COMMUNICATIONS of the ACM

MAY 2008 VOLUME 51, NUMBER 5

# WEB SEARCHING IN A MULTILINGUAL WORLD

TAMING AGENT ARCHITECTURES

HOW INTUITIVE IS OO DESIGN?

REDUCING INTERNET AUCTION FRAUD

THE ROLE OF CS IN SYNTHETIC BIOLOGY

ACM ELECTION BALLOT

EMERGING TRENDS IN M-GOVERNMENT

Association for Computing Machinery

essays, {craft, art, science} of software, python, eclipse, agile development, onward!, {generative, functional} programming, .net, open source, concurrency, smalltalk, aspects, second life, ruby, service-orientation, objects, embedded, ultra large scale {model, test}-driven passion, fun!, agents, domain-specific use cases, movies, lightning talks, systems, objective-c, development, c#, design patterns, languages, wiki, product-lines, java, refactoring, plop

# OOPSLA

2008 NASHVILLE

## DEFINE THE FUTURE OF SOFTWARE

## www.oopsla.org/submit

### CONFERENCE CHAIR
**GAIL E. HARRIS**
Instantiated Software Inc.
chair@oopsla.org

### PROGRAM CHAIR
**GREGOR KICZALES**
University of British Columbia
papers@oopsla.org

### ONWARD! CHAIR
**DIRK RIEHLE**
SAP Research
onward@oopsla.org

### CALL FOR PAPERS
March 19, 2008
Due date for Research Program, Onward!, Development Program, Educators' Symposium, Essays and proposals for Tutorials, Panels, Workshops and DesignFest

July 2, 2008
Due date for Development Program Briefs, Doctoral Symposium and Student Volunteers


Association for Computing Machinery

**NASHVILLE CONVENTION CENTER, NASHVILLE, TN**
**October 19 - 23, 2008**

# Symposium on Computer Human Interaction for Management of Information Technology

CHIMIT 08

*A Turning Point in IT*

November 14/15, 2008 San Diego, CA



## Dates

**Papers Submission**
April 7th, 2008

**Posters and Experience Reports**
May 30th, 2008

## Invitation to Participate

CHIMIT is the leading forum for discussing topics on IT management with a focus on people, business, and technology. Researchers and practitioners are invited to share issues, solutions, and research drawing upon fields such as human-computer interaction, human factors, computer systems, and management and service sciences.

## Topics

Original unpublished contributions are sought in areas including but not limited to:

***Workspace Studies***
- Ethnographic studies of IT work in context
- Patterns of work for various IT processes

***Processes and Practices***
- Development and use of processes in IT
- Impact of business decisions on IT

***Organizational Knowledge***
- Studies of collaboration and coordination
- Knowledge management and training

***Design***
- Design of human-centered IT systems
- Architectural considerations for user experience

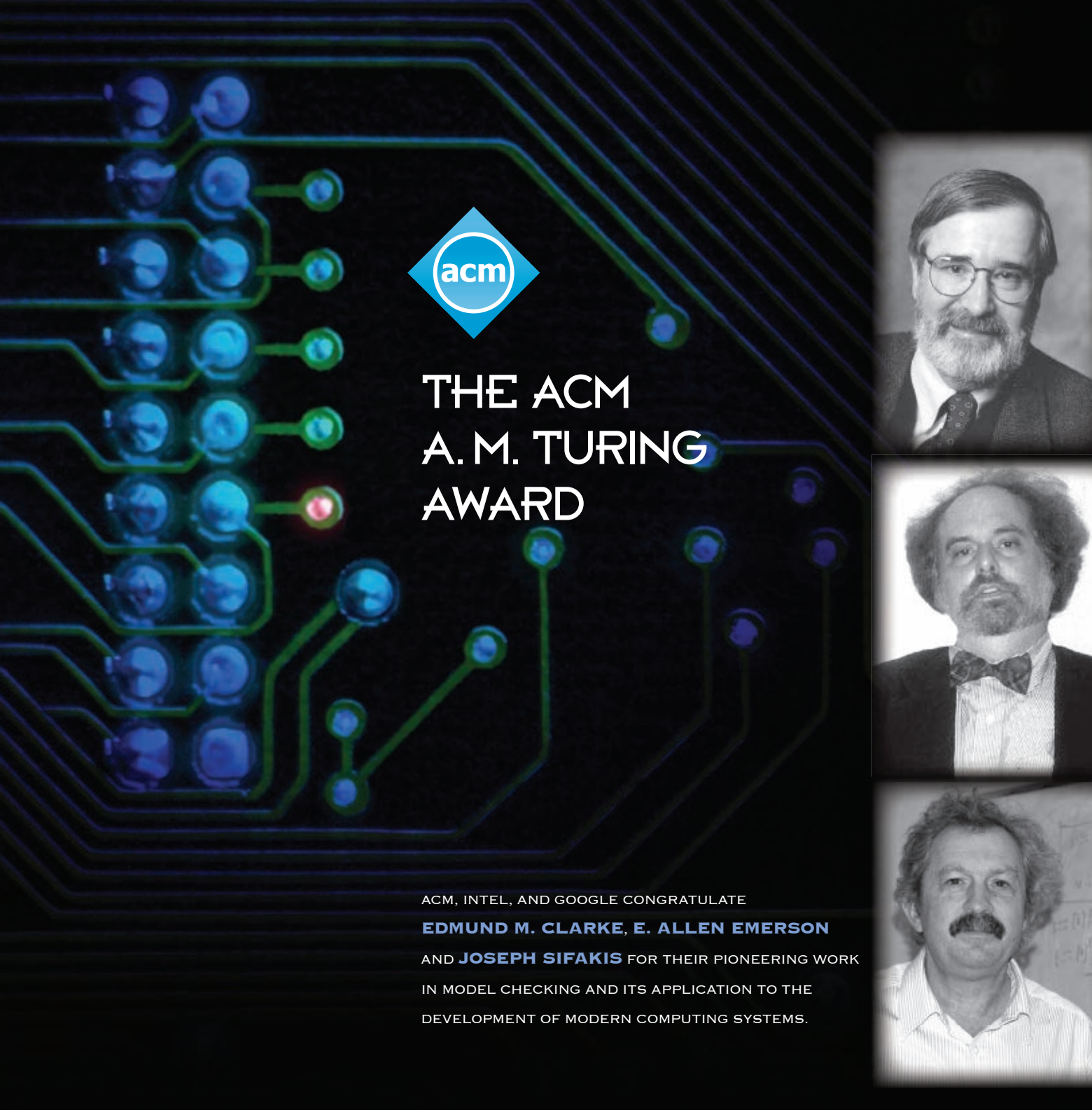
***Tools and Techniques***
- Collaborative system administration workspaces
- Visualizations of complex system behavior

***Automation***
- Automation/Policy languages
- Human interfaces to automation

*In cooperation with* USENIX

**www.chimit08.org**

November 14/15, 2008 | San Diego, CA



# THE ACM A. M. TURING AWARD







ACM, INTEL, AND GOOGLE CONGRATULATE

**EDMUND M. CLARKE**, **E. ALLEN EMERSON**

AND **JOSEPH SIFAKIS** FOR THEIR PIONEERING WORK IN MODEL CHECKING AND ITS APPLICATION TO THE DEVELOPMENT OF MODERN COMPUTING SYSTEMS.

"Intel and the entire computing industry have been direct beneficiaries of the Clarke, Emerson, and Sifakis' ground-breaking work, and insights from their novel automatic verification results have been widely adopted by the industry. These model checking approaches provide dramatically better coverage when searching for design errors. We are pleased to join the community in congratulating this year's A. M. Turing Award winners. For more information see www.intel.com\research."

**Andrew A. Chien**
**Vice President, Corporate Technology Group**
**Director, Intel Research**

"Google is delighted to spotlight great work in computer science, especially when it has a positive impact on our industry and society. The Turing award, which celebrates work that has greatly improved the quality of hardware and software through theory and systems research, is right in line with our own values. For more information see http://www.google.com/corporate/index.html."

**Alan Eustace**
**Senior Vice President**
**Engineering & Research, Google**

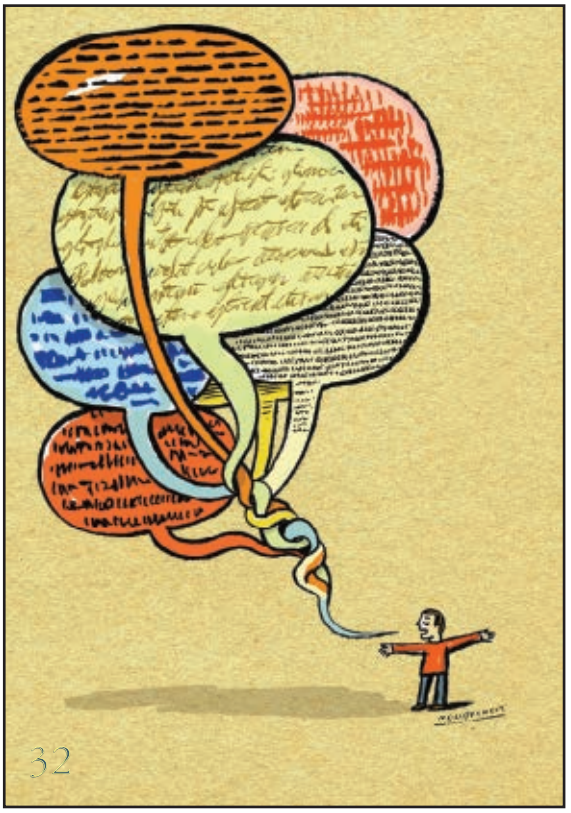
Financial support for the ACM A. M. Turing Award is provided by Intel Corporation and Google.

# Table of Contents

May 2008

## Articles







## Columns



## Departments




**ACM ELECTION BALLOT** P. 22


COVER ILLUSTRATION: ROBERT NEUBECKER

# Tribute to Honor Jim Gray

May 31, 2008
University of California, Berkeley

*A Tribute Honoring Jim Gray:*
Legendary computer science pioneer, known for his groundbreaking work as a programmer, database expert, engineer, and his caring contributions as a teacher and mentor.

## *General Session*

Zellerbach Hall, UCB
9:00am – 10:30am

Speakers:

Shankar Sastry
Joe Hellerstein
Pauline Boss
Mike Olson
Paula Hawthorn
Mike Harrison
Pat Helland
Ed Lazowska
Mike Stonebraker
David Vaskevitch
Rick Rashid
Stuart Russell

*All are welcome.*
*Registration is not required.*

## *Technical Session*

Wheeler Hall, UCB
Please see website for session times.

Presenters:

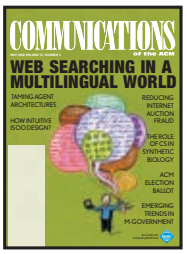
Bruce Lindsay
John Nauman
David DeWitt
Gordon Bell
Andreas Reuter
Tom Barclay
Alex Szalay
Curtis Wong
Ed Saade
Jim Bellingham

*All are welcome.*
*Registration is required, see below.*

*Technical Session registration and additional information:*

http://www.eecs.berkeley.edu/ipro/jimgraytribute

# Editorial Pointers




THE WORLD WIDE WEB MAY BRING THE WORLD TO our fingertips, but for tens of millions, some of the greatest wonders of the Web still require proficiency in English—the dominant language for information seeking. Progress has been slow building search portals in other languages, but with online populations in some regions experiencing triple-digit growth, demand is pushing innovation in a big way.

Wingyan Chung reviews the latest advances in Web search engines in a multilingual world in this month's cover story. He examines three prototype Web search portals—in Chinese, Spanish, and Arabic—and reveals the strengths and weaknesses of each and their potential for opening up informational treasures like never before. He also offers valuable guidance on how systems developers and managers, as well as users, worldwide can best support non-English Web searching.

ALSO THIS MONTH, MANY GOVERNMENTS WORLDWIDE EMPLOY WIRELESS connectivity to extend their services among the citizenry. Silvana Trimi and Hong Sheng consider the advantages and implications of m-government, exploring notable projects, and nations, leading the charge.

Dov Dori details a conceptual modeling framework—Object-Process Methodology—that employs graphics and text to help alleviate cognitive loads. And Irit Hadar and Uri Leron question the intuitiveness of object-oriented design, citing research that traces the inevitable clash between intuitive and analytical modes of thinking.

Auction houses appear to be at a crossroads, with many buyers and sellers losing confidence in the system. Auction fraud is an escalating concern, not only to consumers but to the U.S. Federal Trade Commission where reports of fraudulent transactions grew from 106 in 1997 to more than 24,000 in 2007. Bezalel Gavish and Christopher Tucci examine this trend, focusing on how some buyers are being swindled.

While software configuration management (SCM) practices may not support change management, Mohan, Xu, and Ramesh report all is not lost. Their article indicates how to integrate traceability practice with SCM practice to improve change-management processes in software development.

In "Viewpoint," former *Communications* editor-in-chief Jacques Cohen urges CS support for the emerging field of synthetic biology, emphasizing the idealistic goals of the field will not be feasible without the engaged contribution of computer scientists.

Finally, meet the candidates running for ACM's 2008 General Election, beginning on page 22.

Diane Crawford

EDITOR

DOI: 10.1145/1342327.1342328

COMMUNICATIONS of the ACM

A monthly publication of
the ACM Publications Office

**ACM**
2 Penn Plaza, Suite 701
New York, NY 10121-0701 USA
(212) 869-7440 FAX: (212) 869-0481

**Group Publisher:** Scott Delman

**Editor:** Diane Crawford
**Managing Editor:** Thomas E. Lambert
**Senior Editor:** Andrew Rosenbloom
**Editorial Assistant:** Zarina Strakhan
**Copyright:** Deborah Cotton

**Contributing Editors**
Phillip G. Armour; Hal Berghel; Michael A. Cusumano; Peter J. Denning; Robert L. Glass; Seymour Goodman; Rebecca Mercuri; Peter G. Neumann; Pamela Samuelson; Meg McGinity Shannon

**Art Director:** Caren Rosenblatt
**Production Manager:** Lynn D'Addesio

**Advertising**
ACM Advertising Department
2 Penn Plaza, Suite 701, New York, NY 10121-0701
(212) 869-7440; Fax: (212) 869-0481
**Director of Media Sales:**
Jonathan M. Just: jonathan.just@acm.org

For the latest media kit—including rates—contact:
Graciela Jacome: jacome@acm.org

**Contact Points**
**CACM editorial:** crawford_d@acm.org
**Copyright permission:** permissions@acm.org
**Calendar items:** calendar@acm.org
**Change of address:** acmcoa@acm.org

**Communications of the ACM**
(ISSN 0001-0782) is published monthly by the ACM, 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

*Advancing Computing as a Science & Profession*

**COMMUNICATIONS** OF THE ACM • *A monthly publication of the ACM Publications Office*

ACM, 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA (212) 869-7440 FAX: (212) 869-0481

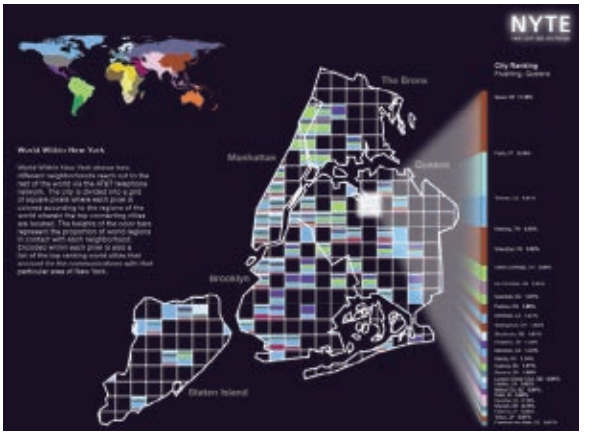
# News Track

### Enrollment Levels

Enrollment in computer science programs may have leveled off after the dot-com downturn, but that leveling happened only after the number of bachelor's degree graduates hit a trough, reports *Computerworld*. In the latest statistics from the Computing Research Association (www.cra.org), which follows year-over-year enrollment and graduate trends from 170 Ph.D.-granting institutions, only 8,021 students graduated with computer science degrees from these schools in the 2006–2007 academic year. By contrast, in 2003–2004—the high point of this decade—14,185 students were awarded bachelor's degrees in computer science. The sharp decline in graduates may be about to level off, according to CRA's latest trend analysis. In the fall of 2006, new CS enrollments topped out at 7,840; new enrollments for fall 2007 were at 7,915. While it's too early to declare a turnaround, CRA analysts say the students should be able to find job opportunities based on the last projections from the U.S. Bureau of Labor Statistics (see box on page 8).

### New York's Global Links

MIT researchers have been collecting the electronic communications of millions of New Yorkers since last January. The round-the-clock effort has nothing to do with national security or covert eavesdropping. Rather it is to build a census that shows, neighborhood by neighborhood, the phone and Internet links New Yorkers have to other cities across the planet and how these connections change over time. "Our cities and the globe are blanketed with flowing bits of digital data, and looking at this data, we're able to better understand the physical world," says Carlo Ratti, director of MIT's SENSEable City Lab. The Associated Press reports that no information about individuals or actual conversations is being collected. AT&T gives MIT only aggregate data from its switches throughout the city. Researchers found New Yorkers who engage in global gab tend to be international business professionals or poor immigrants. Moreover, communication between




Manhattan and the world surges each morning after the New York Stock Exchange opens. The most-called city is London (8% of all overseas calls), followed by Santo Domingo (5%). Half of all calls from Manhattan are to Canada, Great Britain, the Dominican Republic, Germany, and Japan. Visualizations from the New York Talk Exchange (sensible.mit.edu/nyte) project are now part of an exhibit at the Museum of Modern Art called "Design and the Elastic Mind," which is open through May and examines how designers use technology in ways that change lives.

### Etch-a-Circuit


Researchers have demonstrated a new technique inspired by the classic Etch-a-Sketch drawing toy that could be used to create rewritable logic circuits and denser computer memory. *Technology Review* reports that researchers, using an atomic force microscope (AFM), were able to draw nano-size wires and dots that could be repeatedly erased and written. The researchers used an AFM tip like a pencil,

Bottom: Jeremy Levy

# News Track

### Jobs to the Rescue

The U.S Bureau of Labor Statistics released its latest projections of jobs that will experience strong growth and be in great demand through 2016 (www.bls.gov/). The top five jobs in the IT field are:

Network systems and data communications analysts
Growth by 2016: 53%
Median Annual Salary: $64,600
Computer applications software engineers
Growth by 2016: 45%
Median Annual Salary: $79,780
Database administrators
Growth by 2016: 29%
Median Annual Salary: $64,670
Computer systems software engineers
Growth by 2016: 28%
Median Annual Salary: $85,370
Network and computer systems administrators
Growth by 2016: 27%
Median Annual Salary: $62,130

drawing electrically conductive paths on special material. The lines were as thin as three nanometers, making them considerably narrower than the lines drawn using electron-beam lithography—one of the most precise techniques for etching devices out of silicon. The study, recently published in *Nature Materials*, found that the wires and dots stayed in their state for at least 24 hours. The research team believes the etchings will last much longer and is currently testing this theory.

### Doctor Verboten

At least seven U.S. researchers working in Germany have faced criminal probes this year for using the title "Dr." on their business cards, Web sites, and résumés, the *Washington Post* reports. While all hold doctoral degrees from elite U.S. universities, under a little-known Nazi-era law only people who earn Ph.D.s or medical degrees in Germany are allowed to use "Dr." as a courtesy title. The law was modified in 2001 to extend the privilege to degree-holders from any country in the European Union, but Ph.D.s from the U.S. or elsewhere outside of E.U. are still forbidden to use it. Violators can face a year behind bars. While the German doctor rule has been in effect since the 1930s, it has been only sporadically enforced over the decades. What sparked a tipster to file a complaint with German prosecutors against seven U.S. researchers working at the Max Planck Society, which operates 80 scientific research institutes across Germany, is unknown. The criminal investigations have alarmed higher-education officials in Germany, where U.S. researchers are in demand. "This is a completely overdone, mad, and an absolutely ridiculous situation," declared the head of Germany's central office for foreign education. At last report, prosecutors were recommending that charges be dropped and civil fines be imposed.

### Clarke's Three Laws

- When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.
- The only way of discovering the limits of the possible is to venture a little way past them into the impossible.
- Any sufficiently advanced technology is indistinguishable from magic.

*First written in 1962 by legendary science fiction writer* Arthur C. Clarke *who passed away in March at the age of 90.*

Amy Marash

Send items of interest to cacm@acm.org

DOI: 10.1145/1342327.1342329


# Just Say 'A Class *Implements* a Data Type'

I agree with Chenglie Hu's "Viewpoint" "Just Say 'A Class Defines a Data Type'" (Mar. 2008) that data types are of central importance in computer science and should be emphasized throughout the curriculum. In CS 2, I always dedicate a good number of lectures to abstract data types and their (algebraic) specification. Some students are initially resistant, but eventually most recognize that an "interface" in Java's sense is not enough to explain what, say, a priority queue is, whereas an "implementation" (or "class") provides too much detail.

Motivating ADTs by examining "basic types" in a more abstract light, separated from their (accidental) properties in a given language or machine, seems to work well. One of the few advantages of C++ over Java is that "user-defined types" can be made to look exactly like "basic types," thus allowing for continuity.

A word of caution: Despite the educational and practical value of ADTs, an emphasis on them may be seen as "too theoretical" by some. If you indeed consider this route, be prepared for the inevitable discussions.

Finally, I offer three minor disagreements with Hu: Algorithms and data types are a "holistic whole," so there's no need to separate them; variables are instances of an ADT; and just say "a class *implements* a data type."

**Peter H. Froehlich**
*Baltimore, MD*

Chenglie Hu's point about "toy" objects in his "Viewpoint" (Mar. 2008) resonated with me. At the time I graduated from the U.S. Navy's Nuclear Power Program in 1983, the classroom part of the course was a six-month cram session of 76-hour (or more) weeks. We spent the first three learning or reviewing a classic high school science and math curriculum, then moved into the physics, thermodynamics, chemistry, and mechanical engineering needed for us to understand and operate the systems comprising a nuclear-reactor propulsion environment. The only way to coherently tie all these topics together while moving this quickly was to use real data and real examples from working reactors. From beginning to end, we were exposed to real-world examples in our studies.

Even with real data, 75% of the people entering the program did not complete it. I can't imagine even trying if we had only toy data and examples. I encourage all programmers and educators to weigh the benefits of using real-world examples in their courses in the interests of allowing a quicker and more complete understanding of the material.

**Jack Cain**
*Santa Cruz, CA*

PAUL WATSON

I want to express my appreciation to Chenglie Hu for his "Viewpoint" (Mar. 2008). My experience trying to switch from the "pure-OOP" approach in CS1 has been an uphill battle. For example, my department still uses Byron Weber Becker's *Java: Learning to Program with Robots* as the textbook for our CS1 students; while its value is recognized, it blurs most of the important elements for understanding how computers actually work. For example, how can these students be expected to "understand" a constructor with parameters before finally getting around to variables and eventually to methods in Chapter 6? One result is that we must find another way to teach them in CS2. I don't mind having to use object-oriented programming in CS1, but we should not pretend that the algorithmic side of the road doesn't also lead to solutions to real problems. As it is, even after a good part of the semester, the students might have learned only to play a poorly designed video game.

Chung-Chih Li
*Bloomington, IL*

I am an adult learner (age 57) just starting a Java programming course at the University of Phoenix, online. The first assignment was "Describe the difference between an object and a class." As you can imagine, the answers and examples were all over the place. After reading Chenglie Hu's "Viewpoint" (Mar. 2008), I understand why he would not start students out with a discussion about objects and classes.

Of special interest to me was the approach he suggested for teaching programming. Has he used such an approach with his own students? Is it still too early to think about using it for teaching a class? Maybe Hu can write his own textbook to spell it out in detail.

Although I'm likely to have graduated before any changes are made that would help me learn programming, I want to know how this idea plays out so I could make appropriate suggestions to my university. The online discussions among students who have finished the Java courses there suggest they were taught in the best way possible.

I also checked out Hu's Web pages at cscserver.cc.edu/chu/ and was amused to find that some of his students "might not have happily learned." I wish I could afford the luxury of attending classes in person instead of fitting them in online around a full-time job and family. I'm always happy to learn. The biggest lesson is that although the amount of knowledge I never will possess is enormous, I will never be deterred from trying to learn more.

Mark Hanna
*Banks, OR*

***Author's Response:***
In their comments, Froehlich (while I respect his objections) and Li shared what has worked and what hasn't in their respective classes. Cain shared an instance of how important real-world examples are to the learning process. And Hanna, as a novice, was still wondering about the best way for him to learn programming. The common thread is that educators must find ways to give their programming students a coherent learning experience. At the very least, they, along with textbook authors, must provide examples they are most likely to benefit from in their software-engineering classes.

Chenglie Hu
*Waukesha, WI*

**ENFORCE USER COMPLIANCE AS A FIRST LINE OF DEFENSE**
The article "The Myths and Truths of Wireless Security" by Alfred Loo (Feb. 2008) put too much emphasis on passwords. End users are known for their inability to use decent-strength passwords. Even expecting them to care about passwords aims too high. The IT staff must be reminded over and over of such user attitudes and behavior.

When my end users request changes in their passwords, they always want to know the minimum password length and whether they must include nonalphabetic characters. I tell them to use whatever password they want as long as it doesn't have to be written down. I also tell them to use at least one element that changes between servers. The reason for not having to write it down is that most intrusions involving password guessing come from colleagues in the same organization, not from "professional" attackers. Meanwhile, they resist even basic schemes that go beyond the old (flawed) login+password schemes.

Confronted by security challenges, time-based tokens, one-time passwords, and the like, they politely listen to administrators'

security-minded arguments, then dismiss them as not feasible, imagining we don't know the real situation or the true nature of an attacker's motivation.

They think they simply don't need a stronger identification scheme. In Mexico, banks are required to have such a scheme, though when money is not directly involved, users have been known to throw tantrums so they can keep using their everyday passwords.

Seeking to access a certain user's data, an attacker might try to find vulnerabilities in the host system, revealing data about every user in the system, rather than guess probable passwords based on a particular user's public profile.

