Your task is to design a protocol that determines whether he is telling the truth. That protocol should take less than $O(\sqrt{N})$ time. Here are the operations you are allowed (all taking $O(1)$ time): counting a single grain; dividing the bucket into approximately

equal portions (the grains may have different shapes and weights so you can approximate divisions only by two); and asking a question of Amazing whose answer may or may not be true. You have many buckets at your disposal. You win if, in less than $O(\sqrt{N})$ time, you determine he is telling the truth or catch him lying (even once) with probability at least $\frac{1}{2}$. Otherwise, he wins. ■

Reference

1. Goldwasser, S., Micali, S., and Rackoff, C. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing* 18, 1 (Feb. 1989), 186–208.

All are invited to submit solutions and prospective upstart-style puzzles for future columns to upstartpuzzles@cacr.nyu.edu

Dennis Shasha (dennishasha@yahoo.com) is a professor of computer science in the Computer Science Department of the Courant Institute of New York University, New York.

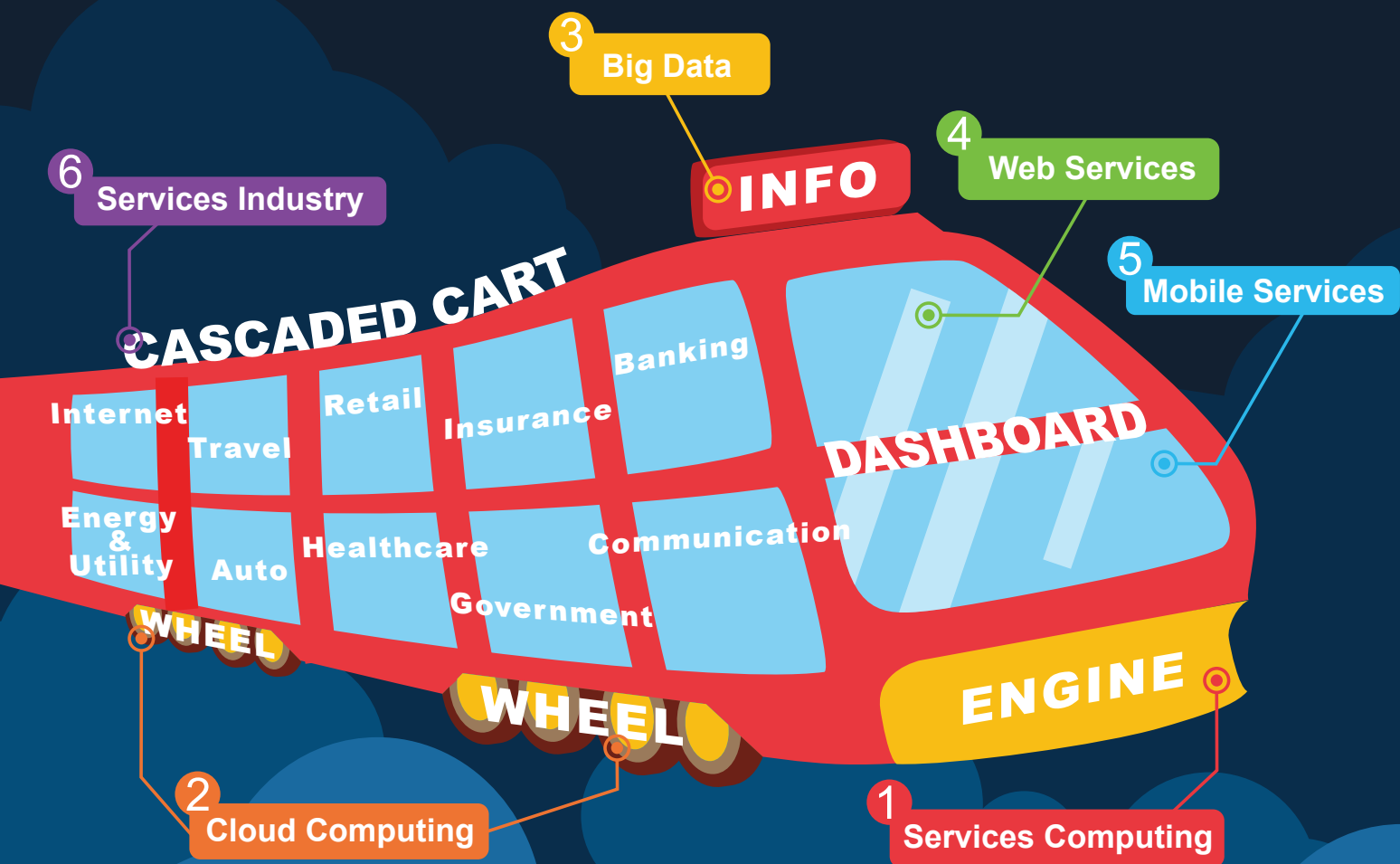
Copyright held by author.

2015 IEEE Services Congress IEEE BigData Congress

June 27 - July 2, 2015, New York, USA



- ① IEEE 12th International Conference on Services Computing (SCC 2015)
- ② IEEE 8th International Conference on Cloud Computing (CLOUD 2015)
- ③ IEEE 4th International Congress on Big Data (BigData Congress 2015)
- ④ IEEE 22nd International Conference on Web Services (ICWS 2015)
- ⑤ IEEE 4th International Conference on Mobile Services (MS 2015)
- ⑥ IEEE 11th World Congress on Services (SERVICES 2015)



Submission Deadlines

- 1/15/2015 ICWS 2015 (<http://icws.org>)
- 1/15/2015 CLOUD 2015 (<http://theCloudComputing.org>)
- 1/31/2015 SCC 2015 (<http://conferences.computer.org/scc>)
- 1/31/2015 MS 2015 (<http://theMobileServices.org>)
- 2/10/2015 BigData Congress 2015 (<http://ieeeBigData.org>)
- 2/10/2015 SERVICES 2015 (<http://ServicesCongress.org>)

Contact: confs@ServicesSociety.org



VEE 2015

11th ACM SIGPLAN/SIGOPS international conference on

Virtual Execution Environments

Istanbul, Turkey March 14-15, 2015 with ASPLOS

Authors are invited to submit original papers related to virtualization across all layers of the software stack, from high-level language virtual machines down to the microarchitectural level. VEE 2015 accepts both full-length and short papers.

Deadline: November 25, 2014



General Chair

Ada Gavrilovska (Georgia Institute of Technology)

Program Committee

Jonathan Appavoo (Boston University)
Haibo Chen (Shanghai Jiao Tong University)
Dilma Da Silva (Texas A&M University)
Amer Diwan (Google)
Daniel Frampton (Microsoft Research)
David Gregg (Trinity College Dublin)
David Grove (IBM Research)
Vishakha Gupta (Intel)

Program Co-chairs

Angela Demke Brown (University of Toronto)
Bjarne Steensgaard (Microsoft Research)

Tomas Kalibera (Purdue University)
Kenichi Kourai (Kyushu Institute of Technology)
Priya Nagpurkar (IBM Research)
Donald Porter (Stony Brook University)
Jennifer Sartor (Ghent University)
Ravi Soundararajan (VMWare)
Gael Thomas (LIP6)
Timothy Wood (George Washington University)



in cooperation with



<http://vee.sigops.org/vee15>