GUNNAR WOLF
*Mexico City, Mexico*

***Author's Response:***

User resistance to doing more to secure a computer system is common. Ways to overcome it and enforce the organization's security policies depend on support from top management, including the CEO.

Malicious hacking takes multiple forms, each involving its own level of technical skill and preferences for attack methods. However, none can ever be ignored. A single security breach might be fatal to an organization, as I described in my article.

ALFRED LOO
*Hong Kong*

## CLARIFY OBJECT IDENTITY IN OBJECT SCHIZOPHRENIA

The article "Patterns, Symmetry, and Symmetry Breaking" by Liping Zhao (Mar. 2008) led me to explore the relationship between the object schizophrenia problem (OSP) and the symmetry-breaking concept it described. OSP in classical object-oriented modeling can be understood as broken semantics for method execution, contracts, and other concerns. (For more on OSP, especially in the context of subject-oriented programming, see www.research.ibm.com/sop/sopcpats.htm.)

The notion of object identity is central to OSP, as the semantics become clearer only when we consider object identity. An object has a set of operations and a state that remembers the state. OSP emerges mainly as a result of message-forwarding mechanisms (such as consultation and delegation), leading to "memory loss," or broken semantics. In delegation, the inability to associate "self" with the appropriate object (method holder or originator) results in broken delegation and (consequently) in OSP.

In the context of the article, OSP can also be understood as symmetry breaking, especially in terms of object identity. Moreover, the notion of object identity—crucial to OSP—might lead to a clearer understanding (and possibly a formal approach) when applied to design patterns and symmetry breaking.

AGNEESWARAN VIJAY SRINIVAS
*Lausanne, Switzerland*

## KEEP FAITH IN AMBIENT INTELLIGENCE

Ambient intelligence (AmI) is often promoted as a breakthrough despite its potential drawbacks. The article "The Illusion of Security" by David Wright et al. (Mar. 2008) analyzed ways to prevent the fictional threat scenario it presented and if indeed the scenario did occur ways its consequences might be addressed. That scenario—three fraudulent insider data administrators working in concert and senior executives trying to conceal their attack—is not AmI-related. Real-world scenarios with more disastrous consequences could occur on any given day in any given bank or critical-infrastructure-related governmental office.

A malicious insider group operating inside such an organization is highly improbable, as its conspirators would be expected to be well-screened, highly paid, and backed up by trusted endorsers.

Meanwhile, the type of attack described in the article is a *force majeure* for any organization—practically impossible to prevent once the conspirators have agreed to cooperate.

The three administrators and the company president described in the article fled to an undeveloped country with no AmI infrastructure. Our hope for a secure society with less of a chance of AmI-related disasters shouldn't depend on downgrading our interest in developing vibrant AmI technology wherever we are.

JAAK TEPANDI
*Tallinn, Estonia*

Please address all Forum correspondence to the Editor, *Communications of the ACM,* 2 Penn Plaza, Suite 701, New York, NY 10121-0701; email: crawfordd@acm.org.

DOI: 10.1145/1342327.1342330

# *ACM, Uniting the World's Computing Professionals, Researchers, Educators, and Students*

Dear Colleague,

At a time when computing is at the center of the growing demand for technology jobs worldwide, ACM is continuing its work on initiatives to help computing professionals stay competitive in the global community. ACM delivers resources that advance computing as a science and profession.

As an ACM member you have access to leading-edge resources including publications and conferences covering the entire computing field. Whether you are pursuing scholarly research, building systems and applications, or managing computing projects, ACM offers opportunities to advance your interests.

## MEMBER BENEFITS INCLUDE:

- Access to ACM's **new Career & Job Center** offering a host of exclusive career-enhancing benefits
- **Free e-mentoring services** provided by MentorNet®
- **Full and unlimited access to 2,500 online courses** from SkillSoft®
- **Full and unlimited access to 600 online books** from Safari® Books Online, featuring leading publishers, including O'Reilly (Professional Members only)
- **Full and unlimited access to 500 online books** from Books24x7®
- A subscription to ACM's flagship monthly magazine, ***Communications of the ACM***
- The option to subscribe to the full **ACM Digital Library**
- The **Guide to Computing Literature**, with over one million searchable bibliographic citations
- The option to connect with the **best thinkers in computing** by joining **34 Special Interest Groups** or **hundreds of local chapters**
- **ACM's 40+ journals and magazines** at special member-only rates
- ***TechNews***, ACM's tri-weekly email digest delivering stories on the latest IT news
- ***CareerNews***, ACM's bi-monthly email digest providing career-related topics
- ***MemberNet***, ACM's e-newsletter, covering ACM people and activities
- **Email forwarding service & filtering service**, providing members with a free acm.org email address and high-quality **Postini** spam filtering
- And much, much more!

ACM's worldwide network of more than 87,000 members range from students to seasoned professionals and includes many of the leaders in the field. ACM members get access to this network and the advantages that come from their expertise to keep you at the forefront of the technology world.

Please take a moment to consider the value of an ACM membership for your career and your future in the dynamic computing profession.

Sincerely,

Stuart I. Feldman

Stuart I. Feldman
President
Association for Computing Machinery

Association for Computing Machinery

*Advancing Computing as a Science & Profession*

Association for Computing Machinery

*Advancing Computing as a Science & Profession*

# membership application & *digital library* order form

Priority Code: ACACM28

## You can join ACM in several easy ways:

**Online**
*http://www.acm.org/join*

**Phone**
*+1-800-342-6626 (US & Canada)*
*+1-212-626-0500 (Global)*

**Fax**
*+1-212-944-1318*

**Or, complete this application and return with payment via postal mail**

**Special rates for residents of developing countries:**
*http://www.acm.org/membership/L2-3/*

**Special rates for members of sister societies:**
*http://www.acm.org/membership/dues.html*

***Please print clearly***

Name

Address

City | State/Province | Postal code/Zip

Country | E-mail address

Area code & Daytime phone | Fax | Member number, if applicable

## Purposes of ACM

ACM is dedicated to:
1) advancing the art, science, engineering, and application of information technology
2) fostering the open interchange of information to serve both professionals and the public
3) promoting the highest professional and ethics standards

*I agree with the Purposes of ACM:*

*Signature*

ACM Code of Ethics:
http://www.acm.org/serving/ethics.html

## choose one membership option:

### PROFESSIONAL MEMBERSHIP:

- ❑ **ACM Professional Membership: $99 USD**
- ❑ **ACM Professional Membership plus the ACM Digital Library: $198 USD ($99 dues + $99 DL)**
- ❑ **ACM Digital Library: $99 USD (must be an ACM member)**

### STUDENT MEMBERSHIP:

- ❑ **ACM Student Membership: $19 USD**
- ❑ **ACM Student Membership plus the ACM Digital Library: $42 USD**
- ❑ **ACM Student Membership PLUS Print *CACM* Magazine: $42 USD**
- ❑ **ACM Student Membership w/Digital Library PLUS Print *CACM* Magazine: $62 USD**

**All new ACM members will receive an ACM membership card.**
**For more information, please visit us at www.acm.org**

Professional membership dues include $40 toward a subscription to *Communications of the ACM*. Member dues, subscriptions, and optional contributions are tax-deductible under certain circumstances. Please consult with your tax advisor.

### RETURN COMPLETED APPLICATION TO:

Association for Computing Machinery, Inc.
General Post Office
P.O. Box 30777
New York, NY 10087-0777

Questions? E-mail us at acmhelp@acm.org
Or call +1-800-342-6626 to speak to a live representative

**Satisfaction Guaranteed!**

## payment:

Payment must accompany application. If paying by check or money order, make payable to ACM, Inc. in US dollars or foreign currency at current exchange rate.

❑ Visa/MasterCard ❑ American Express ❑ Check/money order

❑ Professional Member Dues ($99 or $198) $ _____

❑ ACM Digital Library ($99) $ _____

❑ Student Member Dues ($19, $42, or $62) $ _____

**Total Amount Due** $ _____

Card # Expiration date

Signature

# ACM A.M. TURING AWARD NOMINATIONS SOLICITED

Nominations are invited for the 2008 ACM A.M. Turing Award. 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 award carries a prize of $250,000 and the recipient is expected to present an address that will be published in an ACM journal. Financial support of the Turing Award is provided by the Intel Corporation and Google Inc.

**Nominations should include:**

1) A curriculum vitae, listing publications, patents, honors, other awards, etc.

2) A letter from the principal nominator, which describes the work of the nominee, and draws particular attention to the contribution that 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 usually include substantive letters of support from a group of prominent individuals broadly representative of the candidate's field.

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 at: http://awards.acm.org/homepage.cfm?srt=all&awd=140.

**Nominations should be sent by October 31, 2008 to:**

Professor Brian Randell
School of Computing Science
Newcastle University
Newcastle upon Tyne, NE1 7RU, UK
brian.randell@ncl.ac.uk

## 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 M. Clarke
2007 E. Allen Emerson
2007 Joseph Sifakis

DOI: 10.1145/1342327.1342331

# The Crucial Role of CS in Systems and Synthetic Biology

## Biological cells, coerced to function as hardware and driven by artificial DNA, can perform such nanoscale tasks as detecting toxic substances and manufacturing new drugs.

In the coming decades the hardware-software paradigm that has been central to computer science since its inception may well be challenged—or at least complemented—by an exciting new development: synthetic biology. Biological cells will become an alternative to current hardware, and an analog of software will be engineered to direct cells to produce useful artifacts or substances. The hardware-software model will thus come increasingly close to mimicking miniature (nanosize) robotics that induce living organisms, like bacteria, which have existed in nature for billions of years, to assemble minuscule amounts of compounds. Scientists would thus be able to perform tasks that are still largely unimaginable today, like cleaning the environment, making new drugs, and detecting dangerous chemicals.

Computer science has always been in the vanguard of new scientific and engineering developments. The one I advocate here will help us move into the next frontiers. The emerging field of synthetic biology [1, 3, 4] is the engineering counterpart of designing computer-like biological machinery and its associated software, or wetware. Wetware indicates how programs that assemble the desired artifacts are constructed in a biological wet-lab, in contrast to the dry-lab used to assemble and program electronic components.

The goals of synthetic biology can be accomplished only if biologists and computer scientists fully comprehend the dynamic behavior of living cells. A discipline called systems biology [5], which encompasses synthetic biology, strives to understand dynamic cell behavior. Systems biology is essentially analytical, whereas synthetic biology deals with the engineering issues of changing known cell behavior to accomplish human goals.

My aim here is twofold: urge computer scientists to follow closely what is happening in synthetic and systems biology, participating actively in their development; and emphasize that the idealistic goals of systems and synthetic biology will not be feasible without the engaged contribution of computer scientists. Fulfilling these objectives will open inspiring new vistas for our field.

To illustrate the promise of synthetic and systems biology, consider a prototype experiment that demonstrates what is currently feasible in these new fields. A Petri dish containing a special strain of harmless bacteria sits on a laboratory bench at a research center. A slight fragrance of mint emanates from the dish. A scientist in a lab coat adds a few drops of a chemical, and the mint scent is quickly replaced by a strong odor of...bananas (web.mit.edu/newsoffice/2006/igem.html).

LISA HANEY

A comparable event occurred at the 2006 International Genetically Engineered Machines (iGEM) competition, which had brought together graduate students from 15 countries majoring in biology, computer science, and electrical engineering (parts2.mit.edu/wiki/index.php/Main_Page). The apparently innocuous test of changing a substance's fragrance by adding special ingredients has enormous practical implications, most notably in environmental remediation and pharmaceutical engineering. Imagine that the ingredient being added to the dish contains a dangerous chemical. The experiment allows investigators to detect the presence of the substance simply by exposing it to bacteria and having humans, animals, or even instrumentation detect the chemical changes generating the fragrance.

What does this have to do with computer science? The question and its answer are at the heart of synthetic biology. Its objective is to reengineer cells by changing or supplementing their DNA so the ensuing cell products are sensitive to given substances and can indicate and eventually eradicate their presence. For example, synthetic biology could potentially engineer harmless bacteria capable of detecting and absorbing oil spills or breaking down carbon dioxide in the atmosphere.

The entity that associates synthetic biology to computer science is DNA, which can be viewed as a program that remains static or dormant in a computer-like memory. Only when it is executed by processors—the equivalent of interpreters and hardware—does its dynamic behavior come to life.

From its origins in the early 1980s, bioinformatics, combining computer science and biology, has dealt mostly with the static properties of DNA and its products, like RNA and proteins. With the automation of DNA sequencing in the mid-1980s, the immediate goal was to obtain sequences of letters—A, C, G, and T—identifying the basic nucleotides that characterize all living matter, from bacteria to humans. The project of sequencing a variety of genomes is still formidable; a recent press release reported "The theoretical price of having one's personal genome sequenced just fell from the prohibitive $20 million to about $2.2 million, and the goal is to reduce the amount further—to about $1,000—to make individualized prevention and treatment realistic" [7].

Following the principles of Darwinian evolution, bioinformatics specialists have scrutinized similarities among the static DNA of various species. In fact, nearly all research in similarities to date has been done at the static level, without great concern for the dynamics triggered when DNA is processed by actors, like polymerases and ribosomes. Such actors are essentially the nanobiological machinery that processes DNA to produce proteins—the building blocks of life.

Current efforts in bioinformatics also seek to determine protein structure and function. Most research has concentrated on identifying the stable 3D shape of a static molecule, even though protein molecules have degrees of flexibility that are relevant in determining a molecule's function. The concern for static sequences and 3D structures is still amply justified, since studying the dynamics of DNA and protein interaction is nearly impossible without a thorough study of their static counterparts.

Protein function is specified through informal natural language sentences that describe the role of a protein in a living cell. This description must still be complemented with formal specifications by, for example, indicating the protein's role in a network of protein interactions. Computer scientists are needed to design these specifications.

Systems biology studies the dynamic properties of the interactions between DNA and its products. Even if this new field is viewed as a branch of bioinformatics, it is already an area of significant interest to biologists, computer scientists, control engineers, and mathematicians. In dealing with static DNA and its products, including protein sequences, fundamental computer science algorithms operate on strings of symbols. They search for approximate patterns in very long sequences, compare multiple sequences, and combine overlapping sequences. Optimization is the objective of the approximate pattern-matching of sequences; the algorithms are designed to minimize the cost of abstractly transforming one sequence into another.

The goal of systems biology is to scrutinize the

dynamics of cell behavior. For example, a certain protein *P* (produced by a gene *G*) is used to prevent the production of another protein *P'* by blocking the processing of gene *G'*. The computer science analog is keeping some parts of a program from being executed once the execution reaches a certain stage. This behavior is akin to Edsger Dijkstra's semaphores in regulating the dynamic behavior of concurrent programs [2].

Both systems and synthetic biology will require expertise in computer science way beyond the expertise necessary for processing sequences. In turn, these disciplines will challenge computer scientists with problems that are at the forefront of computer science today, including how to develop nanotechnology hardware, fault-tolerant circuit design, program verification, model checking, program synthesis from data, and data mining.

Systems biology deals with gene interactions, some involving hundreds, if not thousands, of genes. When some interactions go awry, cell behavior changes dramatically, resulting in situations like the uncontrollable growth of a cancer. Thus, the computer algorithms in systems and synthetic biology are akin to those used for finding bugs or incorrect behavior in large complex programs. Debugging is one of the most arduous tasks in program development. Nonetheless, computer scientists have already developed sophisticated tools to facilitate debugging, some applicable for finding faulty configurations in biological networks.

A microarray, or genome chip, is a silicon chip that has become a common tool in systems biology. It includes tens of thousands of minute wells, each with multiple short strands of DNA material representative of each gene (www.affymetrix.com/index.affx). Each strand is matched with its counterparts obtained through a wet-lab experiment involving the genes of the cell being studied. The degree of matching is measured by special scanners linked to computers. The measurements estimate the amounts of products generated by a gene. Microarrays can also be used by biologists to dynamically record the changes in gene products over time, as a cell is subjected to some external influence (such as food starvation or the effect of a drug).

Microarrays are also expensive, each costing hundreds of dollars, and tens or possibly hundreds of them may be necessary to study the gene interactions of a single cell. But, as with sequencing costs, microarray costs are decreasing, and the volume of data being generated by microarray experiments is gigantic, possibly surpassing the size of the available DNA data. Even though results of microarray experiments are coarse and may contain laboratory errors, they represent challenging problems for data-mining experts.

The computer algorithms in systems and synthetic biology are akin to those used for finding bugs or incorrect behavior in large complex programs.

Several research groups led by prominent computer scientists have immersed themselves in fascinating research involving the amalgamation of computer science and systems biology. For example, a team at the Weizmann Institute led by Ehud Shapiro has designed nanobiological processors that function as finite-state-automata to recognize desired sequences of DNA and eventually deliver drugs capable of correcting cell behavior that could lead to disease (www.wisdom.weizmann.ac.il/math/profile/scientists/shapiro-profile.html) [8].

As mentioned earlier, synthetic biology is a notable leading-edge effort in systems biology. Its aim is to use the existing "processing" capabilities of a cell (such as yeast or *E. coli*) to perform such tasks as cleaning the environment, detecting dangerous chemicals, and manufacturing drugs. J. Craig Venter, a pioneer in sequencing the human genome, is pursuing the goal of generating (in a wet-lab) very long sequences of artificially produced DNA (www.jcvi.org). The artificial DNA is being designed to perform the tasks involved in reengineering cell behav-

ior. Computer scientists in synthetic biology teams help design the DNA fragments that must be inserted into a living cell and verify that the resulting genetic network is robust in the sense that small variations within the engineered cell mechanism are tolerated and will not result in malfunction.

The annual iGEM competition mentioned earlier is backed by a number of corporations, including Microsoft, which has established a systems biology group at its Research Center in Cambridge, England; researchers there explore the applications of Milner's π-calculus—developed to check the properties of mobile hardware, like cell phones—in systems biology. Ensuring that within a certain area a given number of calls can be handled properly by a provider's technology has counterparts in systems biology. As such, a team led by Luca Cardelli, a Microsoft computer scientist in Cambridge, who previously engaged in research aimed at ensuring the correctness of distributed programs, is today developing formal languages to describe cell behavior [6].

Even if systems and synthetic biology experiments look simplistic, they are indeed the early prototypes of major advances in the field. With the help of computer science, researchers and engineers will achieve the progress needed to transform systems and synthetic biology from pure science to industrial-scale reality. C

## References


1. Campbell, A.M. Meeting report: Synthetic biology jamboree for undergraduates. *Cell Biology Education 4,* 1 (Spring 2005), 19–23.
2. Dijkstra, E.W. Solution of a problem in concurrent programming control. *Commun. ACM 8,* 9 (Sept. 1965), 569.
3. Gardner, T.S., Cantor, C.R., and Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature 403,* 6767 (Jan. 2000).
4. Hopkin, K. Life: The next generation: Engineers and biologists team up to create synthetic biological systems. *The Scientist 18,* 19 (Oct. 2004).
5. Kitano, H. Systems biology: A brief overview. *Science 295,* (Mar. 2002).
6. Phillips, A., Cardelli, L., and Castagna, G. A graphical representation for biological processes in the stochastic pi-calculus. *Transactions in Computational Systems Biology LN in CS 4230* (Nov. 2006), 123–152.
7. Pollack, A. The race to read genomes on a shoestring, relatively speaking. *New York Times* (Feb. 9, 2008).
8. Shapiro, E. and Benenson, Y. Bringing DNA computers to life. *Scientific American 294,* 5 (May 2006), 45–51.



**Jacques Cohen** (jc@cs.brandeis.edu) is the TJX/Feldberg Professor of Computer Science in the Department of Computer Science at Brandeis University, Waltham, MA.





DOI: 10.1145/1342327.1342332

# This Menu Has Changed

## Innovative interfaces will give wireless the golden touch.

I admit it. I really wanted one. So much so that I was one of those half-crazy people trolling for treasures before dawn the day after Thanksgiving last year. "You don't have any left, do you?" I asked meekly, convinced my request would be met with uproarious laughter and public ridicule. "Actually, two," mumbled the sleep-deprived store worker. "But that man is buying one right now."

Arms outstretched, I quickly slid toward the lone, sleek white box with ethereal gray lettering and grabbed hold. I got it. I got it!

When the gift wrap was torn apart in chaos weeks later, the eager hands held that white box just as tightly as I had. "Wii!," three children and one husband exclaimed gleefully. It's been in near constant use ever since. Not just by immediate family, but grandparents, toddlers, teenagers. Anyone who gets their hands on it wants to keep playing. (I admit to having a sore shoulder after taking a game of tennis a tad too competitively.)

A big part of the attraction of Nintendo's crazily popular video

game, introduced a year and a half ago, is the system's motion-sensitive controllers. Rather than the thumb aerobics required of competing game systems, the Wii demands a more physical interaction. To play a game on the Wii the player has to move the console in such a way as to mimic real-life movements. For instance, in bowling, the remote is held, raised up, swung backward and finally swung upward, much like a bowler would move an actual bowling ball. And just like if your wrist tends to twist slightly at the release at the actual lanes, your ball is going to curve toward the gutter in Wii world, too. If you swing the remote a millisecond too late when playing baseball, you're going to get a strike, and if you're too strong with your golf swing, your golfball will land in a sand trap. Unlike the spectacular graphics of its peers, Wii's art and movements are simple and universal.

The Wii uses a wireless, Bluetooth-enabled battery-operated remote control unit that has motion-sensing ability. Meanwhile, the sensor bar that sits atop the television has 10 infrared LEDs spaced along the bar. Using accelerometer and optical sensor technology inside the remote, and tapping into Bluetooth, the motion is detected. The remote

LISA HANEY

control unit reads the movement's distance from the different lights on the bar, and then that motion is calculated through triangulation. Industry experts say it is the accelerometer that contributes most to the Wii's $250 price tag, which while costly is not as expensive as other gaming systems.

Likewise, even in an unstable economy, another high-priced gadget that caught consumers' interests and landed on their wish lists this past holiday season was Apple's iTouch. The mobile device's multi-touch technology had consumers buzzing and willing to spend nearly $500 before a price cut was announced.

These two devices are just two examples of renewed enthusiasm being pumped back into the consumer tech sector. Why? It's in the interface.

SUCH INNOVATIVE INTERFACES, LIKE THE WII'S MOTION-SENSITIVE CONTROLLER, HELP TO BREAK DOWN THE BARRIERS BETWEEN MAN AND MACHINE, SAY ANALYSTS, AND HELP TO FIRE UP APPLICATION DEVELOPERS.

## GAMES PEOPLE PLAY

Even venerable cable news channel CNN is not immune to the charms of an innovative interface. Watching "The Situation Room" during election coverage you'd think star reporter John King was as enamored of the multi-touch presentation screen as my family has been of the Wii. King could be seen narrowing in on a graphic with a pinching motion of his fingers, moving over delegate territory with a wave of his hand, making the colored map come to life. CNN might have called it "The Magic Wall" when it debuted in January during the Iowa caucuses, but readers of this magazine know it as multi-touch, which is credited to Jeff Han and a handful of companies for its development and evolution.

The influence of innovative interfaces on the wireless industry has been increasing for a while. Consider GestureTek, for instance. The Sunnyvale, CA-based company has taken its video gesture recognition patents and partnerships with gaming—Microsoft and Sony licensed their technology for Xbox and PlayStation respectively—and applied the technology to mobile devices and applications. The bridge between gaming and mobile devices comes because GestureTek's technology works with cameras, and most mobile devices have cameras installed. Deals with NTT DoCoMo and Verizon have helped to put GestureTek on the mobile map two years ago, even though they've been around for 20 years.

"It's disruptive technology," says Ed Fowler, director of business development at GestureTek, of the video gesture technology. "People see it once, and they get it." Fowler says that since wireless operators are drooling at opportunities to keep customers connected more, downloading more bits, the lure of fusing gaming components and smooth interfaces with mobility is enticing.

With its technology inside, GestureTek sees the phone of the near future being shaken or rolled slightly, like a gaming console, so users can quickly bring up information needed. There are also commercial applications that make sense, like navigation. But the ease of use of this type of interface could have social uses as well, says Fowler. An elderly person living alone, for instance, could shake their mobile phone once to dial 911, or could do physical therapy for their arms using their mobile device, Fowler says. The technology is also enabling digital signage, or interactive billboards.

## CAN'T TOUCH THIS

Apple's aggressive acceptance of multi-touch shows no signs of waning. In fact, recent reports say a next generation of the iPhone, bigger in size but still smaller than a laptop, will come out next year with even more multi-touch functionality. Fowler

adds that more tablet devices with shake and roll capabilities are in product development. Blackberry, too, looks like it will undergo an interface change. Industry reports say the device might go one better than the multi-touch with a "multi-pressure" touch screen makeover, due out next year. Meanwhile, also raising the interface bar, Microsoft filed a patent in October last year for "extensive filtered lists for mobile device user interfaces," where selected data and service provided by multiple software applications can be accessed through a group or list of items, according to reports.

Such innovative interfaces, like the Wii's motion-sensitive controller, help to break down the barriers between man and machine, say analysts, and help to fire up application developers. The act of scrolling through a hierarchy of menus is considered an outdated way to get information. Interface innovation, as exemplified by the iPhone, has positively influenced sales of mobile devices. The number of mobile phones sold in the U.S. increased 16% from the second to the third quarter of 2007, to the tune of $3.2 billion, according to NPD Group. Smartphones, specifically, saw a 163% increase year over year, the research firm said. The iPhone helped to bridge the gap between traditional phones and more business-oriented phones, the firm adds. By the end of January 2008, there were some 4 million iPhones sold, according to reports.

As the Consumer Electronics Show in January showcased, competitors aren't letting Apple do all the touching. Sleek devices, including universal remotes, got the multi-touch treatment. HTC Touch, Verizon's LG Venus (at $200) and Voyager to name a few have added to the realm of devices with these touchy-feely interfaces at price breaks. Devices like these helped sell 1.14 billion phones last year, up 12% from the previous year, according to reports.

At the 3GSM Mobile World Congress held in mid-February in Barcelona, industry bigwigs, like Arun Sarin, CEO of colossal carrier Vodafone, paid homage to Apple's influential interface. During his keynote speech at the show, Sarin told the audience to keep it simple and advised for a slash in operating systems. "The simpler we make it, the more we sell," Sarin says on a video blog on Vodafone's Web site (captured by a camera phone, by the way) taken just after the speech.

Once user-friendly fronts become the norm in mobile devices, operators will still have work to do. While the "Internet on the mobile is really beginning to happen," Sarin says, operators must innovate so they are not relegated to just becoming "bit pipes." He advised operators to consolidate the number of operating systems that are used, according to reports. His own company counts between 30 to 40 operating systems up and running on its network right now and he suggests the number be reduced to five or under, say reports. With fewer operating systems, developers can create more compelling applications to help drive usage.

Even as market watchers say eventually the number of operating systems will dwindle, increased interest in the OS space has drawn big names. Last fall Internet behemoth Google unveiled Android, its mobile operating system. The Linux-based mobile software solution is designed under the Open Handset Alliance. Other players in this area include Symbian, Linux, and Microsoft Windows Mobile.

While operating systems duke it out, the conference showed that content remains a key issue. That this year's show got some star power when movie mogul Robert Redford and actress Isabella Rossellini attended only underscores the importance of forging deeper relationships with the music and video industries to the mobile world.

Even with the voluminous expansion of content and move toward streamlining operating systems, with increasingly innovative interfaces becoming so addictive, it looks like whatever wireless brings us, getting there will be half the fun. C

**Meg McGinity Shannon** (shannon7575@verizon.net) is a technology writer based on Long Island, NY.



DOI: 10.1145/1342327.1342333

# *candidates for* **PRESIDENT** (7/1/08-6/30/10)

**WENDY HALL**

Professor of Computer Science

School of Electronics and Computer Science

University of Southampton

Southampton, UK

## BIOGRAPHY

■ Dr Hall received a BSc with honors (Mathematics, University of Southampton, 1974), PhD (Pure Mathematics, University of Southampton, 1977) and MSc (Computer Science, City University, London, 1986). Lecturer, Oxford Polytechnic (1977), Lecturer, LSU College of Higher Education (1978), Lecturer (1984), Senior Lecturer (1990), Professor (1994) Computer Science, Head of School (2002-2007), University of Southampton.

Dr Hall's team developed the well-known Microcosm open hypermedia system, which was patented and spun-off into a commercial company (1994), and won an ITEA award (1995) and a BCS IT award (1996). Current research interests include web science, the semantic web, and digital libraries. Dr Hall has published over 350 journal and conference papers, which can be found at http://www.ecs.soton.ac.uk/~wh/

One of the founding directors of the Web Science Research Initiative, a joint venture between the University of Southampton and MIT, launched in 2006. MIT Research Affiliate (2007-present).

CBE in Queen's Birthday Honours list (2000), EPSRC Senior Research Fellow (1996–2002), member of UK Technology Foresight Panel (1995–98), member of the council of EPSRC (1997–2002), member of executive of UK Computing Research Committee (2002–present), member of UK Prime Minister's Council of Science and Technology (2004-present), member of Scientific Council of the European Research Council (2005–present), chair of UKCRC Grand Challenges Committee (2005–present). Honorary DSc (Oxford Brookes University, 2002, University of Glamorgan, 2005, University of Pretoria, 2007), Honorary Fellow University of Cardiff (2004). Anita Borg Award for Technical Leadership (2006).

ACM Pubs Board (1999–2005), programme co-chair ACM Multimedia'96, conference co-chair ACM Hypertext'97, conference co-chair ACM Multimedia'98, chair WWW2006, executive committee SIGMULTIMEDIA (1998–2003), member SIGWEB and SIGMM.

President BCS (2003–2004), Senior VP Royal Academy of Engineering (2005-present), FREng (2000), FBCS (1996), FIEE (1997), FCGI (2002), CEng (1990). VP Publications BCS (1998–2002), member of IW3C2 (1997–present), member of British Library Board (2007-present), member of several editorial boards and numerous programme committees, non-executive director of several companies.

## STATEMENT

■ I am extremely honoured to have been asked to stand for election as President of the ACM. I have been a member of the ACM for most of my professional career, and I am an active member of SIGWEB and SIGMM. I served on the ACM Publications Board for six years and on ACM Council for 2 years before being elected ACM Vice President in 2006.

ACM is increasingly well established as the premiere computing society. Membership has grown steadily over the past five years, making ACM the largest computing society in the world. Our Digital Library is regarded as the best in computing. We are publishing more journals and running more conferences than at any point in ACM's history, and taking the lead on many important policy issues in the computing field. But to reach our full potential, there are three things I believe the ACM must do: 1) really grasp what it means to be an international society; 2) take the lead in terms of increasing diversity in all aspects of our field; and 3) provide something new and unique for ACM practitioners.

While ACM's impact outside the US is increasing, I believe we must do more in order to become the leading computing society worldwide. We need to develop further our initiatives in India and China, rethink our relationship with Europe and explore how to be more relevant in South America and other parts of the world. As a European, a Past President of the BCS and a researcher with many international connections, I believe I have the background, experience, and perspective to help lead ACM in these important efforts.

I have actively worked throughout my career to increase diversity in computing. I chair the BCS Women's Forum and the Royal Academy of Engineering's Diversity Committee, and have been on several government working parties in this area. As ACM Vice President I have taken an active role in the development of ACM-W to give it a greater voice in the organisation, and I am co-chairing the Presidential Committee on Gender Diversity. I am very keen to see this work come to fruition and to establish diversity as a fundamental part of the ACM agenda.

I am committed to continue the initiative ACM started three years ago to better serve the practitioners community. Queue magazine is a part of this, as is the new Professions Board. The successful future of the ACM depends on its ability to support the academic and research communities around the world, to reach communities typically underrepresented in computing, and to do something unique for practitioners both within and outside ACM.

The revitalization of CACM is a crucial element in the process of increasing our international presence and expanding the service we provide to all of our communities. It is critical that the next President works to ensure the success of this initiative. I believe my track record in academic publishing, my leadership experience within the UK and Europe, and my knowledge of, and experience in, ACM can benefit both ACM and the community.

**J STROTHER MOORE**

Computer Science – Department Chair

University of Texas at Austin

Austin, TX, USA

## BIOGRAPHY

■ J Strother Moore is CS department chair at UT Austin.

He got his SB degree in math from MIT in 1970 and his PhD from Edinburgh in 1973. He is an AAAI Fellow, an ACM Fellow, a member of the NAE, and a recipient of the ACM Software System Award.

Since his PhD he has spent 17 years in industry and 17 years in academia, successively at Xerox PARC, SRI, UT Austin, Computational Logic, Inc. (which he co-founded), and UT Austin again, where he now holds the Admiral B.R. Inman Centennial Chair in Computing Theory. His primary research is in automatic theorem proving and the verification of hardware and software but his contributions are broader.

Moore loves programming. He put himself through MIT working for the MIT Laser Research Group (where he wrote FORTRAN to solve differential equations), for IBM (page fault simulation), and for TRW Systems Group (debugging Apollo lunar orbit insertion and navigation). During his last year at MIT, he worked full time at State Street Bank and Trust (coding mutual funds services in PL/1). Later, he and Bob Boyer invented a structure shared approach to resolution theorem proving, which helped usher in "logic programming," and also a structured shared representation of edited text which became the basic representation in the Bravo and Word editors. Moore specified the INTERLISP virtual machine. Boyer and Moore are the co-authors of a fast string searching algorithm, a linear-time majority vote algorithm, and (with Kaufmann) a series of theorem provers used in commercial hardware and software verification. Moore led the effort to produce the first verified software stack and designed, implemented, and verified the assembler, linker, and loader for it. He led the effort to verify the floating point division microcode for the AMD K5 and other commercial projects.

Moore has been department chair for 7 years. He is on the Board of Directors of the Computing Research Association, is co-chair of the 2008 CRA Snowbird conference, and co-authored an influential CRA best practices paper on how universities should handle CS IP from industrially sponsored research (www.cra.org/reports/ip). Moore serves on the ACM Educational Policy Committee, is active in NCWIT, has co-chaired the Habermann Award committee for 3 years, and is on the advisory boards of the CS departments of Prairie View A&M, New Mexico State, and UT El Paso.

## STATEMENT

■ Computing is the transformative science of our age. Its effects on the other sciences, engineering, medicine, and business are acknowledged. But it will similarly transform all aspects of our lives and cultures in the decades ahead. Computing is transformative in part due to deep scientific reasons.

The workforce and diversity issues we face are exacerbated by the misconception—which we helped promulgate—that computing is merely a technological enterprise as opposed to a deep scientific one as well.

Our machines and their capabilities are among the most exciting intellectual ventures in history.

I want to lead the ACM because I am passionate about the transformative nature of computing and think the ACM is the organization that can get this vision out globally. But we must target the right audiences and deliver a compelling message. The right audiences are national leaders, industrial leaders, and academic leaders (including K-12). The message is that the science of computing will transform everything they care about, whether it is a national economy, a business paradigm, or the meaning of excellence in a given discipline.

Other initiatives that are important to me include: the cultivation of top-notch volunteer ACM leadership in India and China, broadening participation in the CS community at all levels, the branding of ACM as the unbiased source of trusted reports on computing issues of global importance (e.g., the globalization report), and the opportunity to become the publisher of choice in computing.

Most of my professional life has been as a programmer and researcher. But my biography supports my dedication to teaching, my abiding concern for opening the CS workforce to all, and my interest in the healthy collaboration of academia and industry.

I have 7 years of executive experience as chair of a large, top ranked CS department, and ten years of experience as Chief Scientist of a company I co-founded. I have served 3 years on the Board of Directors of the CRA. That experience has taught me some of the keys to institutional leadership.

A critical skill is listening more than one talks. Another is seizing the opportunity to identify and articulate shared goals. A third is assembling a trusted team of advisors and being open to their advice. A fourth is delegation of authority in areas where the institution already performs appropriately. And a fifth is leadership focus and follow-through on shared goals—even if it means foregoing potential new goals to achieve still-desired old ones.

I am realistic enough to understand that in two years it is not possible to accomplish all the initiatives above. But I can promise to bring the full force of my experience, skill, energy, and passion to the leadership of the ACM.

# *candidates for* VICE PRESIDENT (7/1/08-6/30/10)

**JOSEPH A. KONSTAN**

Professor of Computer Science and Engineering

University of Minnesota

Minneapolis, MN, USA

## BIOGRAPHY

■ Joseph A. Konstan earned an A.B. magna cum laude with highest honors in Computer Science from Harvard College in 1987; he earned an M.S. (1990) and Ph.D. (1993) in Computer Science from the University of California, Berkeley, with a dissertation on user interface toolkit software architecture, focusing on event handling, layout management, and data change propagation. He immediately joined the faculty of Computer Science at the University of Minnesota, where he was most recently promoted to Full Professor in 2005. He has served the University in several roles, including as Director of Graduate Studies for Software Engineering, DGS for Computer Science, and as the University's representative to (and vice-chair of) the Federal Demonstration Partnership (where he works to reduce administrative burden related to research grants).

Dr. Konstan is an ACM Distinguished Scientist, and has served as an ACM Distinguished Lecturer. He has authored over 85 peer-reviewed papers, articles, and book chapters, and holds six U.S. patents. His recent work spans three research challenges: Recommender Systems—the design and development of systems that personalize content based on user preferences; Online Community—designing the structure and implementation of online communities to better elicit user contributions; and E-Public Health—specifically a current clinical trial of an online "experience" designed to reduce HIV risk-taking in high-risk men.

Konstan is currently Chair of ACM's SIG Governing Board. In this role (and as part of ACM's Executive Committee) he has taken on challenges related to ACM-W and the support of women in computing and to the internationalization of ACM. Previously, he served as President of ACM SIGCHI, as vice-Chair of both the SGB and the Membership Services Board, and on the Executive Committee of ACM SIGMultimedia. He was general chair of UIST 2003 and ACM Recommender Systems 2007; program co-chair for ACM Multimedia 2000; doctoral symposium chair for three ACM conferences; and a volunteer in many roles for a variety of conferences.

## STATEMENT

■ I am honored by this nomination. Since first volunteering in 1994, I have enjoyed working with talented and dedicated ACM volunteers and staff. Since entering leadership five years ago, I've seen up-close many of the tremendous things we're doing for the field and our members.

From the Digital Library and Portal to our Professional Development Center, we provide great online resources for computing professionals.

Through CSTA and the Education Board and Council, we're nurturing a motivated and capable pipeline of students. Our SIG conferences are the technical and professional networking highlight for thousands. And our transactions, journals, and magazines are the best in the business.

At the same time, we must improve. We're investing in India and China—working with local leaders to understand how to support the advancement of the science and practice of computing. We've redesigned Communications of the ACM to make it more relevant to our membership, more timely, and more representative of the best work in all of computing.

I'd like to pursue four other important issues. First, our local activities program. While ACM has many chapters worldwide, our support for local activities and local volunteers needs enhancement: increased access to technical programs, better online tools to help local leaders manage membership and communications, and local volunteer development and support.

Second, better support outside North America. While our SIGs and staff have done an excellent job bringing our conferences around the world, we too often operate as an "American presence" in "foreign lands." We must gain the expertise, the tools, and the worldwide volunteer leadership to help us operate in locally appropriate ways while maintaining ACM's high technical quality standards.

Third, better inreach and outreach. We must improve our ways of informing members of the activities and content that interest each of you. And we need to do a better job reaching out to conference attendees, chapter meeting attendees, and online visitors to help them find ACM activities and content of interest, including membership in ACM and its SIGs.

And finally, I feel ACM has a role in guiding our field as we struggle with the relationship between "computer science" and broader "computing." I look forward to these challenges, and to working together on them.

# *candidates for* VICE PRESIDENT (7/1/08-6/30/10)

**ALAIN CHESNAIS**

CTO

SceneCaster.com

Richmond Hill, Ontario

Canada

## BIOGRAPHY

■ ACM and SIGGRAPH Activities
- – Since July 2006 ACM secretary/ treasurer
- – Since July 2005 ACM SIGGRAPH past president
- – 2005 - 2006 SIG representative to Council
- – 2002-2005 ACM SIGGRAPH president
- – 2002-2004 SGB past chair
- – 2000-2002 SGB chair
- – 2001-SIGGRAPH Conference International Resources Chair
- – 1999-2000-SIG Governing Board (SGB) Vice Chair for Operations
- – 1999-ACM EC Nominating Committee
- – 1998-ACM Executive Search Committee
- – 1998-SIGGRAPH Nominating Committee
- – 1997-SIGGRAPH Conference International Chair
- – 1995-1999-ACM SIGGRAPH Vice Chair
- – 1993-1995-Member at Large ACM SIGGRAPH Chapters Committee
- – 1993-1995-ACM Director of Professional Chapters
- – 1993-1995 - ACM Local Activities Board/SIG Board Liaison
- – 1992-Organized the local groups' booth for the SIGGRAPH 92 conference
- – 1991-1995-Chair of the Paris SIGGRAPH chapter

Professional Experience
- – Since June 2007 CTO at SceneCaster.com
- – July 2005-May 2007 Vice President Product Development at Tucows Inc
- – March 2004-July 2005 Director of Engineering at ATI
- – May 2000-November 2003 Vice President, Engineering at TrueSpectra
- – 1996-2000-Director of Engineering, AliasIWavefront
- – 1993-1996-Rendering Manager, Wavefront Technologies
- – 1992-1993-CG Consultant
- – 1987-1992-CTO, Studio Base 2
- – 1983-1987-Research Scientist, Centre Mondial Informatique
- – 1982-1985-Research Assistant, CNRS

Education
- – 1981-IÉcole Normale Supèrieure de l´Enseignement Technique
- – 1980-Diplome d´Etudes Approfondies, Université de Paris XI
- – 1979-Maîtrise Structure Mathématique de l´Informatique, Université de Paris VII
- – 1979-Maîtrise de Mathématiques, Université de Paris VII

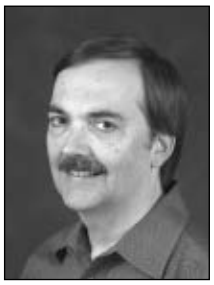
Awards/Achievements
- – Articles in various journals and conferences
- – Systéme particulier selected for film show SIGGRAPH 87.
- – Opening Sequence selected for film show SIGGRAPH 91.
- – Participant in the SIGGRAPH Future Search conference in Snowbird, 1994.
- – Participant in the SIGGRAPH strategy meetings in 2000, 2001, 2005 and 2006.

## STATEMENT

■ I was first introduced to ACM through the annual SIGGRAPH conference over 20 years ago. I immediately joined the local SIGGRAPH chapter in Paris, France and started volunteering for various activities. I haven't stopped since. Throughout my volunteer career I have focused on many aspects of ACM activities, ranging from chapters, to SIGs, to Council. I have also been a very active volunteer in SIGGRAPH, our largest SIG. With over 20 years of management experience in the software industry, I feel well equipped to tackle the issues that the vice president of this organization is expected to manage. In my role as SIG Governing Board chair, I led the task force that proposed and implemented a new allocation model for the SIGs, anticipating the technology downturn that was to come when the bubble burst and the impact we expected it to have on ACM finances. That model was adopted by the SIGs and helped the organization weather the storm despite multi-million dollar losses engendered by our largest SIG. I then focused my efforts on ACM SIGGRAPH, after being elected president, and took the SIG's finances from a 2 ½ million dollar loss when I came in, to generating a modest surplus in my last year.

Today, ACM is a healthy organization that has weathered the tech downturn and currently sees membership rising. The key challenge that I see for ACM in upcoming years have to do with our becoming a truly international organization and attracting younger members into the organization. As a French citizen residing in Canada, I have first hand experience of what it means to be a non American member of our organization. I believe that I can use that experience to help ACM truly become the international organization it should be. I also believe that we need to do much more in terms of expanding our online presence to better cater to the needs of younger researchers and practitioners. Recently at SIGRRAPH, I created a Facebook group for SIGGRAPH members to be able to exchange ideas and communicate with each other using the social networking opportunities that Facebook provides. I believe that there is much more that we can do along these lines to further raise the level of awareness of ACM and significantly grow our membership by becoming more relevant to the needs of students as well as young researchers and practitioners.

# *candidates for* SECRETARY/TREASURER (7/1/08-6/30/10)



**NORMAN P. JOUPPI**

Fellow and Director, Exascale Computing Lab

HP Labs

Palo Alto, CA, USA

## BIOGRAPHY

■ Norm received a Ph.D. in Electrical Engineering from Stanford University in 1984, an M.S.E.E. from Northwestern University in 1980, and a B.S.E.E. from Northwestern University in 1979. After graduation he joined DEC's Western Research Lab, and through acquisition Compaq and then Hewlett Packard. From 1984 through 1996 he was also a Consulting Assistant/Associate Professor in the Department of Electrical Engineering at Stanford University. He currently heads the Exascale Computing Lab at HP Labs.

His current research interests include many aspects of computer system software, hardware, and architecture. He has also led the development of advanced telepresence systems, and has contributed to the architecture and implementation of advanced graphics accelerators. He was the principal architect and lead designer of the DEC MultiTitan and BIPS microprocessors. While at Stanford he was one of the principal architects and lead designer of the Stanford MIPS microprocessor, and developed CAD techniques for VLSI timing verification. He holds more than 35 U.S. patents and has published over 100 technical papers. Norm is a member of ACM SIGARCH, SIGMICRO, SIGGRAPH, SIGMM, and SIGMETRICS.

**ACM Service:** ACM SIG Governing Board (SGB) Representative to ACM Council (2007+), ACM Council Representative to Computing Research Association (CRA) Board (2008+), Editorial board of Communications of the ACM (2008+), SIG Vice Chair for Operations (2006-2007), ACM SGB Member at Large and Conference Advisor (2005-2006), Past Chair of SIGARCH (2007+), Chair of SIGARCH (2003-2007), Vice Chair of SIGARCH (1999-2003), Member of the SIGARCH Board (1993-1999). Program Chair, 1996 International Symposium on Computer Architecture.

**Other service:** Editorial board member of IEEE Computer Architecture Letters 2001+. IEEE TCCA advisory board (2002-2005). Guest editor of IEEE Micro magazine (1992, 1996, 1999).

**Awards:** ACM Fellow. IEEE Fellow. 2005 ISCA Influential Paper award. Compaq 2002 Key Patent award. Two SIGGRAPH/Eurographics Workshop on Graphics Hardware best paper awards.

## STATEMENT

■ Through my service on the ACM Council as SIG Governing Board Representative, I've had the opportunity to learn much about the operation of the ACM. Part of this experience has included finance discussions. I've been a strong proponent of adding value to ACM membership through enhanced services while keeping dues low. Over the last several years this strategy has played an important part in growing ACM membership in an era when membership in related professional societies has been waning. If elected as ACM Secretary/Treasurer, I will continue to look for ways to add value to ACM membership (such as revitalization of the Communications of the ACM and enhancements to the ACM Digital Library) at no marginal cost to our members.

Another important issue going forward is international expansion of our membership while continuing to operate on a sound financial footing. Business and research contributions to computing are becoming increasingly global. Potential members in developing countries can derive significant benefits from relatively low-cost services such as the ACM Digital Library if ACM membership is priced appropriately. Besides mere membership, I would strive to foster volunteer development and encourage service through the world.

During my career I've worked in both industry and academia. I have experience managing the budgets of ACM SIGARCH and my lab at HP, which would serve me well as Secretary/Treasurer. As part of this diverse experience I've developed an appreciation of the range of communities served by the ACM, and I believe I can serve them well as Secretary/Treasurer. Finally, I'm prepared to commit the time required to serve at the best of my ability. And I am always eager to listen to your suggestions!

# *candidates for* SECRETARY/TREASURER (7/1/08-6/30/10)

**BARBARA G. RYDER**

Professor

Department of Computer Science

Rutgers University

New Brunswick, NJ, USA

## BIOGRAPHY

■ A.B. in Applied Math, Brown University (1969); M.S. in Computer Science, Stanford University (1971); Ph.D. in Computer Science, Rutgers University (1982). Associate Member of Professional Staff at AT&T Bell Labs, Murray Hill (1971-1976). Assistant Professor (1982-1988), Associate Professor (1988-1994), Professor (1994-2001), Professor II (2001-), Rutgers University. http://www.cs.rutgers.edu/~ryder/

Fellow of the ACM (1998) for seminal contributions to interprocedural compile-time analyses. Member, Board of Directors, Computer Research Assn (1998-2001). SIGPLAN Distinguished Service Award (2001). Rutgers Graduate Teaching Award (2007). Rutgers Leader in Diversity Award (2006). Professor of the Year Award from Rutgers CS Grad Students (2003).

ACM Council Member-at-Large (2000-2008). Chair, Federated Computing Research Conf (FCRC 2003). SIGPLAN Chair (1995-1997), Vice Chair for Confs (1993-1995), Exec Comm (1989-1999). General Chair of: SIGSOFT Int'l Symp on Software Testing and Analysis (ISSTA, 2008), SIGPLAN Conf on History of Programming Languages III (HOPL-III, 2007), SIGPLAN Conf on Programming Language Design and Implementation (PLDI, 1999, 1994). Program Chair of: HOPL-III (2007), PLDI (1991). Member, Outstanding Contribution to ACM Award Comm and ACM-W Athena Award Comm; Member, SIGSOFT IMPACT Project Steering Comm (2001-). ACM National Lecturer (1985-1988).

Member, Editorial Board of ACM Trans on Programming Languages and Systems (TOPLAS, 2001-2007) and IEEE Trans on Software Engineering (2003-).

Selected panelist: CRA Workshops on Academic Careers for Women in Comp Sci (1993, 1994, 1996, 1999, 2003), SIGSOFT New Software Engineering Faculty Symp (2003, 2005, 2008). Member, Rutgers Advisory Faculty Council on Women in Science, Engineering and Math (2006-). Member: SIGPLAN, SIGSOFT, SIGCSE, ACM, IEEE Computer Society, American Women in Science, EAPLS.

## STATEMENT

■ As ACM Secretary/Treasurer, I will work to ensure good communication among the Exec Comm, Council, SIG leadership, members and staff, and to monitor the financial health of ACM. My extensive experience as a SIG leader and my eight years on ACM Council have prepared me well for these tasks. As General Chair of FCRC 2003, I organized 24 meetings sponsored by 7 SIGs into a coherent, financially sound research conference with 2500 attendees and a $1M budget.

I am determined to maintain ACM as the leading computing society, and our representative on issues of public policy world-wide. There are three key current challenges: providing better services to our practitioner members, expanding ACM into a truly international organization, and supporting the SIGs.

Recent efforts expanded our Local Chapters program, offered new opportunities to mentor younger professionals through MemberNet, enhanced the Digital Library and Portal, and began redesigning CACM to meet member needs better. Such efforts must continue.

Initial efforts at internationalization of ACM established relationships with professionals in India and China that must be strengthened and widened to include areas such as Russia/Eastern Europe, and South Asia. Geographically diverse ACM members should be recruited for ACM and SIG leadership. More ACM meetings outside of North America should be co-sponsored with sister societies.

The SIGs are crucial to ACM: training volunteer leaders, providing valuable research content and tools for the Digital Library, and recruiting students to ACM membership. The SIGs must remain a strong, integral part of ACM. My 10 years of SIG service and 33 years of active ACM membership attest to my commitment to SIG concerns.

I ask for your vote to work for all of these goals.

# *candidates for* MEMBERS AT LARGE (7/1/08-6/30/12)

**ANTHONY JOSEPH**

Chancellor's Associate Professor

Electrical Engineering and

Computer Science Department

University of California, Berkeley

Berkeley, CA, USA

## BIOGRAPHY

■ Dr. Joseph received his S.B. and S.M. degrees in EECS in 1988 and his Ph.D. degree in Computer Science in 1998, all from MIT. He has been on the faculty at the University of California, Berkeley since 1998, holding the Chancellor's Associate Professor Chair since 2007. Starting in June, he will be the director of the Intel Research Berkeley Laboratory.

His research interests include network and computer security, the security of machine learning in decision-making environments, critical cyber-infrastructure protection, mobile computing and networking, wireless telecommunication networks, overlay networks, and computer architectures for distributed systems.

Dr. Joseph is a recipient of the Diane McEntyre Award for Excellence in Teaching, a National Science Foundation CAREER Award, an Okawa Foundation Research Grant in Telecommunications and Information Processing, an IBM Faculty Development Award, an Intel Foundation Graduate Fellowship, an IBM Corporation Graduate Fellowship, and a Shell Corporation National Achievement Corporation Scholarship. He is a lifetime member of ACM, and also a member of USENIX, IEEE, and Sigma Xi.

He has already served ACM in multiple roles: past associate editor for ACM SIGMOBILE Mobile Computing and Communications Review; past technical program co-chair for the ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2003); past editorial board member for ACM/Kluwer Wireless Networks journal; past general chair for the Workshop on Mobile Computing Services and Applications 2006 (co-sponsored by ACM SIGMOBILE); past member of the Defense Science Study Group (2004-2005).

## STATEMENT

■ Our field of computer science faces a changing world. Economic and demographic changes are reshaping the role of CS in today's society. If elected Member at Large, I will focus on:

• Exploring ways to expand distance learning for undergraduate courses and continuing education around the world. I'm interested in increasing the use of distance learning, especially as a distribution mechanism for reaching and educating students and professionals in developing nations. At Berkeley, I put my course lectures online, and I've received positive global feedback. While the cost of computing is ever-declining, the "human" overhead of creating online content has not decreased as much. ACM should help develop the necessary common technologies and services to enable organizations around the globe to distribute their knowledge and engage students.

• Identifying opportunities and mechanisms for using K-12 outreach to increase computer science undergraduate enrollments, especially for women and minorities. Our reduced enrollments are leading to a shortage of CS professionals. I will explore ways to connect K-12 students with CS professionals and academics. By publicizing trends in social networking, online games, and distance learning, ACM could play a vital role in attracting students who are more interested using IT than in the technology itself.

• Working with elected officials to educate them about CS topics, issues, and funding needs. CS has an increasing role in other academic and professional fields around the world, ranging from science (gene sequencing) to art history (image processing). CS is a necessary tool for these fields and government officials need to be educated about the need for balanced applied and basic CS research funding portfolios.

**DANIEL T. LING**

Corporate Vice President

Microsoft Research

Redmond, WA, USA

## BIOGRAPHY

■ Dan Ling received his BS ('74), MS ('75) and PhD ('80) in Electrical Engineering from Stanford University. Dan currently serves as Corporate Vice President at Microsoft. From July 1995 until March 2007 he served as the director of the main Microsoft Research laboratory in Redmond, WA. He has since transitioned to a part-time role on special assignment. Dan joined Microsoft in 1992 to help found a new corporate research laboratory focused on computer science. When he arrived, Microsoft Research consisted of fewer than half a dozen staff in Redmond. Since that time Microsoft Research has grown to over 800 staff in five laboratories around the world. Dan has played a key role in the transformation of Microsoft Research into a leading corporate research lab that has made key contributions to the field of computer science as well as to many Microsoft products and services. In addition to leading the Redmond laboratory he also helped to found the laboratories in Beijing and Bangalore.

Prior to joining Microsoft, Dan was at the IBM Thomas J. Watson research center where he served in multiple roles. His early research work resulted in the co-invention of the dual ported video memory which enabled inexpensive high performance frame buffers. He subsequently managed research projects in computer aided design and computer organization that contributed to the first commercial IBM RISC processor, the PowerPC chip. He also led research efforts in virtual reality, HCI and data visualization.

Dan has served on advisory committees to the School of Engineering at Stanford, UC Berkeley and the University of Washington as well as for the Directorate of Computer and Information Science and Engineering at the National Science Foundation and the Computer Science and Telecommunications Board at the National Academies.

## STATEMENT

■ The ACM has been a vital organization to the field of computer science over the past 60 years. Today the ACM is more important to our field than ever organizing outstanding conferences and publications, pioneering the digital library, serving as a gathering place through the SIGs, and being an important voice for our community. However our field faces a number of challenges where ACM can help. While the technical advances of information technology and its contribution to both society and the world economy have been breathtaking, many students today are turning away from the discipline. The ACM needs to help clearly articulate the excitement of the field and reach out beyond colleges and universities to high schools so that younger students continue to enter the field. Similarly the ACM needs to strengthen its efforts to increase the diversity of people in the profession via initiatives such as ACM-W and CDC.

Much of the excitement today is working across disciplines not only within the field of computer science but understanding how the principles and ideas of computer science can influence other scientific fields. I believe students are particularly excited about addressing major issues facing our society such as climate change, sustainable development and global health. Computer science can play an important role not simply by providing technology but through what Jeannette Wing has called "computational thinking". The ACM can be a strong voice demonstrating these exciting opportunities.

Finally, it is important for ACM to be widely inclusive of our community; to broaden its international perspective through stronger international chapters, to provide additional value to the practitioner, and to be meaningful to new professionals just starting their careers.

# *candidates for* MEMBERS AT LARGE (7/1/08-6/30/12)


**KELLY LYONS**

Associate Professor

Faculty of Information Studies

University of Toronto

Toronto, Ontario

Canada


## BIOGRAPHY

■ Kelly Lyons received a BSc in Computing Science from Queen's University at Kingston, Canada, in 1985. She then joined the IBM Toronto Lab before taking a leave of absence to complete her MSc (1989) and PhD (1994) at Queen's in graph layout algorithms. She held a PhD Fellowship at the IBM Centre for Advanced Studies (CAS) and returned to IBM in 1994 as the Principal Investigator of the Data Management for eCommerce project in CAS.

In 1997, Kelly left CAS to work on complex query performance in the IBM DB2 Performance and Advanced Technology department. She then became a manager with responsibility for projects in self-managing databases, next generation I/O, XML performance, and performance regression testing. She managed the CAS/DB2 projects which involved 12 PhD students and 10 professors. In 2004, Kelly became the Head of CAS with responsibility for over 60 collaborative projects with university researchers. In 2008, she joined the University of Toronto's Faculty of Information Studies as an Associate Professor to teach and conduct research in services science, social computing, data management, and business intelligence.

Kelly has co-authored 16 refereed papers, filed 2 patents, served on program committees for various conferences, given keynote and invited presentations, and co-chaired several workshops. She has been a member of ACM since 2002 and joined the ACM Grace Murray Hopper Award Committee in 2007. Kelly is on the Board of Management of the Ontario Centres of Excellence Centre for Communication and Information Technology, is an IBM Faculty Fellow, and an adjunct professor at Dalhousie and York Universities. She is very active in Women in Technology initiatives and has given several presentations to young people and teachers on this topic.

## STATEMENT

■ I first joined ACM in 2002 as a member from industry and am now enjoying the many resources available through my membership as a university faculty member. Over the years of my membership, I have seen ACM evolve to better service its diverse constituencies, namely, software developers, practitioners, technical managers, researchers, faculty members, and students. ACM's goal of advancing computing as a science and a profession clearly addresses the unique needs and issues of each of these groups as well as their shared desire to ensure a healthy computing field.

As a passionate computer scientist, I have been troubled by the large decline in interest in our field by young people, especially girls. Over the years I have participated in numerous workshops and panels addressing this issue. As an ACM Council Member-at-Large I would like to contribute to ACM's efforts in this important area. Through ACM's dedication to enhancing the image of the computer scientist, and ensuring a robust educational pipeline, its members can play a role in reversing this decline and in helping interest young people in the computing field.

I also feel strongly that a collaborative community environment is an important attractor and retention feature for people in the field of computing. As a Member-at-Large, I would like to find new ways to connect ACM members through social computing technologies and communities of practice across all constituencies. Finally, I want to encourage membership from under-represented groups, especially young women, to join and participate in ACM's activities and communities.


**CHUANG LIN**

Professor, Chairman of Steering Committee

Department of Computer Science and Technology

Tsinghua University, China


## BIOGRAPHY

■ Chuang Lin is a professor at the Department of Computer Science and Technology at Tsinghua University of China since 2000, where he served as department chair from 2003 to 2007. He received a PhD in Computer Science from Tsinghua University of China in 1994, and a MS from Chinese Academy of Sciences in 1981. Upon joining Tsinghua University, he was a professor and chief of Division of Information Technology, Information Economics and Technology Institute, State Information Center of China since 1981.

Lin has devoted his career to Quality of Services and security control of Internet and Computer networks. His research centers on modeling, simulation and performance analysis of computer systems. In these areas, he has published more than 300 papers in research journals and conference proceedings. He is also the author of three books. Lin led the project "Stochastic Modeling and Performance Evaluation of Computer System", which was awarded the first place in the first prize of National Science and Technology Award in 2005 by the Ministry of Education of China.

Lin has been an ACM China Task Force Member since 2006. He is responsible for the ACM China Secretariat and Education Summit in China. In 2007, he was named ACM Council Member. In April 2005, he organized the ACM SIGCOMM Asia Workshop in Beijing China as the general chair and received the Recognition of Service Award from ACM.

Since 2004, Lin has been a Chinese Delegate in TC6 of IFIP and an IEEE Senior Member. He serves as a counselor for the key research project "Network and Information Security" of China National Science of Foundation since 2000. From 2006 to 2010, he is on the Advisory Committee on Elementary Undergraduate Computer Science Education of the Ministry of Education of China.

## STATEMENT

■ I am running for the position of ACM Member at Large to enhance the presence of ACM in China and enable ACM to help the Chinese computing community.

China is growing faster. Its computing society, including both academia and industry is also growing faster. There are more than 100 universities in China which have Ph.D. programs in computer science. Almost all major multi-national IT companies, such as Microsoft, INTEL, IBM, Google, opened branches in China. The Chinese computing community is really large and still growing faster.

ACM is the "first society in computing". Its journals and conferences are the publishing venues of choice for computer scientists, and the Digital Library is a fantastic resource. However, the presence of ACM in China is not satisfactory. There are only less than 1,000 ACM members in China.

I have tried to promote ACM in the past few years. In April 2005, I organized the ACM SIGCOMM Asia Workshop in Beijing China as the general chair and received the Recognition of Service Award from ACM. I have been an ACM China Task Force Member since 2006. I am responsible for the ACM China Secretariat to handle membership issues in China locally. In 2007, I was named ACM Council Member.

If elected Member at Large, here are some issues I would promote on behalf of the ACM community:

—Create greater awareness of ACM in China and help increase membership among the computing community here, including both academia and industry.

—Draw on ACM's resources, talents, and SIGs to help China, especially to setup a team to translate some important English contents, such as Communications of ACM, into Chinese to attract broader readers.

—Organizing education summit to bring ACM's curriculum resources to China.

# *candidates for* MEMBERS AT LARGE (7/1/08-6/30/12)


**MARY LOU SOFFA**

Chair, Department of Computer Science

Owen R. Cheatham Professor of Sciences

University of Virginia

Charlottesville, VA, USA


## BIOGRAPHY

■ Mary Lou Soffa is the Chair of the Department of Computer Science and the Owen R. Cheatham Professor of Sciences at the University of Virginia. She received her Ph.D. in computer science from the University of Pittsburgh in 1977. Prior to her current appointment, she was a Professor at the University of Pittsburgh and also served as the Graduate Dean in Arts and Sciences. Mary Lou's general research interests are in programming languages/compilers and software engineering. Her current focus is on optimizing compilers, program analysis, virtual execution environments, testing and debugging.

In 1999, Mary Lou was selected as an ACM Fellow and received the Presidential Award for Excellence in Science, Mathematics and Engineering Mentoring. She received the SIGPLAN Service Award in 2003 and the Nico Habermann Award from the Computer Research Association (CRA) in 2006.

Mary Lou has been active in ACM for many years. She is currently serving on the Publications Board. She was the ACM SIG Board Council Representative from 2000- 2002 and served as SIGPLAN Chair, Vice-Chair and Treasurer. She also served as member-at-large on the SIGSOFT executive committee.

Mary Lou served on the CRA Board for ten years. She has worked to increase the participation of women and underrepresented minorities for many years. She serves on the CRA-W Committee and was the co-chair from 1999-2002. She co-founded the CRA-W Graduate Cohort Program and the CRA-W Cohort for Associate Professors. She serves on the Leadership Team for NCWIT. She has been a member of a number of Editorial Boards, and conference chair, program chair, and on the program committee for numerous conferences. She recently was the Program Co-Chair for ICSE2006 and is currently the Conference Chair for CGO-08 and ASPLOS-09.

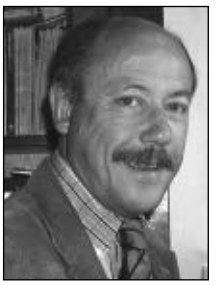
## STATEMENT

■ ACM has a tremendous influence on computing research, development, education and public policy. Much of this influence has come from the SIGS, through their high quality journals and conferences and the digital library, a most important resource. ACM has also developed excellent services through the Professional Development Center and the Career Resource. It is imperative that we continue to support and enhance these important activities.

ACM should make a greater effort in other areas, including providing services to practitioner members and students, and improving the diversity of our community. ACM should expand its leadership role by carefully considering and implementing the types of publications, conferences and services that would be useful for practitioners. Students are a critical component of ACM membership. It is imperative that ACM works with academia, industry and government to demonstrate the excitement and relevance of our field to students. We need to nurture and support our students, including our international students. We can do more to create opportunities for interaction between students from different countries, especially by providing resources enabling them to attend meetings and conferences. ACM should work hard to become a truly international society and enhance the presence of ACM in other countries. One strategy is to determine how ACM can be an important resource for each country's needs.

We need to continue the efforts that ACM already has in place to attract more women and underrepresented minorities and to mentor those who are in the field. Various support strategies and activities are needed at all stages of education, in industry, and in the government.

I will work to continue to support, enhance, and develop these vital activities.




**CARLO GHEZZI**

Chair of Software Engineering

Politecnico di Milano

Milan, Italy


## BIOGRAPHY

■ Carlo Ghezzi is a Professor and Chair of Software Engineering at Politecnico di Milano, Italy. He is the Rector's Delegate for research, past member of the Academic Senate and of the Board of Governors. He has been Department Chair and Head of the PhD Program. He held temporary positions at University of California at Los Angeles, University of North Carolina at Chapel Hill, University of Padova, ESLAI-Buenos Aires, University of California at Santa Barbara, Technical University of Vienna, University of Klagenfurt, University of Lugano.

Ghezzi is an ACM Fellow, an IEEE Fellow, and a member of the Italian Academy of Sciences. He was awarded the ACM SIGSOFT Distinguished Service Award. He has been a member of the ACM Nominating Committee, and is presently a member of the committee for the ACM Software System Award. He has been on the evaluation board of several international research projects and institutions in Europe, Japan, and the USA.

Ghezzi is a regular member of the program committee of important conferences in the software engineering field, such as the ICSE and ESEC/FSE for which he also served as Program and General Chair. In 2006 he was General Co-Chair of the International Conference on Service Oriented Computing.

Ghezzi has been the Editor in Chief of the *ACM Trans. on Software Engineering and Methodology* (2001-2006) and an Associate Editor of *IEEE Trans. on Software Engineering*. He is currently an Associate Editor of *Science of Computer Programming, Service Oriented Computing and Applications*, and *Software Process Improvement and Practice*.

Ghezzi's research has been focusing on software engineering and programming languages. He co-authored over 150 papers and 8 books. He coordinated several national and international (EU funded) research projects.

## STATEMENT

■ ACM is the leading professional society in the field of computing. For over 60 years, it has been serving the scientific community: researchers, educators, engineers, and professional developers. The history of the field can be traced back through the history of the ACM. This history documents how our society has been shaped through continuous innovations. The ACM will continue to assist us in the future advances in hardware, software, networks, and devices, which will make computing ubiquitous and pervasive.

New problems, however, are also emerging and some older problems are still with us. The ACM should play a proactive role to help us facing and solving them.

A lot remains to be done to spread computing science and education in all parts of the world. By its very nature, computing is at the heart of technology that connects the world. It should by no means become the source of deeper divisions and discriminations. The ACM should elaborate policies to support worldwide knowledge sharing and scientific cooperation, breaking all barriers of race, culture, economy, gender, and age.

University education also has problems, particularly in North America and Europe, because it is not attracting enough brilliant young people, and females are a small minority. They often view computing as uninspiring hacking, lacking deep challenging underpinnings. The society at large often has misconceptions, both about our profession and about its roots in science. We must find ways to better communicate. And we must focus our education on long-lasting principles, more than on the mundane and transient surface of technology. The ACM should take the lead of a pride initiative at all levels, finding new ways to communicate with the society, and especially with the new generations.

**MATHAI JOSEPH**

Advisor, Tata Consultancy Services

Pune, India

## BIOGRAPHY

■ Mathai Joseph was Executive Director of the Tata Research Development and Design Centre in Pune, India, until 2007 and is now an Advisor to Tata Consultancy Services. He earlier held a Chair in Computer Science at the University of Warwick from 1985-97. Before that, he was a senior research scientist at the Tata Institute of Fundamental Research in Mumbai, India. He has been a Visiting Professor at Carnegie-Mellon University, Eindhoven University of Technology, University of Warwick and University of York (U.K.).

He was a founder of some long-running international conference series, such as Foundations of Software Technology and Theoretical Computer Science, Formal Techniques in Real-Time and Fault-Tolerant Systems and, more recently, Software Engineering Approaches to Outsourcing and Offshoring. He has set up workshop series (TECS Week and TECS Workshop) to promote computer science research in India, especially among teachers in India's engineering colleges, and established research collaborations with several prominent universities in different countries.

His main research interests have been in real-time and fault-tolerant computing, with emphasis on formal specification and analysis. He has written and edited a number of books and conference and journal papers in these areas. He has worked on software systems for many years and wrote an early book on the design of a multiprocessor operating system.

Mathai Joseph did a Master's degree in Physics at Bombay University and a PhD at the Mathematical Laboratory, University of Cambridge, UK. He has worked in several countries and now lives in Pune, India. He has been investigating early work in computing and computer science in India and the transition over the last few decades of the Indian software industry.

## STATEMENT

■ The remarkable growth of the Indian software industry has been a major stimulus for the recent spread of computing in India. It has led to a massive increase in the number of engineering colleges preparing students for entry into the industry. However, this growth has not been accompanied by any comparable growth in computer science research: there are few institutions where high grade research is done and only a few additions (notably the India research centres of large computing corporations) over this period; there are still very few PhD students in the country.

I believe ACM has a major role to play in creating links between the best in computer science worldwide and education and research in India. ACM publications and conferences have set an unquestioned standard for representing the highest grade of computer science research. ACM awards, like the Turing Award, are universally recognized for honoring the most significant contributions to computing. Yet, as a professional society, ACM is still very under-represented in India and no major ACM meetings are held here.

A lot of my work has been to create forums for the exchange of ideas, new research and people among industrial and educational institutions in India and overseas. I would like to work with ACM to grow these and other initiatives for promoting computer science education and research in India.

I would also like to work with ACM to promote the understanding of the software industry as a global manufacturing industry, with a consequent redefinition of software engineering technology. ACM has already made a valuable contribution through the report of its Task Force on the subject and is well-placed to make the findings a source for policy makers and a part of every computing professional's education.

DOI: 10.1145/1342327.1342334

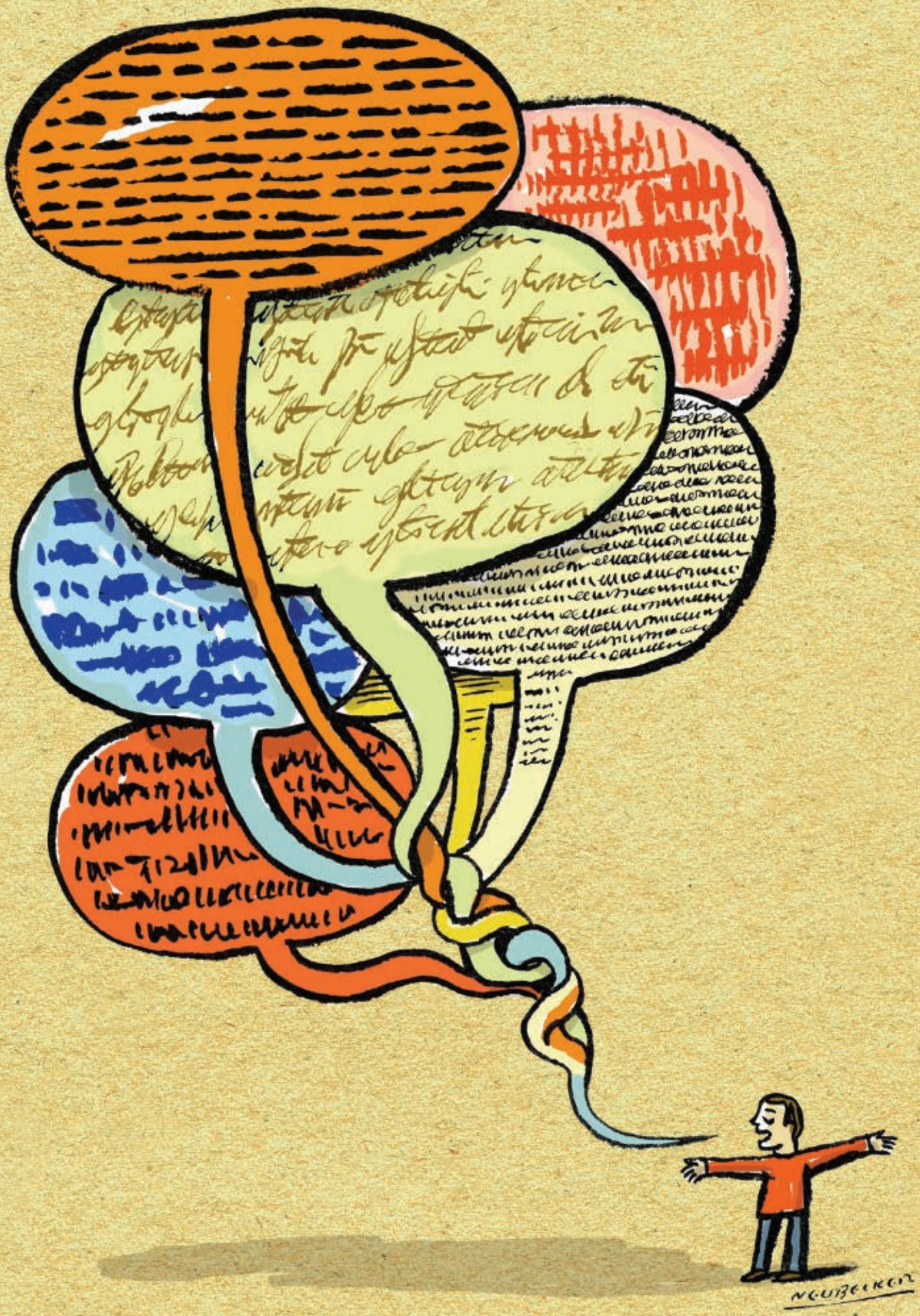
Studies of three prototype Web search portals—in Chinese, Spanish, and Arabic—reveal how to best support non-English Web searching.

# WEB SEARCHING IN A MULTILINGUAL WORLD

*By Wingyan Chung*

Worldwide Internet use has grown tremendously in recent years, most rapidly in non-English-speaking regions. For example, from 2000 to 2007, the online populations in Latin America and the Middle East grew by 577.3% and 920.2%, respectively [9]. At the same time, the number of registered domain names in mainland China (.cn) surged by 137.5% annually [2], fueling the growth of Web pages in Chinese, the second most popular language on the Web. Meanwhile, Arabic Web content was estimated to be doubling every year [1]. Such growth has created demand for better Web searching and browsing in some non-English languages. However, existing Web portals may be unable to meet it because they primarily serve English-speaking users.

*ILLUSTRATION BY ROBERT NEUBECKER*



Arabic, the fifth most popular language in the world, is spoken by more than 284 MILLION PEOPLE IN SOME 22 COUNTRIES, yet the Arabic Web is still in its infancy, constituting less than 1% of total Web content.

While many research findings on Web searching are available, little has been done on the theoretical and empirical aspects of non-English Web searching. Here, I review Web-search engines in a multilingual world and describe a framework that tries to address these issues. Experimental studies of three prototype Web search portals—in Chinese, Spanish, and Arabic—reveal how to best support non-English Web searching.

English has been the dominant language for information seeking on the Web. But this is not the case for many non-English-speaking users who rely on their native languages to search and browse the Web. The process of information seeking consists of various stages of problem identification, definition, resolution, and solution presentation [12]. Two major information-seeking activities are searching and browsing. In searching, users first decompose their goal into smaller problems, then formulate keyword queries, and finally evaluate the results through serial search or systematic sampling. In browsing, users first transform their general information needs into a problem, then explore the Web content and hyperlinks through such browse-support tools as automatic summarization, clustering, visualization, and Web directories, ultimately evaluating the results by scanning through them.

Techniques proposed to support Web searching and browsing include meta-searching and Web-page preview and overview. Because different search engines employ different methods for page collecting, indexing, and ranking, they may include systematic bias in their search results [10]. Meta-searching is a promising method for alleviating this problem [4]. By sending queries to multiple search engines and collating the set of top-ranked results from each engine, meta-searching can greatly reduce bias in search results and improve coverage. In addition, post-retrieval analysis provides added value to results returned by search engines. Text-categorization techniques help filter Web-page content and provide previews of individual Web pages in the form of summaries. Document-categorization techniques help group Web pages, and document visualization techniques help amplify human cognition in browsing Internet search results. Though used in some search engines, including excite.com and vivisimo.com, meta-searching and information previews and overviews are rarely applied in non-English search engines.

Web searching in a multilingual world is characterized by cross-region and cross-country use of a language, producing regional effects in Web-site design and functionality. For example, Spanish is widely used in Europe, North America, and South America. Arabic is the primary language in the Middle East and North Africa. Chinese is the primary language in mainland China, Hong Kong, and Taiwan. The users of the Fast search engine (www.fastsearch.com), mostly European, input queries more frequently than Excite search-engine users, who focus more on e-commerce topics [11]. These results suggest regional differences on the Web.

## SEARCH ENGINES

Several major search engines provide search services to non-English-speaking users. Having more than 160 local domains, Google allows users to restrict search results to pages in 117 languages, providing translation services between English and eight European languages (Dutch, French, German, Greek, Italian, Portuguese, Russian, and Spanish), three Oriental languages (Chinese, simplified and traditional, Korean, and Japanese), and Arabic. AltaVis-

ta's Babel Fish (babel.altavista.com) provides more pairwise translation services between languages (except Arabic). Similar translation services are also provided by Yahoo!, which has regional sites in 24 countries supporting Web search in 37 languages used by 411 million unique users each month. Yahoo!'s diversified services, including online shopping, auctions, email, news, blogs, partnerships with content providers, and instant messaging, enable it to fit comfortably into most aspects of users' lives. Meanwhile, MSN Search has 42 regional sites located in different countries. Its U.S. site supports a local search service for searching information in the user's geographic area. Like Yahoo!, MSN also provides such services as email, instant messaging, news, and entertainment information. Its connection with Microsoft Windows and Internet Explorer has helped it earn an important share in the search market (monitored by SearchEngineWatch.com, a Web site that provides lists and reviews of major and specialized search engines).

| Region | Greater China Regions | | | | | | U.S, Europe, and Latin America | | | | | | | | Middle East | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Main language | Simplified Chinese | | Traditional Chinese | | | | Spanish | | | | | | | | Arabic | | | |
| Search engine or portal (headquarters) / Content and Functionality Features | Baidu.com (mainland China) | Sina.com.cn (mainland China) | Yahoo! HK (hk.yahoo.com, Hong Kong) | Timway.com (Hong Kong) | Yam.com (Taiwan) | Openfind.com.tw (Taiwan) | Terra.com (Spain) | Orange.es (France) | Auyantepui.com (Venezuela) | Conexcol.com (Colombia) | Bacan.es (Ecuador) | Yahoo! Telemundo (telemundo.yahoo.com, Spain) | BIWE.com (Spain) | Quepasa.com (Mexico & U.S.) | Ajeeb.com (Kuwait) | Albawaba.com (Jordan) | Weyak.ae (U.A.E.) | Ayna.com (U.S.) |
| Membership services | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Newsgroup/ Weblog search | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Web directory | | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Search for Web sites | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Search stock prices | | ✓ | ✓ | | ✓ | | | ✓ | | | | ✓ | | | | | | |
| Online translation | | | ✓ | | | | | | | | | ✓ | | | ✓ | | | |
| Search for news | ✓ | ✓ | ✓ | | ✓ | ✓ | | ✓ | | | | ✓ | ✓ | | ✓ | ✓ | | ✓ |
| Multimedia search (image, music, video, software) | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Size of collection | Very good | Good | Very good | Fair | Good | Very good | Very good | Very good | Fair | Fair | Fair | Very good | Good | Good | Very good | Good | Very good | Very good |
| User interface | Good | Fair | Very good | Fair | Very good | Good | Very good | Very good | Good | Good | Fair | Very good | Good | Good | Very good | Very good | Good | Good |

**Table 1. Search engines and portals in Chinese, Spanish, and Arabic.**

While an exhaustive review of search engines in all languages is beyond my scope here, I have reviewed major Web search engines in three emerging languages: Chinese, Spanish, and Arabic. Table 1 lists major search engines and portals in these languages, highlighting important content and functionality features. Chinese is the primary language used by people in mainland China, Taiwan, and Hong Kong. Language encoding, vocabularies, economies, and societies of these regions differ significantly. In mainland China, Baidu.com is a major search engine serving many large enterprises, including Dell (China), Lenovo, and Yahoo! China. It has collected more than a billion Chinese Web pages from mainland China, Hong Kong, Taiwan, and other regions, a collection that grows by several hundred thousand Web pages per day. Another major Web portal in China, Sina.com.cn, provides comprehensive services, including Web search, email, news, business directory, entertainment, and weather forecasts. Leveraging on its rich content and large user base, Sina has its own search engine, iAsk.com, that uses both Web content and usage information to rank Web pages. Other search engines in mainland China include Sogou.com and Zhongsou.com.

The two major search portals in Taiwan are Openfind and Yam. Openfind.com.tw, established in 1998, suggests relevant terms to refine users' search queries, allowing them to find other related items from each search result. Yam.com, established in 1995, provides comprehensive online services involving search in various media in Taiwan, including Web sites and pages, news, forum messages, and local activities. Since 2000, Yam.com has partnered with Google to provide search services.

Due to Hong Kong's bilingual culture, people there rely on both English and Chinese when searching the Web. Yahoo! Hong Kong (hk.yahoo.com) returns results in a variety of categories, including Web sites, pages, and news. Established in 1997, Timway.com searches more than 30,000 Hong Kong Web sites categorized into more than 3,000 groups that attract 2.6 million visits per month.

Spanish is the second most popular language in the U.S., as well as the primary language in Spain and some 22 Latin American countries where regional search engines provide search and browse services. With 19 regional sites, Terra.com offers its services to more than 3.1 million Internet users in the U.S., Spain, and Latin America. A Gallup poll in 2002 described Terra as the most popular search engine in Spain; Orange.es (formerly Wanadoo), a subsidiary of

I recommend that system developers and IT managers INCORPORATE BROWSE SUPPORT AND ANALYSIS TOOLS into their online search systems and portals to augment traditional textual list displays.

France Telecom, was second. Yahoo! Telemundo (Spain, telemundo.yahoo.com), the Spanish version of Yahoo! serving the U.S. and Latin America, provides a Web directory compiled by human editors categorizing millions of listed sites. Yahoo! Telemundo supplements its results with those from Inktomi and Google. Established in 1995 as one of the first search engines to search Spanish information on the Web, BIWE.com provides a variety of services, including a Web directory, email, entertainment, and market information for Spanish-speaking users. Meanwhile, Quepasa.com, with headquarters in the U.S., is a bilingual Web portal (Spanish/English) serving Spanish-speaking populations in the U.S. and Latin America.

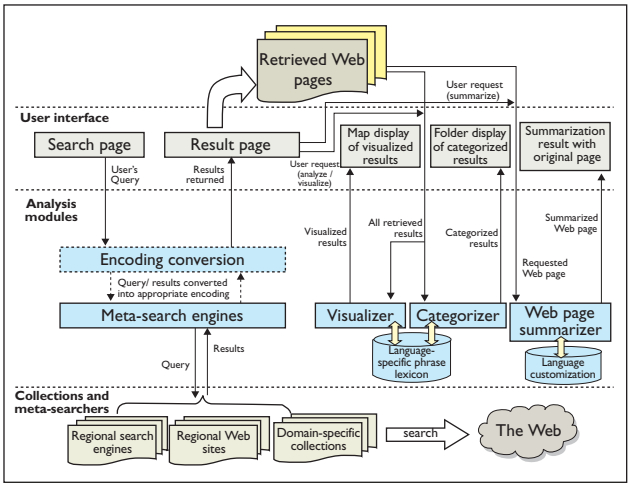



**Figure 1. Framework for Web searching in a multilingual world.**

Arabic, the fifth most popular language in the world, is spoken by more than 284 million people in some 22 countries, yet the Arabic Web is still in its infancy, constituting less than 1% of total Web content. Four major search engines offer Arab speakers comprehensive services and extensive content coverage. Ajeeb.com, a bilingual Web portal (English/Arabic) launched in 2000 by Sakhr Software Company, includes a multilingual dictionary (Arabic/English/French/Turkish/German) and a Web directory, "Dalil Ajeeb," which Ajeeb claims to be the world's largest online Arabic directory. Albawaba.com, another Arabic search portal providing comprehensive services, supports searching for both Arabic and English pages; results are classified according to language and relevance. It also meta-searches other search engines—Google, Yahoo!, Excite, Alltheweb, and Dogpile—and provides a comprehensive directory related to 22 Arab countries. Launched in 2000, United Arab Emirates-based Weyak (www.weyak.ae/) offers a range of online services covering more than 1.25 million Arabic Web pages. Based in the U.S. (in New Hampshire), Ayna.com provides an Arabic Web directory, an Arabic search engine, and other services, including a trilingual (Arabic/English/French) email system, chat, greeting cards, personal homepage hosting, and commercial classified ads. Claiming more than 700,000 registered users, Ayna provides access to more than 25 million pages per month. Alexa Research ranks Ayna among the three leading Web sites in the Arab world.

## FRAMEWORK FOR WEB SEARCHING

My review found that existing search engines in Chinese, Spanish, and Arabic typically present results in the form of long lists of textual items. While such presentation is convenient for viewing, it may limit users' ability to understand and analyze the results. The collections searched by the search engines are often region-specific and lack a comprehensive understanding of the environment in which they operate. Major English-language search engines, including Google, support searching through non-English resources but fall short of covering domain- and region-specific information. There is a need to better support Web search in some emerging non-English languages. Here, I

describe a framework that addresses some of the needs for Web searching in a multilingual world. As outlined in Figure 1, the framework consists of domain collections, meta-search, statistical language processing, and Web-page summarization, categorization, and visualization.

Reflecting regional and language differences, a careful domain analysis must be conducted by any prospective search-engine developer before a Web portal is built in any particular language. To ensure comprehensive coverage, the analysis should review existing Web portals and technologies, including the characteristics of the language, and select an area or theme for which significant Web resources in the language have been developed. The review should cover regional search engines, government and business Web sites, and news Web sites to select the relevant Web content needed to build a domain-specific collection or for meta-search. Important keywords and URLs relevant to the chosen domain are gathered as seed queries or hyperlinks to build the collection.

Managing a large user base and growing Web content, many non-English Web search engines and portals are challenged to organize their content properly to support convenient browsing and searching. For example, Sina.com.cn includes more than 700 hyperlinks on its home page, each annotated with long textual descriptions in a small font, making browsing difficult, especially for inexperienced Web users. Pre- and post-retrieval analysis is thus needed to alleviate information overload.

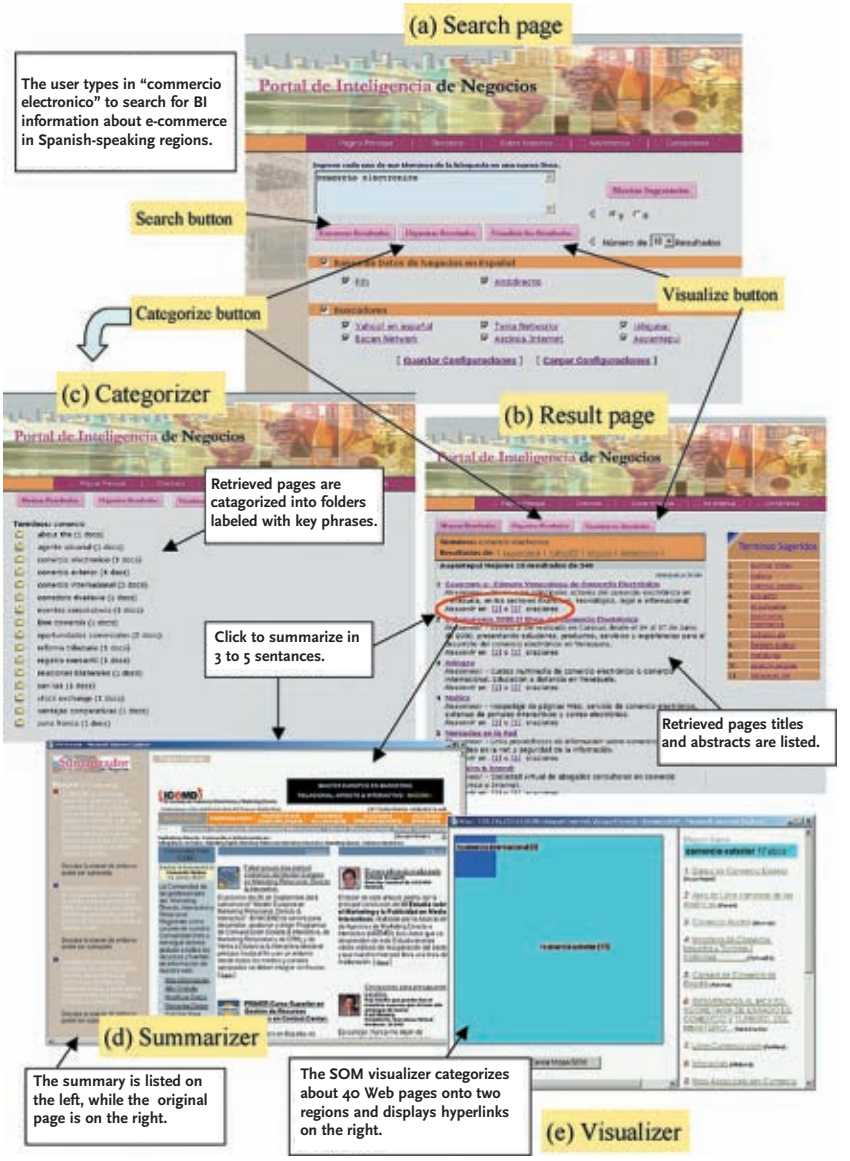



**Figure 2. Screenshots from SBizPort.**

Modules supporting such analysis include encoding conversion, summarization, categorization, and visualization. Encoding conversion is necessary when a language is used by people in multiple regions and countries using different versions of the same language. For example, the traditional and simplified versions of Chinese differ enormously in written formats, leading to two different input formats in information-retrieval systems; hence, they require encoding conversion for searching across the two language versions. Web-page summarization uses linguistic and heuristic techniques to extract key sentences from the page to represent a summary of the article [8].

Categorization helps organize search results in different groups that are understood more easily. To assist the categorization process, lexicons built by a statistics-based mutual-information approach can provide meaningful phrases in different languages. A neural-network approach called "Kohonen self-organizing map" can be used to categorize and visualize Web pages, helping users navigate on a 2D jigsaw map to identify the set of similar pages or find relevant pages.

Based on the framework, three prototype search portals—in Chinese, Spanish, and Arabic—were developed [3, 6]. The Chinese Web portal (CBizPort) helps users search and browse for business intelligence (BI) in mainland China, Hong Kong, and Taiwan.

Here, BI refers to the product of acquisition, interpretation, collation, assessment, and exploitation of information in the business domain [4]. CBizPort includes two versions of its user interface—one for simplified Chinese, one for traditional Chinese—each with the same look and feel. Relying on a conversion dictionary with 6,737 Chinese characters in each of the two encodings (Big5 and GB2312), the encoding converter converted all Chinese characters into the encoding of the interface version. The eight information sources used in the portal's meta-searching are major Chinese search engines and business-related portals from the three regions. Relying on two Chinese-phrase lexicons to extract phrases, the portal's categorizer organizes retrieved Web pages into various folders labeled with the key phrases in the page summaries and titles.

The Spanish Web portal (SBizPort) supports searching and browsing of business information from 22 Spanish-speaking regions. In addition to keyword searching, summarization, and categorization, like those in CBizPort, SBizPort provides a comprehensive collection of business Web pages for searching and supports visualization of retrieved pages (see Figure 2). Users visualize Web pages by clicking on a region to see a list of pages on the right and open pages by clicking the link-embedded titles.

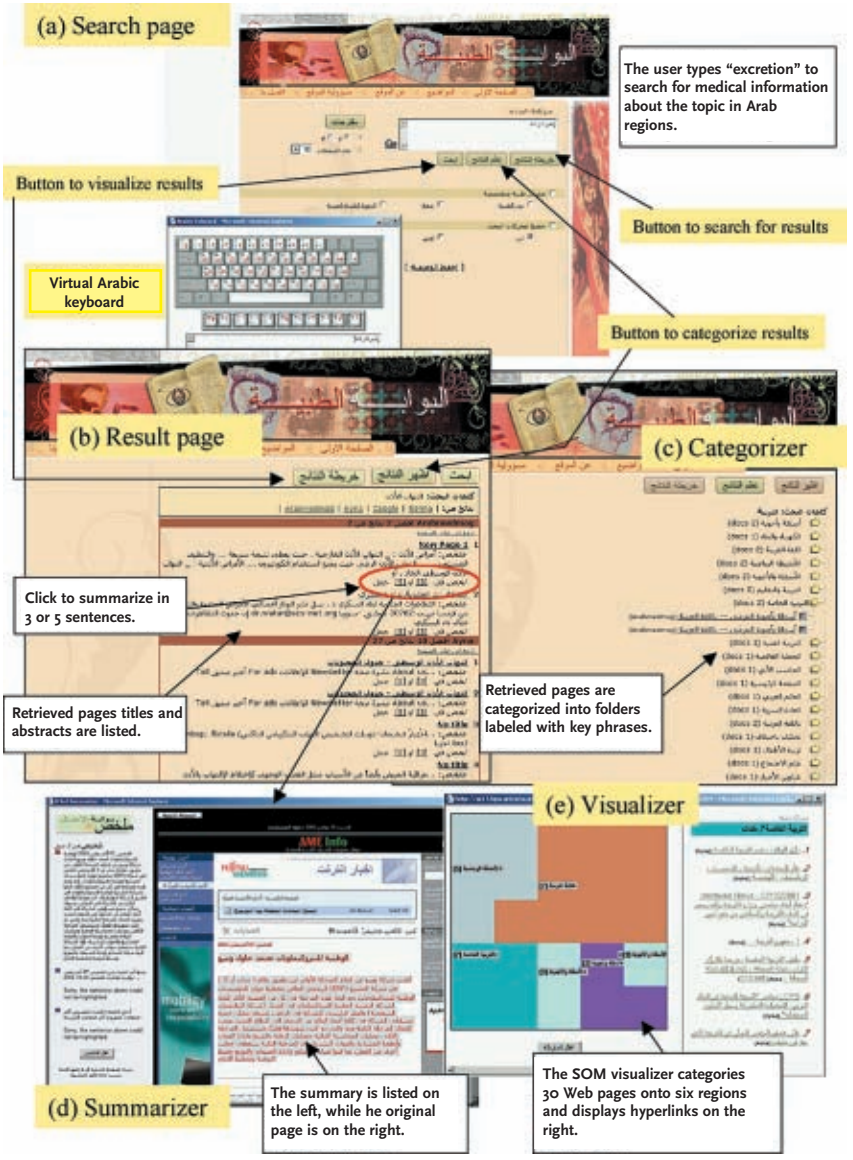



**Figure 3. Screenshots from AMedPort.**

The Arabic Web portal, AMedPort (see Figure 3) focuses on the medical domain in some 22 Arab regions and supports all search and browse functions available in SBizPort. AMedPort includes a customized user interface with a right-to-left text display and a virtual keyboard to facilitate Arabic input.

## EXPERIMENTAL FINDINGS

Sixty native speakers participated in the experiments (detailed in [3, 6]) of the three Web portals to evaluate the framework's usability in supporting Web searching in a multilingual world (see Table 2). In each experiment (about an hour), a Web portal was compared with a benchmark search engine in each of the three languages. Each subject was introduced to the portal and the benchmark search engine and was randomly assigned different task scenarios (one scenario per system). Each scenario contained three or four search and browse tasks based on standards developed through the National Institute of Standards and Technology Text Retrieval Conference (trec.nist.gov). All questionnaires used in the experiment were administered in the subjects' native language. Subjects spent an average of three minutes to finish a search task and eight minutes to finish a browse task. The order in which the systems were used was randomly assigned to avoid bias due to the sequence of use.

Several information-retrieval measures—precision, recall, and F value—revealed the search and browse effectiveness of the system by computing the ratios between the number of relevant results found by a

Users must be cautioned that THE TOOLS ARE STILL PRONE TO ERROR due largely to ambiguities in natural-language processing and high computational costs.

subject and the number of all results found by the same subject or by an expert. Domain experts provided answers for judging the subjects' browse-task performance. After using a system, a subject filled out a questionnaire with comments and satisfaction ratings on a seven-point Likert scale.

In the CBizPort experiment with 30 Chinese subjects from mainland China, Hong Kong, and Taiwan and with three Chinese business academics and practitioners serving as experts, the results show that the effectiveness of existing Chinese search engines can be improved significantly by adding CBizPort. However, no significant difference was found in the effectiveness of the summarizer and categorizer in the Chinese portal and in their user-satisfaction ratings. Despite this, 11 subjects commented that the summarizer and categorizer facilitated their understanding and searching. These results indicate that CBizPort could augment these search engines in searching and browsing Chinese business information but also that the summarizer and categorizer need further improvement in precision and browse support.

In the SBizPort experiment with 19 Spanish subjects from six countries—Colombia, Mexico, Panama, Peru, Puerto Rico, and the U.S.—and a veteran Spanish business consultant serving as the expert, the SOM visualizer achieved significantly better browse effectiveness than BIWE, the benchmark search engine, showing that the tool helps alleviate information overload and supports browsing effectively. The use of a domain-specific collection achieved higher mean accuracy and search efficiency than not using it, though the differences were not significant. Subjects rated SBizPort significantly better than BIWE, citing the precision and relevance of returned results. These results further indicate that the information-visualization tool can be an alternative to presenting search results as text in a list.

In the AMedPort experiment with 11 Arab subjects from five countries—Iraq, Jordan, Lebanon, Mauritania, and Morocco—and an Arab microbiology researcher serving as the expert, the AMedPort achieved a significantly higher mean accuracy, efficiency, and satisfaction rating than (and comparable browse effectiveness to) Ayna, the benchmark search engine. Nine subjects said AMedPort was useful and provided more topics and information than the benchmark. The portal provided high-quality information from many sources but needs improvement in both its summarizer and categorizer.

There was a probability of 0.05 (or less) that the confirmed results were actually not statistically valid in the experiments in which the best search engines in the respective languages were used as benchmarks.

| Portal | CBizPort | SBizPort | AMedPort |
|---|---|---|---|
| Research questions | • How does the framework support Web searching in a multilingual world?<br>• How can Web searching and browsing be made more effective by using the portals developed by the framework?<br>• What are the lessons learned and implications for non-English Web searching? | | |
| Languages | Traditional Chinese is used in Hong Kong and Taiwan, while simplified Chinese is used in mainland China. They have different encoding formats. Word segmentation is a problem. | Spanish is used in many regions and countries in North and South Americas and Europe; Catalan (another version of Spanish) is widely used as well. | Arabic is used in more than 22 countries in the Middle East and North Africa. It is written from right to left. Web content in Arabic is generally not rich. |
| Regions | Mainland China, Hong Kong, Taiwan | Mexico, Honduras, Costa Rica, Panama, Colombia, Venezuela, Ecuador, Peru, Chile, Argentina, Spain, Bolivia, Paraguay, Uruguay, Guatemala, Nicaragua, U.S. | Lebanon, Saudi Arabia, Bahrain, Canada, Tunisia, Kuwait, Egypt, Switzerland, United Arab Emirates, Russia, U.K., U.S. |
| Benchmarks | Sina.com.cn, Openfind.com.tw, HK.yahoo.com | BIWE.com | Ayna.com |
| Result highlights | CBizPort's meta-searching, summarization, and categorization are helpful for searching and browsing Chinese Web pages in the three regions. CBizPort needs improvement on search precision and information quality. | SBizPort achieved a better browse performance and satisfaction rating than the benchmark but needs to be improved on its domain-specific collection. | AMedPort achieved better search accuracy than comparable browse performance in Ayna. AMedPort obtains better user ratings in user satisfaction. |

**Table 2. Experimental studies on developing Web search portals in various languages.**

## LESSONS LEARNED

The results from these experiments demonstrate that

the framework supports Web searching in a multilingual world. Post-retrieval analysis techniques (such as summarization and visualization) were found to alleviate information overload but also that the extent of such improvement varies across domains. Summarization and categorization did not achieve significant improvement in the CBizPort study. In the SBizPort and AMedPort studies, information visualization achieved significant performance improvement in Web-search results. The ability to visualize a large number of search results was essential for good performance in all three portals.

I recommend that system developers and IT managers incorporate browse support and analysis tools into their online search systems and portals to augment traditional textual list displays. Such tools can be used to summarize Web-page textual descriptions [6], support query formulation [7], visualize emerging events related to their environment and organizations [5], and categorize search results into hierarchies or maps [4]. However, users must be cautioned that the tools are still prone to error due largely to ambiguities in natural-language processing and high computational costs that may not be economical for small Web sites.

Factors to be considered when adopting the tools include the extent to which the Web-page collection provides sufficient statistical information for machine learning, adequate hardware and software to support intensive computation, availability of a work force to improve the Web-site interface and accommodate new presentation choices, characteristics of the language used, and user IT literacy.

Across a variety of languages and domains, I found significant differences in the development of Web-search portals, technologies, and language use. For instance, the growth of Internet use in mainland China (but relative lack of comprehensive Web search and browse support) strongly suggests the need for future improvements. While Web-search technologies in Taiwan are more mature, there is likely room for new technologies developed specifically for processing Chinese. The strong growth of the Chinese- and Spanish-speaking online populations will likely persist in the coming years, further emphasizing the need for better, more integrated Web-search portals that deliver results in a variety of formats and provide richer information for the regions and the communities that use the languages. The increasing amount of Arabic Web content and online population, along with economic and political developments in Arab regions, will continue to fuel the growth of many Arabic Web sites that remain mostly underdeveloped today. The research I've reported here will likely contribute to a better understanding of related developmental and experimental issues.

My ongoing work includes developing scalable techniques to collect, analyze, and visualize Web information in different languages, studying user needs in non-English Web search, and exploring the effect of new techniques in information exploration and analysis. This effort will contribute to Web searching and browsing in a multilingual world. C

## References


1. Abbi, R. *The Current Status of the Internet in the Arab World.* UNESCO Observatory on the Information Society (2002); www.unesco.org/cgibin/webworld/portal_observatory/cgi/jump.cgi?ID=2329.
2. China Internet Network Information Center. *The 20th Statistical Survey Report on the Internet Development in China;* Beijing, China, 2007; www.cnnic.net.cn/uploadfiles/pdf/2007/7/18/113918.pdf.
3. Chung, W., Bonillas, A., Lai, G., Xi, W., and Chen, H. Supporting non-English Web searching: An experiment on the Spanish business and the Arabic medical intelligence portals. *Decision Support Systems 42,* 3 (Dec. 2006), 1697–1714.
4. Chung, W., Chen, H., and Nunamaker, J. A visual framework for knowledge discovery on the Web. *Journal of Management Information Systems 21,* 4 (Spring 2005), 57–84.
5. Chung, W., Chen, H., Chaboya, L., O'Toole, C., and Atabakhsh, H. Evaluating event visualization: A usability study of the COPLINK spatio-temporal visualizer. *International Journal of Human-Computer Interaction 62,* 1 (Jan. 2005), 127–157.
6. Chung, W., Zhang, Y., Huang, Z., Wang, G., Ong, T.-H., and Chen, H. Internet searching and browsing in a multilingual world: An experiment on the Chinese Business Intelligence Portal. *Journal of the American Society for Information Science and Technology 55,* 9 (July 2004), 818–831.
7. Leroy, G., Xu, J., Chung, W., Eggers, S., and Chen, H. An end-user evaluation of query formulation and results review tools in three meta-search engines. *International Journal of Medical Informatics 7,* 11–12 (Nov.–Dec. 2007), 780–789.
8. McDonald, D. and Chen, H. Summary in context: Searching versus browsing. *ACM Transactions on Information Systems 24,* 1 (Jan. 2006), 111–141.
9. Miniwatts International. *Internet Usage Statistics: The Internet Big Picture* (updated Nov. 30, 2007); www.internetworldstats.com/stats.htm.
10. Mowshowitz, A. and Kawaguchi, A. Bias on the Web. *Commun. ACM 45,* 9 (Sept. 2002), 56–60.
11. Spink, A., Ozmutlu, S., Ozmutlu, H., and Jansen, B. U.S. versus European Web searching trends. *SIGIR Forum 36,* 2 (Fall 2002).
12. Wilson, T. Models of information behavior research. *Journal of Documentation 55,* 3 (June 1999), 249–270.



**Wingyan Chung** (wchung@scu.edu) is an assistant professor in the Department of Operations and Management Information Systems of the Leavey School of Business at Santa Clara University, Santa Clara, CA.





DOI: 10.1145/1342327.1342335

*Intuition is a powerful tool that helps us navigate through life, but it can get in the way of more formal processes.*

# HOW INTUITIVE IS OBJECT-ORIENTED DESIGN?

By Irit Hadar and Uri Leron

The object-oriented programming paradigm was created partly to deal with the ever-increasing complexity of software systems. The idea was to exploit the human mind's natural capabilities for thinking about the world in terms of objects and classes, thus recruiting our intuitive powers for building formal software systems. Indeed, it has commonly been assumed that the intuitive and formal systems of objects and classes are similar and that fluency in the former helps one deal efficiently with the latter. However, recent studies show that object-oriented programming is quite difficult to learn and practice [1, 3, 7]. In this article, we document several such difficulties in the context of experts participating in workshops on object-oriented design (OOD). We use recent research from cognitive psychology to trace the sources of these difficulties to a clash between the intuitive and analytical modes of thinking.

Recent research in cognitive psychology shows that people consistently make mistakes on simple everyday tasks, even when the subjects are knowledgeable, intelligent people, who undoubtedly possess the necessary knowledge and skills to perform correctly on those tasks. The source of these mistakes is often shown to be the insuppressible influence of intuitive thinking. This research, the *heuristics and biases program*, has been carried out by Kahneman and Tversky and others during the last 30 years, and has led to Kahneman's receiving the 2002 Nobel Prize in economics.[1] In his Nobel Prize lecture, Kahneman opened with the following story:

A baseball bat and ball cost together one dollar and 10 cents. The bat costs one dollar more than the ball. How much does the ball cost?

Almost everyone reports an initial tendency to answer "10 cents" because the sum $1.10 separates naturally into $1 and 10 cents, and 10 cents is about the right magnitude. Indeed, many intelligent people yield to this immediate impulse: 50% (47/93) of Princeton students and 56% (164/293) of students at the University of Michigan gave the wrong answer [2, 4].

What are our mind's mechanisms that may account for these empirical findings? One current influential model in cognitive psychology is *Dual-Process Theory* [4, 10, 11]. According to this theory, our cognition and behavior operate in parallel in two quite different modes, called System 1 (S1) and System 2 (S2), roughly corresponding to our common sense notions of intuitive and analytical thinking.

These modes operate in different ways, are activated by different parts of the brain, and have different evolutionary origins (S2 being evolutionarily more recent and, in fact, largely reflecting *cultural* evolution). S1 processes are characterized as being fast, automatic, effortless, unconscious, and inflexible (difficult to change or overcome). In contrast, S2 processes are slow, conscious, effortful and relatively flexible. In addition, S2 serves as monitor and critic of the fast automatic responses of S1, with the "authority" to override them when necessary. In many situations, S1 and S2 work in concert, but there are situations (such as the ones concocted in the heuristics and biases research) in which S1 produces quick automatic non-normative responses, while S2 may or may not intervene in its role as monitor and critic.

A brief analysis of the bat-and-ball data can demonstrate the usefulness of dual-process theory for the interpretation of empirical data. According to this theory, we may think of this phenomenon as a "cognitive illusion" analogous to the famous optical illusions from cognitive psychology. The surface features of the problem cause S1 to jump immediately with the answer of 10 cents, since the numbers one dollar and 10 cents are salient, and since the orders of magnitude are roughly appropriate. The roughly 50% of students who answer 10 cents simply accept S1's response uncritically. For the rest, S1 also jumps immediately with this answer, but in the next stage, S2 interferes critically and makes the necessary adjustments to give the correct answer (five cents).

While people in everyday situations prefer responses over careful systemic reasoning, students solving mathematical problems would be expected to consciously train their methodical thinking to check, and override if necessary, their immediate intuitive responses. From these findings we may understand the strong influence intuition.

Recently, a similar phenomenon has been found in advanced mathematical thinking, with college students learning abstract algebra [6]. While it seems natural that people in everyday situations prefer (however unconsciously) quick approximate responses that come easily to mind over careful systematic rule-bound reasoning, students solving mathematical problems during a university course would be expected to consciously train their methodological thinking to check, and override if necessary, their immediate intuitive responses. From these findings we may understand the strong influence intuition, especially its tendency to be influenced by surface clues, has on our thinking. In this article we demonstrate that a similar phenomenon—and a similar explanation—may also hold for OOD tasks carried out by experts in industry.

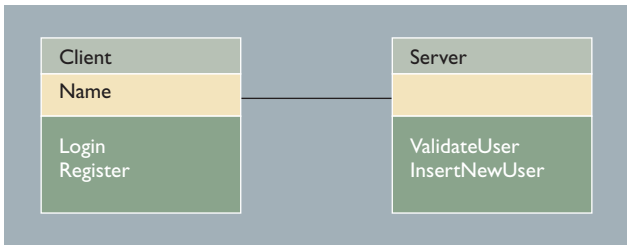
A note on terminology: We follow Kahneman and

[1] Tversky unfortunately died several years earlier.

other cognitive psychologists in using "intuition" in its folk meaning of everyday thinking. This meaning is elaborated in the description of System 1, and is mainly used in contradistinction to analytical thinking or to reasoning. The title of this article should thus be understood as an inquiry into the nature of the gap between the everyday "natural" meaning of objects and categories vs. their formal meaning in OOD. It should further be noted that intuition may have different meanings in different contexts. For example, our use of the term is quite different from the way a mathematician might use it when he or she says: "I had the intuitive idea of the proof long before I was able to complete the formal proof."

## INTUITIVE THINKING IN OOD

OOD is a complex domain, requiring formal training and effortful thinking, which is just the kind of process System 2 would be expected to appropriate. However, our research indicates that here too the automatic, quick, and effortless operation of System 1 may hijack software developers' attention and lead them to decisions that are not adequate and may even clash with their own knowledge.




A design of authorization system.

We discuss several examples of this phenomenon exhibited by experienced software developers in industry while practicing design activities, and explain them in light of the dual-process theory. We invoke this theory in the domain of OOD in an attempt to understand the relatively elementary mistakes we observed in the responses of intelligent capable professionals, even in cases when they have the necessary knowledge to avoid such mistakes.

Our observations took place within advanced UML workshops [8] conducted in the industry. During these workshops the participants were asked on several occasions to analyze simple design tasks. The participants worked on these tasks either individually or in small groups and their solutions were subsequently discussed within the whole group. Our data includes the written solutions of the participants in the workshops, documentations of their group discussions as observed and documented by the researchers, and transcripts of class discussions. The research population included 41 software developers with experience of 2–12 years in OO development. Because our objective was to describe a complex situation in its natural settings and its full complexity, we have used the qualitative research paradigm [12], which focuses on case studies for obtaining specific insights rather than on large populations, simplified experiments, and statistical methods for discovering universal laws. (This is analogous to the methods used by anthropologists studying unfamiliar cultures.) During the research we documented, videotaped, and analyzed many relevant incidents and processes. The data analysis included coding the data obtained, and characterizing and classifying it to emerging categories. The full research findings and evidence will be described elsewhere. Here, we provide a selection of examples to demonstrate our findings.

*Confusing the direction of inheritance.* A design task concerning a hotel reservation system was presented for a discussion to a group of experienced engineers participating in a UML workshop. The instructor suggested using three classes (email, fax, and phone), to represent the three corresponding modes of entering reservations. The possibility then arose of using inheritance relations between concrete classes to exploit shared functionality and features, such as checking the availability of a room.

For example, the class fax could inherit from the class email, since a fax object requires more handling (such as scanning and digitizing), hence has more functionality, than an email object.

*Instructor:* Under the restriction that for now we only use these three classes, can any of these classes inherit from another class? Can we use the fact that they have many things in common?
[The participants hesitate]…
*Instructor:* For example, fax is like email, only with a few more tests.
*Dan:* Email inherits from fax, because email is the same as fax, only with fewer tests.
*Instructor:* So, email has less functionality than…
*Dan* [hastily]: Oh, right, it should be the other way around.

In view of this and similar observations, we presented a group of 10 software developers with a similar question in order to check this phenomenon more directly. The answers were divided 5:5 between the two possible directions of inheritance. Significantly, as in the bat-and-ball and in Dan's case, the participants who chose the wrong direction required only a small nudge (with no informational or explanatory content) to quickly change their mind.

*Analysis:* All the participants in the research have

several years of experience in OO software development. Why do intelligent and experienced professionals have difficulties with such an elementary issue? We propose that the same mechanism used by Kahneman to explain the bat-and-ball phenomenon is also in operation here. Specifically, S1 with its quick and effortless operation "hijacks" the thinking process and produces a response that seems roughly appropriate, while the slow and effortful S2 remains dormant. This analysis gets additional support from the observation that the small cue offered by the instructor didn't teach the participant anything new, only served to wake up S2; the necessary knowledge was there all along, but the dual system analysis is needed to explain why it was not mobilized.

**Under the demands of abstraction, formalization, and executability, the formal OO paradigm has come to sometimes clash with the very institutions that produced it.**

Why would S1 and S2 clash about the meaning of inheritance? In people's everyday intuition (S1), inheritance is about transferring "stuff" (such as property or money), and the direction is usually from the person who has more to the one who has less. For example, in an informal poll we asked students, in the context of OOD, what is the relation between a doctor and a paramedic in an ambulance? A typical reaction was, "paramedic inherits from the doctor because the doctor has more qualifications." Similarly, we predict that most people would say that a student "inherits" from the professor (because the professor has more knowledge) and not vice versa. But in the OOD formalism (S2), the reverse is true: the class with more functionally inherits from the one with less.

*Difficulties in identifying objects.* One of the first tasks in OOD is "carving a given scenario at its joints" in terms of objects and classes. In one of the workshops the participants were asked to design an authorization system that will route users as follows:

- An existing user will login into the system.
- A new user will register and receive authorization.

A typical design would look like the accompanying figure. The following discussion took place while the participants were working in pairs on the task.

*Ron:* Let's define login and register as objects.
*Sharon:* Do login and register seem like objects to you?
*Ron:* Why not?
*Sharon:* An object is a client, for example.
*Ron:* Client is also an object. Login and register are activated and operate within the system; therefore they can be defined as objects.
*Sharon:* I've never seen an object login.
*Ron:* Don't worry, it will be okay. You'll see how I design the system; it will be just fine.
*Sharon* [hesitates, at last reluctantly giving in]: Okay, fine, although it doesn't sound good.

*Analysis:* Ron's decision is a typical S1 behavior, similar to that observed in the bat-and-ball task. In searching for objects he is influenced by the surface features of the task (the salience of the terms login and register in the task description) rather than its essential (though implicit) components. Unlike the bat-and-ball phenomenon, Ron requires more than a nudge to change his mind, which seems to imply that his S2 knowledge in this regard is not too firm either.

Sharon, in contrast, seems to have a firmer sense of the right objects, but this too is S1 knowledge, in the sense that she cannot explain her choice. Her attempts at convincing Ron involve expressions like "I've never seen an object login," and "it doesn't sound good," which show that she relies on her vast past experience (S1) rather than on analytical rule-based reasoning (S2). Sharon's example, in contrast to the other examples presented in this article, demonstrates how using intuition may in fact contribute positively, even in situations of formal problem solving.

## CONCRETIZING ABSTRACT CLASS

*Confusing characteristics of abstract and concrete classes.* Abstract class is a class with at least one virtual function. Thus one can't instantiate concrete objects directly from an abstract class, but only through a (concrete) inheriting class. In this example, Rebecca chose to define an abstract class car and the following discussion ensued.

*Rebecca:* Let's say car is an abstract class. Then, in one design I can inherit from it Chevrolet and Rolls-Royce, and in another design I will instantiate an object car with manufacturer value Chevrolet.
*Instructor:* Is car an abstract class?
*Rebecca:* No, yes, that's not the point...

In a subsequent interview with Rebecca, the researcher probed the matter further.

*Researcher:* Rebecca, what did you mean by the car example?
*Rebecca:* I just tried to show that there are two design possibilities using an abstract class, but I got mixed up.
*Researcher:* What was the problem?
*Rebecca:* I wanted to show that you can instantiate objects with parameters instead of using inheritance tree… but it didn't work out.
*Researcher:* Why?
*Rebecca:* Because the moment I instantiate objects, I cannot define the class as abstract.

We note that this was not an isolated case. While it seems that the participants in this study recognize the distinction between abstract and concrete classes in theory, several cases were observed where they referred to abstract classes as if they had the characteristics of concrete classes. Even in some written solutions, we found cases where an abstract class was defined but was subsequently used as a concrete class.

*Analysis:* Rebecca knows the difference between concrete and abstract classes, but this is S2 knowledge. Our interpretation of how S1 worked in this example follows from the dual nature of the relationship between the natural and the formal conceptual framework concerning categories and objects. On the one hand, OOD builds on the intuitions of the natural concepts, but on the other hand, the natural system sometimes clashes with the formal one. We propose that this is what happened in Rebecca's case. Specifically, in the natural categorization system [5], there is no parallel for the formal OOD concept of abstract class (a class from which no concrete objects can be instantiated). Hence, when Rebecca's S2 was not on guard, her S1 took over and slipped from abstract to concrete class. As before, a small nudge was enough to wake up S2 and lead Rebecca to make the necessary distinction.

*Identifying software development with coding.* Coding is an important software development activity, but other no less important activities contribute to software development, such as requirements analysis, design, and testing. We observed participants underweighting these other activities, to the extent of identifying software development with coding.[2] The following discussion occurred in an interview regarding time invested in different activities:

*Ann:* Most of the time I was occupied with development.
*Researcher:* What do you mean development?
*Ann:* You know, writing the code. For me coding and developing are the same thing, even though I know this is not correct.

*Analysis:* Ann's first automatic response, that developing is the same as coding, is an S1 response. S1 consists of what is most accessible and what comes most easily to mind; here the view of development as coding comes to mind, presumably because the code is the final and most tangible product of the whole process, while the other components (such as design and requirement analysis) are less conspicuous. The interviewer's question served as a nudge that woke up S2, hence the utterance: "though I know it is not correct." In fact, her second pronouncement is a good demonstration of an actual clash between the two systems: S1 expressing the view that "coding and developing are the same thing," but simultaneously, S2 objecting that "I know it is not correct."

## HOW INTUITIVE IS OO DESIGN?

So, how intuitive is OOD? Well, in a certain sense it indeed is intuitive: our cognitive system certainly makes extensive use of objects and categories, on which this paradigm is built. However, as often happens in the evolution of formal systems, this relationship has a flip side [9]. Under the demands of abstraction, formalization, and executability, the formal OO paradigm has come to sometimes clash with the very intuitions that produced it. Thus, while objects, classes, and inheritance certainly have an intuitive flavor, their formal version in OOD is different in important ways from their intuitive origins.

Dual-process theory, imported from contemporary cognitive psychology, highlights the underlying mechanism of those situations where our intuitions clash with our more disciplined knowledge and reasoning. Or, put in Kahneman's words [4]: "Highly accessible features will influence decisions, while features of low accessibility will be largely ignored. Unfortunately, there is no reason to believe that the most accessible features are also the most relevant to a good decision."

[2]This observation was obtained in a joint study with Peleg Yiftachel.

Indeed, we have seen that, under the force of these general cognitive mechanisms, deciding on appropriate objects, classes, and relations is sometimes influenced by irrelevant surface clues or everyday meanings of these concepts, thus leading to inappropriate choices. Intuition is a powerful tool, which helps us navigate successfully through most everyday tasks, but may at times get in the way of more formal processes. We hope this article may contribute to better understanding of this problem, and point the way to thinking about its resolution. C

## References


1. Armstrong, D.J. The quarks of object-oriented development. *Commun. ACM 49*, 2 (Feb. 2006), 123–128.
2. Gilovich, T., Griffin, D., and Kahneman, D., Eds. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press, 2002.
3. Holmboe, C. A cognitive framework for knowledge in informatics: The case of object-orientation. *ITiCSE'99 Conference Proceedings*, (June 1999), 17–20.
4. Kahneman, D. (Nobel Prize Lecture). Maps of bounded rationality: A perspective on intuitive judgment and choice. In Les Prix Nobel, T. Frangsmyr, Ed. (2002), 416-499; www.nobel.se/economics/laureates/2002/kahnemann-lecture.pdf.
5. Lakoff, G. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. The University of Chicago, 1987.
6. Leron, U. and Hazzan, O. The rationality debate: Application of cognitive psychology to mathematics education. *Educational Studies in Mathematics 62*, 2 (2006), 105–126.
7. Morris, M.G., Speier, C., and Hoffer, J.A. An examination of procedural and object-oriented systems analysis methods: Does prior experience help or hinder performance? *Decision Sciences 30*, 1 (Winter 1999), 107–136.
8. OMG Object Management Group. *UML Notation Guide. Version 1.3*, 1999.
9. Paz, T. and Leron, U. The slippery road from actions on objects to functions and variables. *Journal of Research in Mathematics Education*; http://edu.technion.ac.il/Faculty/uril/papers/Paz_Leron_Actions_vs_Functions.pdf.
10. Stanovich, K.E. and West, R.F. Individual differences in reasoning: Implications for the rationality Debate. *Behavioural and Brain Sciences 23* (2000), 645–726.
11. Stanovich, K.E. and West, R.F. Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. Psychology Press, 2003, 171–230.
12. Strauss, A. and Corbin, J. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage, Newbury Park, 1990.



**Irit Hadar** (hadari@mis.haifa.ac.il) is a lecturer at the Department of MIS, University of Haifa, Israel.
**Uri Leron** (uril@technion.ac.il) is a Churchill Family Professor (Emeritus) of Science and Technology Education at the Technion—Israel Institute of Technology, Haifa, Israel.





DOI: 10.1145/1342327.1342336

By Dov Dori

# WORDS FROM PICTURES
## FOR DUAL-CHANNEL PROCESSING

Text and graphics are complementary modalities our brains process interchangeably. Conceptual modeling, recognized as a critical step in architecting and designing systems, is an intellectual activity that would greatly benefit from the concurrent utilization of the verbal and visual channels of human cognition. A conceptual-modeling framework that employs graphics and text would help alleviate cognitive loads. Object-Process Methodology (OPM) is a bimodal graphics/text conceptual-modeling framework catering to these needs. Here, I argue on behalf of the OPM holistic approach in addressing assumptions about the dual channel, as well as limited-channel capacity and active processing. To help make the case, using a running example of a car's emergency braking system, I demonstrate bimodality and complexity management via hierarchical decomposition and animated simulation to address these cognitive needs. Meanwhile, work is under way to employ some of these ideas in a future version of the Systems Modeling Language (SysML) (www.sysml.org).

*Combining graphics and text representations of complex systems, Object-Process Methodology makes it easier to understand technical ideas, whether or not one is technically oriented by nature or training.*

*When the user expresses a piece of knowledge in one modality—graphics or text—* the complementary one is automatically updated so the two remain coherent at all times.

Humans assimilate data and information, converting it simultaneously into meaningful knowledge and understanding of systems through words and pictures. During eons of human evolution, the human brain has been trained to capture and analyze images, enabling us to escape predators and capture food. In contrast, processing spoken words, let alone text, is a product of a relatively recent stage in that evolution. As our brains are hardwired to process imagery, graphics naturally appeal to the brain more immediately than words. However, words can express ideas and assertions that are way too complex or even impossible to express graphically; as an example, just try graphically representing this sentence to sense the validity of this claim. While a picture may be worth a thousand words, a word or sentence is indeed sometimes worth a thousand pictures. A problem with the richness of natural language is the potential ambiguity that arises from its use. This does not imply that pictures cannot be ambiguous as well, but graphic ambiguity is greatly reduced, even eliminated, by assigning formal semantics to pictorial symbols of things and to the relationships among them.

Diagrams aid cognitive processing due to their specificity [11], a theory proposing that graphical representations limit abstraction and thereby aid "processibility." That is, diagrams, because they usually involve fewer interpretations than free text, are more tractable than unconstrained textual notation. When corresponding words and pictures are presented near each other, learners are better able to hold corresponding words and pictures in working memory at the same time, enabling the integration of visual and verbal models [8]. A contribution of diagrams may be that they reduce the cognitive load of assigning abstract data to appropriate spatial and temporal dimensions; for example, whereas information about temporal ordering is only implicit in text, a flow diagram reduces errors in answering questions about that ordering [6].

A theory called "multimedia learning" proposed in [8, 9] is based on three main research-supported cognitive assumptions:

*Dual channel.* Humans have separate systems for processing visual and verbal representations [1, 3];

*Limited capacity.* The amount of processing that can take place within each information-processing channel is limited [1, 2, 10]; and

*Active processing.* Meaningful learning occurs during active cognitive processing, paying attention to words and pictures and mentally integrating them into coherent representations. The active-processing assumption is a manifestation of the constructivist theory in education, which focuses the construction of knowledge by one's mind as the centerpiece of the educational effort [12]. That is, in order for learning to be meaningful, learners must engage physically, intellectually, and emotionally in constructing their own knowledge.

As the literature suggests, there is great value in designing a modeling approach and supporting tool to meet the challenges posed by these assumptions. While [9] used them to suggest ways to reduce cognitive overload while designing multimedia instruction, they can also be a basis for designing an effective conceptual-modeling framework. Indeed, conceptual modeling is the active cognitive effort of concurrent diagramming and verbalization of one's thoughts. The resulting diagrams and text together constitute the system's conceptual model. A model based on a set of the most primitive and generic elements is general enough to be applicable to a host of domains yet simple enough to express the most complex systems. A sufficiently expressive model can help detect design-level errors, be reasoned about, make predictions, be communicated to other stakeholders, and evolve throughout a system's life cycle.

Such an environment would help us take advantage of the verbal and visual channels and relieve cognitive loads while designing, modeling, and communicating complex systems to stakeholders. These were key motivations some 15 years ago in my design of OPM [4]. The OPM modeling environ-

ment implementation by OPCAT[1] [5] embodies the assumptions. Stateful objects (things existing in some state) and processes (things that transform objects by creating or destroying them or by changing their states) are the building blocks of OPM. Structural and procedural links express static and dynamic relations among entities—objects, object states, and processes—in the system, and a number of refinement/abstraction mechanisms are built into OPM for complexity management.

## DUAL-CHANNEL PROCESSING

Following the dual-channel assumption, the brain simultaneously engages the visual and verbal channels (likely the two brain hemispheres) for conveying ideas regarding the system's architecture. Indeed, OPM represents knowledge about the system's structure and behavior—pictorially and verbally—in a single unifying model. When the user expresses a piece of knowledge in one modality—graphics or text—the complementary one is automatically updated so the two remain coherent at all times.

To illustrate how to account for the cognitive assumptions, I follow a stepwise example of the modeling of a car's anti-lock brake system (ABS). Figure 1(a) outlines OPCAT's graphical user interface, simultaneously displaying the graphic (top) and text (bottom) modalities needed to exploit human dual-channel processing. The top-right pane presents the model graphically in an Object-Process Diagram (OPD); the one below it in Figure 1(a) lists the same model textually in Object-Process Language (OPL). OPCAT recognizes OPD constructs (symbol patterns) and generates their OPL textual counterparts. OPL is a subset of natural English, and each OPD gives rise to a textual OPL equivalent sentence or phrase.

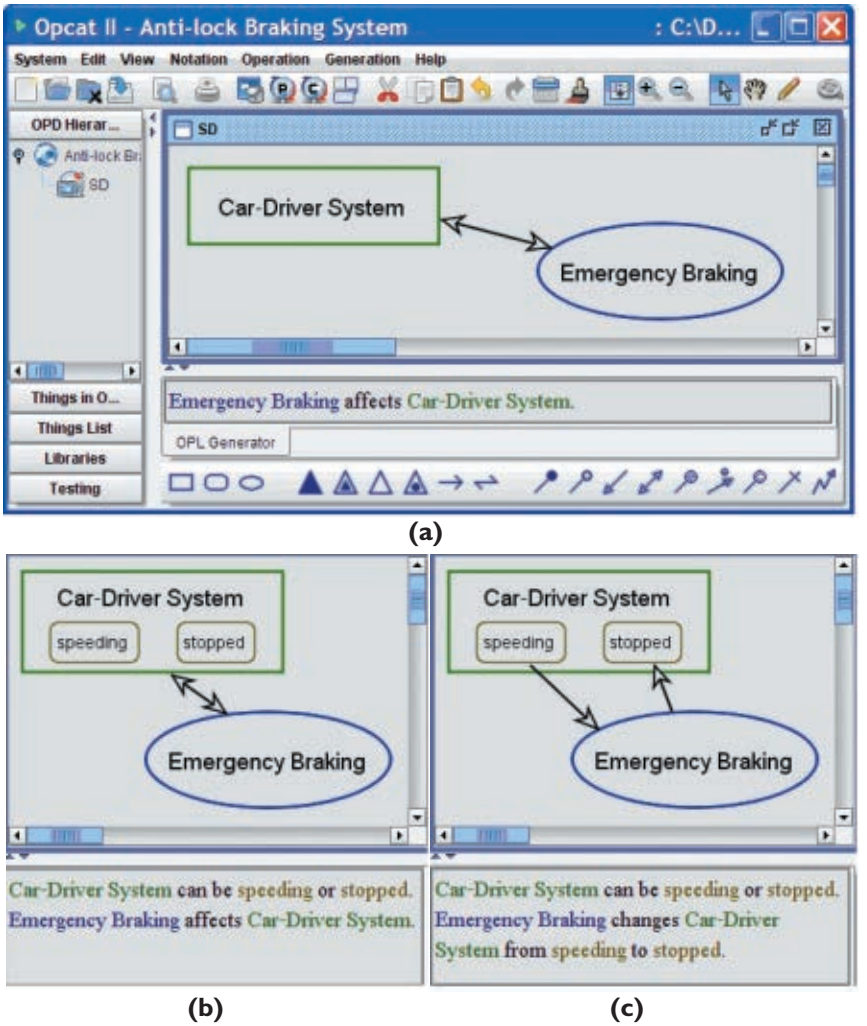



Figure 1. Top-level OPD built in stages. (a) OPCAT user interface, showing the initial system diagram of the anti-lock braking system (top) and its OPL textual specification (bottom) in which the object Car-Driver System is affected by the process Emergency Braking; (b) the states speeding and stopped are added to the Car-Driver System; (c) the input/output link pair is added from the input state to the process and from the process to the output state.

For example, `Emergency Braking`, the central system's process, is the blue ellipse in Figure 1(a), and `Car-Driver System` is an object (green box) affected by and benefiting from `Emergency Braking`. This object-process connection is expressed by linking `Car-Driver System` to `Emergency Braking` via an effect link—a bidirectional arrow indicating that the process affects the object by changing its state from unspecified input state to unspecified output state. As soon as the modeler joins the object with the process through the link, the first OPL sentence, "`Emergency Braking` affects `Car-Driver System`," shows up in the OPL pane of Figure 1(a).

As the example shows, the OPL syntax is designed to generate sentences in plain, natural (albeit restricted) English. Unlike programming languages, OPL names can be phrases (such as `Emergency Braking`). As a subset of English, OPL is accessible to nontechnical stakeholders, and other languages can serve as the target OPL. To enhance the text-graphics

[1] A research version of OPCAT is available for free download at www.opcat.com/downloads/restricted.

link, the text colors of the process and the object names in OPL match their colors in the OPD. Since graphics is more amenable to cognitive processing than text, modelers favor modeling the system graphically in the OPD pane, while the textual interpretation is continuously updated in the OPL pane and can also be continuously referenced to verify that the modeler's intent is captured.

The OPL sentences constructed or modified automatically in response to linking graphical symbols on the screen provide immediate feedback to a modeler, as well as to his/her audience. This real-time human-like response "tells" the modeler what the modeling environment "thinks" he/she meant to express in the most recent graphic-editing operation. When the text does not match the modeler's intention, the modeler can take corrective action. Such feedback is indispensable for spotting and correcting errors at an early stage in any system's life cycle, before they have a chance to propagate and cause costly downstream damage. Any correction of the graphics changes the OPL script; changes can be applied iteratively until a result satisfactory to all stakeholders is obtained. While generating text from graphics is the prevalent working mode, OPCAT also generates graphics from text.

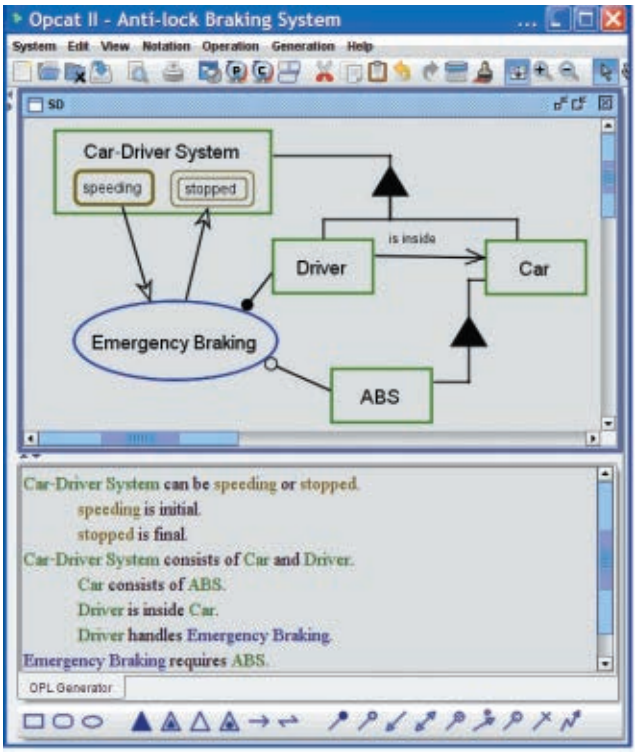



Figure 2. Top-level OPD (system diagram) resulting from adding Car, which consists of an ABS, and characterized by Speed and the effect of Emergency Braking on changing Speed from high to zero.

The System Diagram is constructed such that it contains a central process, which in this case carries out the system's main function and delivers its main value to the beneficiary for whom the system is built. In the case of anti-lock braking, `Emergency Braking` is the process that provides value to the `Car-Driver System`, the beneficiary.

Having established this basic conceptual design, we can now be more specific. In Figure 1(b), I specify the two states—`speeding` and `stopped`—of `Car-Driver System`. This specification triggers generation of the OPL sentence "`Car-Driver System` can be `speeding` and `stopped`."

By replacing the effect link with an input-output link pair consisting of an input link (from `speeding` to `Emergency Braking`) and an output link (from `Emergency Braking` to `stopped`), Figure 1(c) explicitly shows that "`Emergency Braking` changes `Car-Driver System` from `speeding` to `stopped`," as specified (equivalently) in the OPL sentence at the bottom of the figure.

The System Diagram is elaborated further in Figure 2. First, the modeler "unfolds," or specifies, the parts of the whole `Car-Driver System`. The black triangle is the aggregation-participation symbol, specifying that "`Car-Driver System` consists of `Car` and `Driver`." `Driver` is linked to `Emergency Braking` via an agent (human enabler) link (the line ending with a black circle), and `ABS`, a part of `Car`, via an instrument (nonhuman enabler) link (the line ending with a blank circle). The relationship between `Driver` and `Car` is expressed by the "`is inside`" tagged structural relationship between `Driver` and `Car`, and the states `speeding` and `stopped` are marked respectively as initial and final.

To aid active processing, humans use pencil and paper to flesh out their thoughts. OPCAT accepts pencil strokes as graphic input for objects, processes, states, and links. Pattern-recognition techniques help it recognize these graphical elements, even for sketchy pencil strokes.

## LIMITED CAPACITY

Figure 2 is the final System Diagram, the bird's-eye-view model of the system. This OPD contains about seven entities and seven links, pushing the limit of our cognitive capacity, as determined by the "magic number seven plus or minus two" concept [10]. However, we have not yet specified the subprocesses comprising the `Emergency Braking` process or the parts of `ABS`. Addressing our limited human capacity, OPM advocates keeping each OPD simple enough to enable the diagram reader to quickly grasp the essence of the system by inspecting the

OPDs without being overwhelmed by an overly complicated layout. Overloading the System Diagram or any other OPD with more artifacts would put the viewer's comprehension at risk, so showing additional detail is deferred to lower-level OPDs.

When an OPD approaches the limit of human comprehension, the model must be refined to manage the system's inherent complexity. Figure 3 outlines the newly generated OPD, labeled "SD1—Emergency Braking in-zoomed." In it, in-zooming `Emergency Braking` reveals five subprocesses and an interim object. This view is expressed in the OPL sentence "`Emergency Braking` zooms into `Braking`, `Signal Detecting`, `Boosting`, `Anti-locking`, and `Actuating`, as well as `Signal Set`." The modeler is now able to specify that `Driver` is the agent (in charge) of the `Braking` subprocess, and the `Actuating` subprocess is the one that actually changes `Car-Driver System` from `speeding` to `stopped`. The time within a zoomed-into process flows from top to bottom, so `Braking` happens first, `Boosting` and `Signal Detecting` are executed in parallel, and `Actuating` is last. `ABS` is unfolded to reveal its constituent parts (such as `Brake Assembly` and `Mechanical Subsystem`), making it possible to express procedural relations between the subprocess `Emergency Braking` and the parts of the `ABS`.

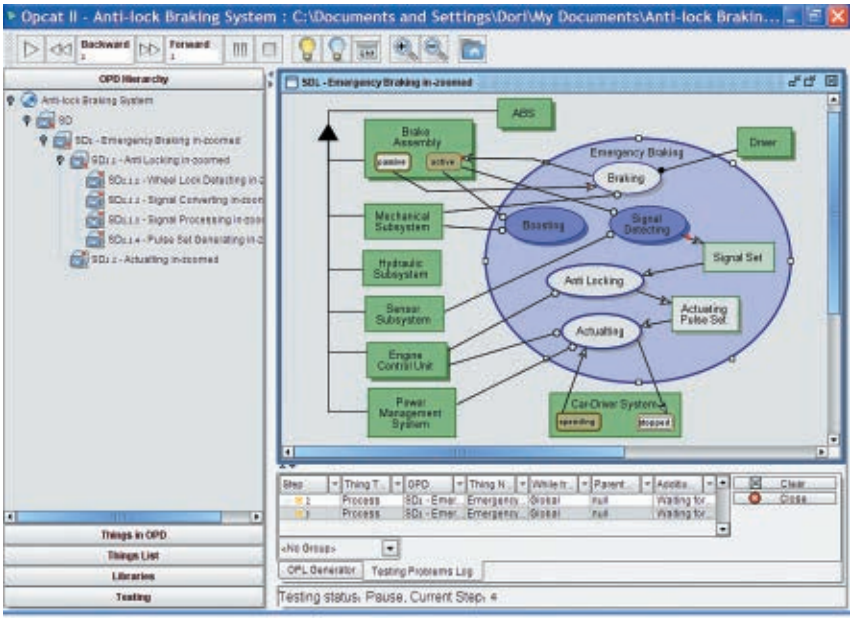



**Figure 3. The Emergency Braking process, revealing five subprocesses and two interim objects. This snapshot of the system's animated simulation serves as a design-level visual debugging tool. Shown by red dots (indicating the flow of control), Braking has just changed the state of Brake Assembly from passive to active. Boosting and Signal Detecting are executed in parallel, while Signal Set is generated by Signal Detecting.**

## ACTIVE PROCESSING

The active-processing assumption is tacitly accounted for during the conceptual modeling process in that each and every modeling step requires the complete engagement of the user—the system architect carrying out the conceptual modeling activity. When modeling, the architect places the conceived elements on the screen (possibly through the pencil tool), linking them and inspecting the OPL textual interpretation that is continuously created in response to new graphic inputs. The architect must from time to time rearrange the graphic layout to make it more comprehensible through such actions as grouping entities and moving links to avoid crossings. If the current OPD is too busy, that means it is approaching our limited channel capacity, in which case a new OPD must be created via in-zooming or unfolding.

Animated simulation is another aspect of active processing. Humans have been observed to mentally animate mechanical diagrams to aid comprehension. Using a gaze-tracking procedure, [7] found that inferences were made about a diagram of ropes and pulleys by imagining the motion of the rope along a causal chain. Similarly, an active-processing aspect of OPCAT is its ability to simulate the system by animating it. This animation enables the modeler to simulate the system and see it "in action" at each point in time during its design. Like a program debugger, the modeler carries out "design-time debugging" by running the animation stepwise or continuously (back and forth), inserting breakpoints where necessary.

Figure 3 is a snapshot of the animated simulation. Objects in green exist at this point; the white ones (such as `Actuating Pulse Set`) were either consumed already or are not yet created. Blue processes (such as `Anti-Locking` with `Boosting` and `Signal Detecting` within it) are now taking place. The active participation of the modeler (as system behavior is inspected) has proved highly valuable in communicating action and pinpointing logical design errors (corrected early on), saving precious time and avoiding costly trouble downstream.

## CONCLUSION

Technological limitations can no longer be cited as an excuse for a lack of human-centered design. A

*A conceptual modeling environment should not be merely usable but fun to use, so* users want to use it rather than feel compelled to do so because that's what they've been told.

conceptual modeling environment should not be merely usable but fun to use, so users want to use it rather than feel compelled to do so because that's what they've been told. The OPM modeling approach has adapted this philosophy into the OPCAT environment to cater to our cognitive abilities (dual processing), limitations (limited capacity), and needs (active processing). Dual-channel processing is addressed through the model's bimodal representation. Since technically oriented people usually prefer diagrams while others might favor text, individuals with both preference types are able to benefit from consulting one modality while inspecting the other. Domain experts and executives on the customer side should participate in eliciting and analyzing their system requirements, but programmer-oriented modeling approaches and environments bar such involvement.

Devoid of the cryptic syntax normally found in programming languages, the OPM model is understandable to customer-side stakeholders, allowing them to inspect the model, understand the system they should expect, and verify that the model meets their requirements. Switching between graphics and text, OPM system modelers are less likely to make costly design errors, while model readers are more likely to comprehend the system and detect design mistakes or omissions that might otherwise slip by. OPM addresses our limited cognitive capacity by providing abstraction/refinement mechanisms that enable complexity management. Active processing is facilitated by animated simulation that helps detect costly design errors. Further, OPL's formality is a basis for generating the designed application automatically. This capability reduces the modeler's manual translations of the modeled requirements, narrowing the gap between requirements and implementation.

All professions and organizations today demand lifelong learning, so designers of modeling environments must be able to account for the variety of human preferences and learning styles. The holistic OPM paradigm, with its intuitive implementation, is an example of a forward-thinking approach that could be adopted in future modeling and learning environments. Indeed, work is under way for the Object Management Group's scheduled release of SysML 2.0 in 2009 or 2010 to utilize OPM elements, including the addition of textual representations to SysML diagram types in order to achieve bimodal representation for dual-channel processing. C

## References


1. Baddeley, A. Working memory. *Science 255*, 5044 (1992), 556–559.
2. Chandler, P. and Sweller, J. Cognitive load theory and the format of instruction. *Cognition and Instruction 8,* 4 (1991), 293–332.
3. Clark, J. and Paivio, A. Dual coding theory and education. *Educational Psychology Review 3,* 3 (Sept. 1991), 149–210.
4. Dori, D. *Object-Process Methodology: A Holistic Systems Paradigm.* Springer Verlag, Berlin, 2002.
5. Dori, D., Reinhartz-Berger, I., and Sturm, A. Developing complex systems with object-process methodology using OPCAT. In *Proceedings of Conceptual Modeling—ER 2003, Lecture Notes in Computer Science 2813* (Chicago, Oct. 13–16). Springer Verlag, Berlin, 2003, 570–572, 2003; software available at www.opcat.com.
6. Glenberg, A. and Langston, W. Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language 31,* 2 (Apr. 1992), 129–151.
7. Hegarty, M. Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition 18,* 15 (Sept. 1992), 1084–1102.
8. Mayer, R. The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction 13,* 2 (Apr. 2003), 125–139.
9. Mayer, R. and Moreno, R. Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist 38,* 1 (2003), 43–52.
10. Miller, G. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review 63* (1956), 81–97.
11. Stenning, K. and Oberlander, J. A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science 19,* 1 (Jan.–Mar. 1995), 97–140.
12. von Glaserfeld, E. *The Construction of Knowledge: Contributions to Conceptual Semantics.* Intersystems Publications, Seaside, CA, 1987.



**Dov Dori** (dori@ie.technion.ac.il) is an associate professor in the William Davidson Faculty of Industrial Engineering and Management at the Technion, Israel Institute of Technology, Haifa, Israel.





DOI: 10.1145/1342327.1342337

BY Silvana Trimi AND
Hong Sheng

# *Emerging Trends in* M-GOVERNMENT

*Considering the advantages and implications of increased usage of wireless connectivity for governmental information and services.*

The emergence of new information and communication technologies (ICTs) has not only revolutionized the way business is conducted but also transformed the delivery mechanism of governmental services. Since the 1990s, public-sector organizations across the globe have been applying Internet technology and other ICTs in innovative ways to deliver services, engage citizens, and improve efficiency: a set of practices commonly known as electronic government (e-government). An explosion in the use of mobile technologies (m-technologies), such as mobile phones, laptops, and personal digital assistants (PDAs) to connect to wireless networks has enabled governments to transit from e-government to mobile government (m-government). Here, we present the potential of m-government (compared to e-government), the general trend of m-government practices in key leading m-government countries, some of the challenges and issues involved with those practices, and expected future trends.

E-government, which refers to the use of wired-Internet technology by public-sector organizations to better deliver their services and improve their efficiency, has achieved significant improvements through the deployment of many innovative applications and thus it has become a global phenomenon [5]. Based on the segment served, e-government practices can be classified into government-to-citizen (G2C), government-to-business (G2B), government-to-government (G2G), and government internal effectiveness and efficiency (IEE). Based on the type of transaction performed, e-government functions are categorized into: informational, transactional, and operational. Informational functions provide access to governmental information through Web portals, including online publishing and broadcasting. Transactional functions allow citizens to interact with government agencies via the Web such as online procurement and payments. Operational functions refer to internal governmental operations that focus on internal efficiency and effectiveness of operations and the interoperability across different e-government practices at different levels [5].

Citizens use government Web sites as central points of access to government information and services across different agencies and they normally do so via wired Internet access. In general it is the end users who are the initiators of transactions that involve e-government services. There are two technological challenges to widespread adoption and implementation of e-government. First, on the provider's side (government), technological infrastructure needs to be built to support the transformation to e-government [5]. Second, on the end user's side, there exists an inherent inequitable access to e-government services due to the digital divide among demographically, economically, and socially diverse groups of population within a country as well as among different countries.

M-government is an extension or supplement of e-government. M-government is the strategy and its implementation for providing information and services to government employees, citizens, businesses, and other organizations through mobile devices [6]. M-technology has emerged as the next wave in the IT revolution as its advantages come from two unique characteristics: "mobility" and "wireless." Mobility is the most touted advantage of m-technology. Mobile devices, typically including PDAs, laptops, cellular phones, tablet PCs, and Blackberries, free users of physical ties to the desktop. The "wireless" characteristic of m-technology refers to the method of transmitting information between a computing device and a data source without a physical connection. In recent years there has been a phenomenal growth in the use of mobile devices due to several factors including: low cost of the devices, the only infrastructure choice for many undeveloped/developing countries, changes in lifestyles, and increased functionality. According to eMarketer data, the number of wireless Internet users will surpass the number of wired users by the end of 2007, with 56.8% of all Internet users connecting wirelessly.

M-technology enables location-based services (LBS)—personalized services delivered to a mobile device user at a remote location. Since a mobile device is usually used by one user who carries it at all times, it also provides the "identity" of the user thus making it a venue for personalized services and/or for tracking/identification. The direction of communication flow for m-devices can be truly two ways, as compared to typically one way (user-to-government) for wired devices. These unique features/advantages of m-technologies increase the opportunities for governments to provide more, better, and different types of services to citizens.

M-government is value-added e-government because it offers the following advantages:

- It improves the delivery of government information and services. Citizens can get immediate access to certain government information and services on an anywhere-anytime basis. For governments, they can use the scalable and swift wireless channels to send time-sensitive information, such as terror and severe weather alerts, to citizens quickly and directly.
- M-technology may be the best solution to overcome Internet connectivity problems and digital divide issues faced by e-government applications. Unlike wired technologies, m-technologies have been more evenly distributed among different layers of society and are growing faster in economically and technologically challenged nations [4]. M-technology helps remove infrastructure constraints for countries with poor or nonexistent wired infrastructure. For example, in Africa, cell phones accounted for 90% of all African phone subscribers in 2007.
- Compared with wired networks, wireless networks appear to be a more cost-effective choice for countries with dense populations and difficult terrain. For example, Macedonia will build a Wi-Fi mesh network that will cover 1,000 square miles of difficult terrain and its service will reach more than 90% of its population for a planned price of $12 a month.
- M-government applications can help avoid problems faced by some countries, such as corruption

and low productivity of governmental agencies. India, for example, has introduced e-payment for some governmental services to avoid corruption and bribery [1].
- M-government increases efficiency and effectiveness of government employees. With the help of m-technology, government employees can access the information needed in real time and update records on the spot. This not only reduces some burden of logistics and decreases data entry errors but also facilitates employees to make informed decisions and actions.

Finally, m-government can open up additional channels for citizen participation, thus increasing constituent participation. Communication through mobile devices encourages citizens to make use of the technology to express their points of view to government officials, lawmakers, and community representatives.

## M-GOVERNMENT INITIATIVES IN THE LEADING COUNTRIES

Promoted by demands for a more responsive government, governments all over the world have recognized the potential of m-technology and are exploring the potential utility and feasibility of m-government. However, m-government development worldwide has been uneven. This is attributed to two factors: the development level of e-government and m-technologies. M-government is a value added of e-government and therefore it will be more advanced wherever a solid foundation (e-government) exists, for example, in developed countries. Second, m-technologies are quite advanced (such as 3G wireless), cost-effective, and widely used in the developed countries. Because of the relatively lower cost (compared to income) of owning and using a wireless device, the percentage of people that own and use wireless devices in developed countries is much higher than developing ones (in Hong Kong for example, mobile penetration is 125%). New important technological advances, such as WiBro (wireless broadband) service recently introduced in South Korea, and a greater variety in both platform (different types of devices) and functionalities (convergence of different devices, such as smart phones), have increased the number and types of services that end users in these countries can utilize. Consequently, while m-technology can help remove infrastructure challenges and present tremendous opportunities for the development and growth of m-government in less-developed countries, m-government applications are prevalent in the developed countries. Thus, we will focus our discussion of m-government practices in the following regions/countries.

**North America.** In the U.S., m-government projects have proliferated. Virginia has been a leader in implementing m-government applications. Its "My Mobile Virginia" was the first wireless state portal in the nation that made government services available via mobile and wireless devices [7]. It offers a variety of downloadable information on handheld devices including emergency weather conditions, legislative information, lobbyist lists, election notices, tax-related information, and tourism information.

Text messages have also been used by government agencies to communicate with citizens. For instance, the Parking Day text service in Iowa sends text messages to drivers reminding them to move their cars on street-cleaning days and opposite side of the street days. These reminders can help drivers avoid getting tickets. "My California on the Go" was introduced in July 2001 as a way for citizens to receive immediate updates on energy warnings, traffic jams, state lottery results, and press releases from the governor's office.

The Global Positioning System (GPS) is used to provide the Mobile Traffic Map in Seattle to help commuters make better drive-time choices using a small-form-factor device. The traffic map covers 176 miles of the area's major freeways and provides instantaneous information on traffic slowdowns, traffic lights, and traffic flows.

In Canada, the "Government of Canada Wireless Portal" is an evolving project that enables citizens to access government information using mobile devices. Services and information currently available include: members of the Parliament contact information, border wait time, economic indicators, passport services, and Canadian government news releases.

The large number of mobile government work forces such as those involved in law and compliance enforcement, transportation and logistics, and health and social services, depend on deployment of IEE and G2G mobile applications. Police officers can be equipped with mobile devices to access databases, issue tickets, check vehicle registrations and license tags, among other things. Examples include Wi-Fi networks in San Francisco, CA and Lincoln, NE where police officers have in-vehicle access to a variety of broadband applications such as crime databases, record-management systems, and traffic-video feeds. Field inspectors can also benefit from IEE applications using mobile devices to check inspection results, issue violations on the spot, print inspection reports, and transmit inspection information wirelessly to the agency's database. The GPS system can track a work-

er's mobile device to pinpoint his or her location and movements. The system can be used to make task assignments to workers who are closest to the job. Table 1 lists current m-government applications in North America grouped according to the segment being served into two categories: G2C and IEE.

**Western Europe.** Europe is quite advanced with regard to m-technology. The penetration rate of mobile phones in Western Europe has grown from 90% in 2004 to 98% in 2006, and is reaching 100% in 2007. The wide acceptance and use of m-technology in Europe has provided a platform for m-government applications. The London Police Department's Short Messaging System (SMS) is used to inform citizens about security threats and emergency alerts. The Bus Operator Metroline uses a mobile tracking system to monitor the status and location of buses, then send text message alerts to its drivers requesting them to speed up or slow down to maintain more even intervals between buses.

M-technology in Western Europe is used in law enforcement to provide instant access to critical data and services. For instance, German police use GPS and mobile phones to track suspects' movements. In Austria, parking inspectors are equipped with handheld devices that connect to a central parking database for an immediate confirmation of whether or not sufficient payment was made by the driver.

| | Applications | Government Agencies | Description |
|---|---|---|---|
| G2C (Government-to-Citizen) | Tracking election returns | Commonwealth of Virginia | - Allow individuals to track election returns for statewide races on the election night |
| | Mobile traffic map | Seattle | - Provide the traffic map<br>- Provide entertainment during slow-moving traffic |
| | Emergency notification | Federal Aviation Administration (FAA) | - Real-time airport status information via email |
| | Parking violation reminder | Iowa | - Reminding messages to parking violators via SMS |
| | Lobbyist-in-a-Box | Commonwealth of Virginia | - Track the progress of bills as they move through the legislative process |
| | Wireless notification | California | - Notification services through PDAs and cell phones for energy alert, lottery results, traffic updates, and articles from the governor's pressroom |
| | Wireless state portal | Virginia, Canada | - Make government services available via wireless and mobile devices<br>- Offer a variety of downloadable information, such as emergency weather situation, terrorism threats, legislative information, tax-related information, tourism information, news releases, and so forth. |
| IEE (Internal Efficiency and Effectiveness) | Field Inspection | U.S. Navy<br>U.S. Department of Energy<br>U.S. Department of Justice<br>U.S. Department of Treasury<br>Texas State Board of Barber Examiners | - Provide online access to database<br>- Input inspection results in an electronic form<br>- Transmit inspection results into the agency's database |
| | Internal communication | New York City Fire Department Commonwealth of Kentucky | - Wireless Messaging Service<br>- Wireless communication platform |
| | Police applications | Lincoln Police Department, NE<br>Michigan Police Department, MI<br>Houston Police Department, TX<br>San Francisco Police Department | Check vehicle registration<br>- Access warrant information, crime database<br>- Issue tickets<br>- Automatic traffic citation<br>- Traffic-video feeds |
| | Enhanced 911 | Maryland | - Dispatchers and rescue personnel to locate emergency callers using cellular phones |
| | Keeping track of vehicles | Traffic Management Operations Center in Portland, OR | - Keep track of vehicles flowing through Oregon<br>- Notify bus riders<br>- Help with traffic flow |
| | Tax collection solution | Texas | - Access taxpayer information remotely<br>- Print vital tax information using wireless tablet computer and bubble-jet printer |
| | Inventory Tracking | U.S. Marine Corps | - Real-time connection with warehouse management software<br>- Input and retrieve data anywhere |
| | Courthouse Wi-Fi | New Mexico | - Internet access in courthouse<br>- Voice-over-IP phone |
| | Voting machine | California | - Wi-Fi enabled system |

**Table 1. Examples of m-government applications in North America.**

Sweden is one of the leading countries of m-technologies in the world with over 95% of the population using mobile phones. Wide-ranging m-government services are available throughout the country. Some of the innovative practices include SMS applications for city job postings in Stockholm, a mobile parking fee payments system, government inspector service, tax services, mobile healthcare providers, and MapMate as a wireless map system, among others [8].

A Finnish telecommunication firm, Sonera, has developed technologies to turn any mobile phone into an identity document. Since every mobile phone has a unique Subscriber Identification Module (SIM) card code, personal identification can be embedded into that code and become as safe as the current paper passport system. The Finnish government is currently upgrading the country's electronic ID card to enable citizens to make secure transactions over mobile phones and also use it as an official travel document (instead of paper passport). Table 2 provides a summary of sample m-government applications in Europe.

**Asia.** Because of the large and dense population,

Asian countries have great potential for the use of m-technology. By March 2007, there were 1.1 billion mobile subscribers in this region. The industrialized countries, such as Japan, South Korea, Hong Kong, Taiwan, and Singapore, have some of the highest penetration rates of mobile phones in the world. Therefore, governments in these countries have implemented various m-government projects.

Japan is one of the leaders in mobile and wireless technology. The number of third-generation (3G) mobile phone users is the highest in the world, making Japan the world leader in the percentage of mobile phones supporting Internet access (94.1%) [2]. However, m-government in Japan is not as widespread as in other developed countries. Deployed m-government applications have been related to tasks that are not legally binding, such as information on tourism, disaster prevention, and child rearing. The Vehicle Information and Communication System (VICS) project provides information to and collects information from vehicles so that citizens can receive timely information services such as traffic congestion, road work, car accidents, availability of parking lots, and weather information. The Japanese government considers m-government a somewhat old-fashioned term and is striving toward u(ubiquitous)-Japan to connect everyone and everything, anytime, anywhere.

In Korea, a project called "M-police" was implemented to assist police officers in capturing suspects and finding missing cars. The project enables the officers to retrieve detailed information on missing vehicles, driver's licenses, vehicles' histories, and pictures of suspects by using mobile devices. In Anyang City, parking inspectors collect parking information and print receipts on the spot using PDAs and small printers. The m-local tax management system, introduced in the cities of Uijeongbu and Kunsan, enables officers to access information on car taxes, obtain data on delinquent taxes, and immediately transfer data to the local tax database [3].

SMS has been widely used in many Asian countries. At the height of the SARS health scare in 2004, the Hong Kong government sent text messages to six million mobile phone users in an effort to calm citizens who were frightened by rumors. Singapore government agencies send citizens text messages such as parking ticket reminders, national service obligations, and passport renewal notifications. Table 3 lists and summarizes sample m-government applications in Asia.

## CHALLENGES AND ISSUES IN M-GOVERNMENT

Despite the great potential and positive expectations, m-government is still in its earliest stage of development and its applications are limited. Governments are proceeding with caution to m-government applications due to issues, risks, and concerns inherited from e-government and some specific to m-technologies. Some of the challenges related to e-government include:

- Improving interoperability and integration since the public sector is legacy-system driven, not process-oriented driven. A radical reengineering of processes and supporting information systems is required that demands many technical, semantic, organizational, and managerial changes.
- Governments are not very receptive to private and public service providers who may contribute to innovative electronic service delivery.
- Many governments lack necessary incentives and institutional structures to realize the full potential of electronic service delivery. For less-developed countries, e-government is either nonexistent or very limited in scope. For example, China currently has more than 465 million mobile phone users, yet its m-government practices are still at the embryonic stage because of the inadequate e-government infrastructure.

M-government also faces some challenges that are unique to m-technologies. Security and privacy are considered to be the major obstacles for m-govern-

| | Applications | Government Agencies | Description |
|---|---|---|---|
| G2C (Government-to-Citizen) | SMS alerting services | London Police Departments | - Inform citizens about security threats and emergency alerts |
| | Mobile tracking systems | Metroline, London | - Track London buses using mobile communication systems<br>- Send messages to control traffic flow. |
| | SMS for job posting | Sweden | - Provide job posting for temporary workers via SMS |
| | MapMate | Sweden | - Wireless map systems |
| | Mobile parking fee payment | Sweden | - Allow citizens to pay parking fee through mobile devices |
| | SIM ID | Finland | - ID cards that serve as an official travel document as a passport does |
| IEE (Internal Efficiency and Effectiveness) | Tracking suspects | German Police | - Use GPS and mobile phones to track suspects' movements |
| | M-parking | Austria law enforcement | - Use handheld devices to connect to central database to monitor parking |

Table 2. Examples of m-government applications in Europe.

ment applications. Wireless network signals are broadcast over the public airwaves, making them vulnerable to hacking and interception. Being small and portable, mobile devices can be easily stolen or lost, putting the data stored in them at constant risk of falling into the wrong hands. Disclosing critical personal information on wireless Internet and location-based services makes citizens more concerned of privacy issues in m-government.

There are a number of inherent limitations for usability of wireless networked devices such as small screens and keypads. Wireless Application Protocol (WAP) phones can only access sites that are written in WML, which is not compatible with HTML. This means e-government Web sites must be rewritten in WML and have just enough information to fit the small screens of wireless devices. Time to translate HTML-WML along with the lack of enough broadband slows the access speed. Limited computational power and memory, shorter battery life, poor display resolution, and limited Web browsing ability are other limitations in using m-devices.

| | Applications | Government Agencies | Description |
|---|---|---|---|
| G2C (Government-to-Citizen) | SMS alerting services | Hong Kong | - Text messages to mobile phone users during SARS |
| | SMS notifications | Singapore | - Provide parking ticket reminders, national service obligations, and passport renewal notification |
| | Download non-legally binding content | Japan | - Download information on tourism, disaster prevention, and child rearing |
| IEE (Internal Efficiency and Effectiveness) | M-Police | Korea | - Police officers retrieve information using mobile devices<br>- Print tickets on the spot |
| | Parking enforcement | Anyang, Korea | - Parking inspectors collect parking lot information using PDAs<br>- Print receipts on the spot |
| | M-local tax management system | Uijeongbu and Kunsan, Korea | - Allow officers to access tax information on the spot<br>- Transfer the data to the local tax database |

**Table 3. Examples of m-government applications in Asia.**

Accessibility and interoperability are also challenges. Wireless Internet access is still costly and the coverage area of existing wireless networks is limited. In the U.S. and throughout the world, a large number of areas do not have Wi-Fi or fixed broadband connectivity. The new technology of WiMAX (which is fixed broadband wireless technology) promises to offer standardization, interoperability, and lower cost. The cities of San Francisco and Philadelphia are planning to offer WiMAX wireless Internet connections for everyone, either with a very low cost or for free.

## CONCLUSION

M-government has demonstrated great potential and shown a promising future. There is no doubt that m-government will be a part of the trend of m-commerce becoming a major part of the proliferating e-commerce in the private sector. As wireless technology rapidly advances, some of the technical obstacles for m-government could be overcome soon. Fourth Generation (4G) technology along with better wireless Internet access technologies will offer increased bandwidth for faster wireless access to the Web. More security features are becoming available, such as the Advanced Encryption Standard (AES). As the number of wireless users continues to increase, issues are resolved, and technology advances, more innovative m-government applications will emerge and m-government will become an increasingly important aspect of government functions. C

## References


1. Banerjee, P. and Chau, Y. An evaluative framework for analyzing e-government convergence capability in developing countries. *Electronic Government 1*, 1 (Jan. 2004), 29–48.
2. Information and Communications in Japan 2005. Feature: Stirrings of u-Japan. June 2005; www.johotsusintokei.soumu.go/jp/whitepaper/eng/WP2005.
3. Jeong, K. and Kim, H. After the introduction of the government portal services: Evolution into the m-government initiatives. In *Proceedings of the ICA 37th Conference*, (2003).
4. Kamssu, A. Global connectivity through wireless network technology: A possible solution for poor countries. *International Journal of Mobile Communications 3*, 3 (2005), 249–262.
5. Lee S., Tang X., and Trimi S. Current practices of leading e-government countries. *Commun. ACM 48*, 10 (Oct. 2005), 99–104.
6. Lee S., Tang X., and Trimi, S. M-Government, from rhetoric to reality: Learning from leading countries. *International Journal of E-government 3*, 2 (2006), 113–126.
7. Moon, J. From e-Government to m-Government? Emerging practices in the use of m-technology by state governments. IBM Center for the Business of Government, 2004.
8. Ostberg, O. A Swedish view on mobile government. In *Proceedings of International Symposium on E- & M-Government* (2003).



**Silvana Trimi** (strimi@unl.edu) is an associate professor of Management Information Systems at the University of Nebraska–Lincoln.

**Hong Sheng** (hsheng@umr.edu) is an assistant professor in the Department of Business Administration at the University of Missouri-Rolla.





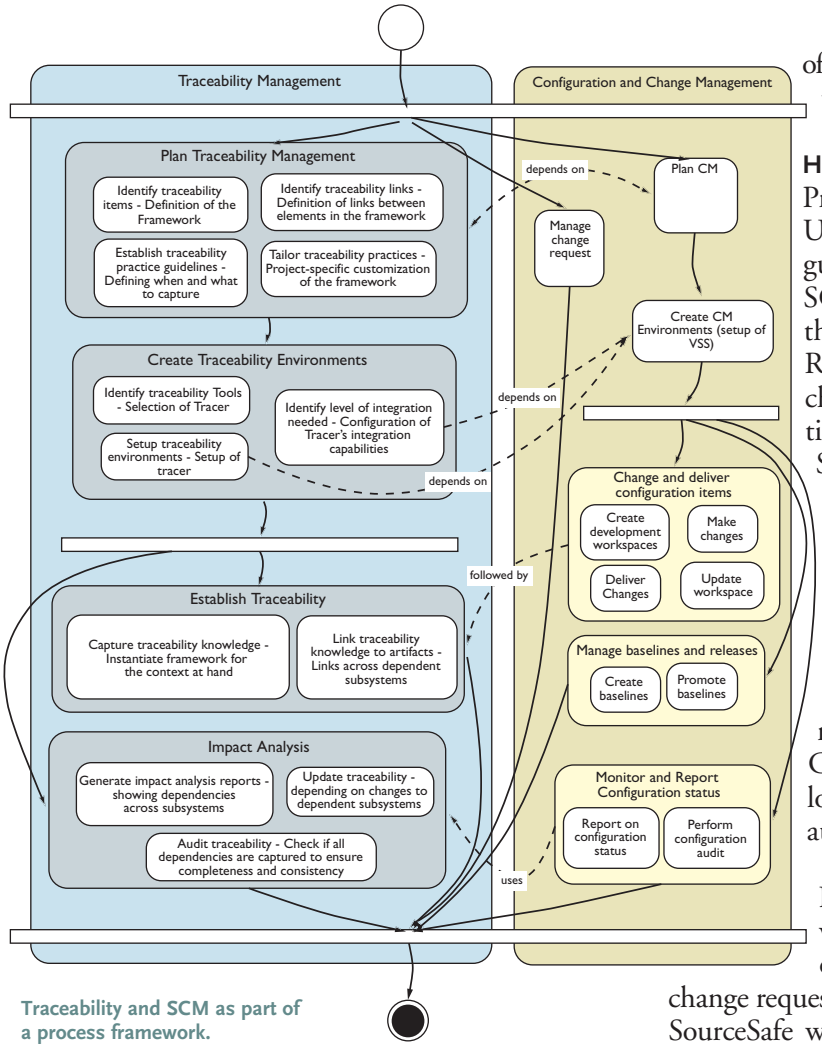
DOI: 10.1145/1342327.1342338

By Kannan Mohan, Peng Xu,
and Balasubramaniam Ramesh

# IMPROVING THE CHANGE-MANAGEMENT PROCESS

*It is crucial to integrate SCM and traceability, but significant challenges slow progress toward this goal.*

A common reason for the failure of a significant number of software development projects is the continuous evolution of software caused by volatility in customer requirements. Sources of volatility are diverse, ranging from changes in technology, evolving end-user needs, and dynamic market pressures. The evolution of software systems consumes significant resources, especially when change-management practices do not adequately support the process. Software configuration management (SCM) practices help in the management, control, and execution of change and evolution of systems [1]. Specifically, SCM helps in identifying the structure of the software product, controlling changes incorporated in software artifacts, maintaining the status of these artifacts, and generating reports for auditing and status reporting




**Traceability and SCM as part of a process framework.**

[4]. Current SCM practices and tools do not adequately support change management [3]. First, SCM often focuses on managing source code files, leaving changes to other artifacts like requirements and design documents unmanaged [2]. Second, the management of dependencies among many artifacts created during the development life cycle, which is essential for ensuring the integrity of the system [7], is typically done outside SCM tools. Traceability tools, which help link conceptual and physical artifacts created during software development are commonly used for this purpose [5]. Thus, though both SCM and traceability have a common objective of facilitating change management, they are often used independent of each other.

This leads us to the question: How can we integrate traceability practice with SCM practice and thereby improve change management processes in software development? We investigate this question by conducting a case study in an organization (hereafter referred to as Hospcom) that develops embedded software systems (see sidebar for details on how the study was conducted). We identify problems in SCM practice and recommend the integration of SCM and traceability as a means to address these problems.

## HOW IS SCM USED?

Process frameworks like the Rational Unified Process (RUP) provide detailed guidelines on how to plan and execute SCM [8]. The box on the right side of the accompanying figure depicts the RUP process for configuration and change management. Many organizations internally develop and use similar SCM processes. At the outset, the project team establishes SCM policies, writes an SCM plan, and establishes a change control process. An SCM tool like Microsoft's Visual SourceSafe is used to support these processes. Development workspaces are created, and change requests are logged, reviewed, verified, and approved. Changes incorporated in the system are logged in the SCM tool. When needed, audit and status reports are generated.

Now let us consider an example from Hospcom to illustrate the problems even with mature SCM practices. Clear procedures for documenting and fulfilling change requests were developed based on RUP. Visual SourceSafe was used to manage versions of various source code modules. Following common industry practices, these dependencies were managed at file level (rather than at the level of individual objects). Reasoning behind changes was documented with minimal details and not linked to requirement specifications, design models, or source code. As a result, when a change request was processed, developers did not fully understand its repercussions on various software artifacts and their versions.

For example, when it was necessary to support a new telephone system, Hospcom incorporated changes in the television control module. The development team documented the changes in the SCM tool using textual descriptions. Later, when incorporating a new feature in the patient billing system, the developers found it very difficult to understand the rationale behind the changes incorporated in the television control module. Traceability matrices that documented the rationale were not linked to specific versions and configurations of various software artifacts. This resulted in erroneous and inconsistent changes to the television control module.

In the absence of well-documented process knowledge (which refers to knowledge about the reasons

behind design decisions and the impact of changes caused by dependencies), developers usually implement changes based on their own (often incomplete) understanding of this knowledge. They rely on their past experience to understand design decisions made during development. This leads to poor maintenance performance, increased effort in identifying and modifying artifacts affected by the changes, low quality of changes incorporated in the system, and incomplete or inconsistent changes. Our study suggests that these problems can be addressed by the integration of traceability and SCM.

## HOW DOES TRACEABILITY ADD VALUE?

Traceability is the ability to describe and follow the life of software artifacts, such as requirements, designs, and implementation of the system [5]. It is used to maintain links or relationships between software artifacts to ensure that the design and implementation satisfy the requirements. Traceability also helps represent design rationale [6]. Since traceability can be used to document complex dependencies across software artifacts, it can help explain how making changes to artifacts affects other artifacts.

Typical traceability practice involves the use of traceability matrices or networks. Traceability matrices identify the correspondence among requirements, design, source code, and test cases, in a tabular format. Traceability networks are semantic networks that identify different types of links among these artifacts. They are often used in large-scale, complex, and mission-critical projects that require a comprehensive understanding of the relationships among artifacts [7]. Given the important role of traceability in managing dependencies, it is a natural complement to SCM in managing system evolution.

While process frameworks like RUP provide guidelines on many key process areas, they do not provide adequate guidelines for traceability practice. Since the overarching objective of SCM and traceability is change management, we suggest the synergistic integration of these practices. For effective change management, clear traceability processes should be defined and associated with appropriate SCM processes. Based on our case study, we have charted the activities involved in a traceability process. We illustrate how they can be integrated with the SCM process described in RUP in the figure. The box on the right side of the figure outlines the activities in SCM as specified in RUP. The box on the left side of of the figure outlines the traceability workflow. The figure also shows how different activities in traceability practice are associated with those in SCM.

Here, we provide examples of problems faced in each activity by stakeholders in our case study and illustrate how SCM integrated with traceability can address a variety of problems.

*Planning.* SCM plans define how to identify and monitor the evolution of configuration items (software artifacts managed within an SCM system). Hospcom developed an SCM plan and followed it in its standardized SCM practice. However, each development team developed its own traceability practice, as no organizationwide traceability plan was developed. Typically, traceability matrices that link requirements to other artifacts were created.

While the benefits of traceability are typically delivered downstream, *the costs are incurred upstream* in the development life cycle.

An integrated approach requires the establishment of a traceability plan that is tightly intertwined with the SCM plan. A traceability plan defines the traceability practice for the organization. It identifies the various traceability items (like requirements, design elements, and source-code modules) and the links among them, as well as project-specific artifacts that need to be tracked. A traceability plan may also specify the level of detail with which traceability knowledge is maintained, possible variations in traceability practice based on project characteristics like size and complexity, and potential applications of documented traceability knowledge. For example, traceability can be established at a coarse level (such as links among related files) or at a more fine-grain level (such as links between a requirement and a class in a design model). The integration of such a traceability plan with the SCM plan is essential for effective change management. Traceability items should be consistent with configuration items managed in the SCM system. The change-management process should define the correspondence between items used in these environments. For example, the correspondence between the file-level items managed by SCM tools and fine-grain traceability items that are managed by traceability networks should be clearly established. This is essential in maintaining dependencies between items at different levels of detail within SCM and traceability tools.

*Managing work-process environments.* After establishing an SCM plan, an SCM environment is created by selecting and installing SCM tools. At Hospcom, Visual SourceSafe was used for SCM. Traceability matrices were created as Word documents. These practices were used independent of each other. Change requests and changes incorporated in source-code modules documented in Visual SourceSafe were not linked to rationale documented in traceability matrices.

For synergistic operation, traceability environments should be integrated with the SCM environment so developers can work in either when incorporating changes in the system. Although such integration can be achieved with some effort through the interfaces provided by either the SCM or traceability system, the development of a common environment will be tremendously helpful. This integration can be achieved at different levels—ranging from just a simple tool invocation to interoperation with the ability to share data and metadata. Integration should be done in such a way that the SCM and traceability tools are aware of the changes in each other and rework is avoided while managing changes in either.

| Activity | SCM | Traceability | SCM integrated with Traceability |
|---|---|---|---|
| Plan | Even if organizationwide SCM plans are developed, detailed and fine-grain knowledge about dependency and rationale for changes are not managed within SCM | Typically, no organization-wide traceability plans are developed. Knowledge about dependencies and rationale are managed in traceability tools or documents in isolation from SCM | SCM and traceability plans are developed in conjunction with one another for effective change management. This approach provides guidance on the level of detail at which knowledge about changes is to be documented and linked to versions and configuration in SCM. |
| Manage work-process environments | SCM tools are typically used to manage source code modules at file level, but not other artifacts like requirements and design. | Traceability tools are typically used to manage finer dependencies across various artifacts. | Integrated work process environments for SCM and traceability enables linking of fine-grain knowledge about dependencies to specific versions and configurations. |
| Manage change requests, make and deliver changes | SCM is used in reviewing, controlling, and approving changes. Only brief textual comments are used to describe changes incorporated in specific artifacts. | Impact of changes can be identified through knowledge about dependencies and rationale behind changes. But these are not linked to affected versions of artifacts. | Specific changes to software artifacts can be documented in detail in traceability tools and linked to specific versions of artifacts. Rationale behind decisions on reviews, control, and approvals can be documented in traceability tools and linked to versions and configurations in SCM. |
| Manage baseline and release | Completeness and consistency checks are difficult by just using information from SCM tools. Artifacts managed in SCM tools are typically not linked to requirements specifications explicitly. | Though traceability tools are used to establish links among requirements, design models, source code, etc., they are not specifically linked to various versions of these artifacts. As changes may affect various versions, it is difficult to use traceability in isolation from SCM to check completeness and consistency. | The integration enables linking requirements in specifications to design elements in versions of design models and source-code modules. This enables checking completeness and consistency of baselines and releases. |
| Monitor and report configuration status | Status of configuration items can be determined using knowledge about changes to various versions documented in SCM tools. Since this is typically done using brief textual comments, this knowledge is usually not comprehensive. | Though traceability tools can be used to check status of various software artifacts, they cannot be used to check status of specific configurations of artifacts. | Configuration status reports can be more complete as they can bring together knowledge about changes to specific versions of software artifacts and how various configurations were derived. Information on composition of configurations from SCM environments and rationale behind the creation of configuration from traceability environment aree integrated. |

**Comparing isolated SCM and traceability practices with an integrated approach.**

*Managing change requests and making and delivering changes.* During SCM, developers create workspaces, make changes, and deliver the changed system. Important tasks in SCM are to accept, review, control, and fulfill change requests. Hospcom created textual documents in the SCM tool that briefly explain the nature of changes incorporated. In several instances, these documents did not provide adequate or specific details on why certain changes were done and how they affect other artifacts.

In an integrated practice, traceability knowledge acquired during development can be used to analyze the impact of changes by tracing dependency links among artifacts and identifying the ripple effect of changes. Traceability knowledge can be used to ensure that changes are made in a consistent manner. Also, when changing software artifacts, developers need to document new traceability knowledge that includes links between the changed artifacts and related artifacts, and they must link this documentation to affected artifacts so the integrity and completeness of artifacts can be ensured.

*Managing baseline and release.* SCM helps create baselines of systems or subsystems for release or reuse. When creating or promoting a baseline, developers need to ensure that all the required artifacts are archived, and the quality and reliability of these artifacts can be guaranteed. At Hospcom, during the creation of baselines and releases, developers were unable to check the completeness and consistency of the system with specific versions of artifacts due to the lack of synergistic use of traceability and SCM. For example, while traceability matrices helped identify software artifacts that implemented specific requirements, they did not provide specific links to the exact versions and configurations that implemented them. Also, when changes were incorporated in specific artifacts to accommodate changes in requirements, the traceability matrices did not direct them to the specific versions of the artifacts that were subject to these changes. Part of this knowledge was documented in the SCM tool. Thus, the knowledge required for performing completeness and consistency checks was fragmented in traceability and SCM environments.

## HOW THE STUDY WAS CONDUCTED

We conducted a case study in an organization (hereafter referred to as Hospcom) that develops embedded systems. The organization has SEI SW-CMM level-three certification and uses RUP. It has several teams working on the development of different embedded systems. The site we studied had about 300 employees, including project managers, systems analysts, and customer representatives. We studied the SCM and traceability practices followed in the development of an embedded system for telecommunication and television management in hospitals (referred to as HSys). HSys manages the communication between telephone and television management systems in hospitals. It also manages the process of billing patients who use the telephone and television in their rooms. The case study spanned a period of about two years, during which data was collected through interviews with one project manager, three project leads, six developers, and two customer representatives. We also observed the various practices followed by the team as it went through the different software development life cycle phases. The purpose of the study was to examine the change-management processes and understand the challenges faced in SCM.

The case study highlights several challenges faced by the development team in managing changes requested by customers. The study highlighted issues in current SCM practice and the need for augmenting this with traceability. It also helped in development of the traceability-process framework and the integration of relevant parts of the SCM and traceability processes.

After developing process guidelines for integrated traceability and SCM practice, we also validated them by obtaining feedback from several experienced software professionals. The ability of the framework to address limitations of current SCM practice through traceability systems was considered useful by these professionals. They also expressed interest in developing or acquiring mediating systems that facilitate interoperation of SCM and traceability tools. The addition of a generic process framework that integrates SCM and traceability into the quality department's repositories of quality guideline documents was considered valuable in shaping organizationwide practice of SCM in conjunction with traceability.

Checking completeness of a baseline or a release can be done more effectively when knowledge from traceability environments is integrated with that documented in SCM tools. For example, dependency knowledge available in traceability environments must be linked to knowledge about changes to specific versions so developers can easily navigate and examine if all applicable requirements and changes are implemented appropriately.

*Monitoring and reporting configuration status.* SCM systems help monitor and generate report projects to check the integrity of configuration items, determine whether they meet requirements, and evaluate the status of configurations. At Hospcom, the developers were unable to create reports on integrity and completeness of configuration items as the knowledge required for this was fragmented across traceability and SCM environments.

Traceability knowledge can be used to complement SCM systems in performing these activities. Traceability knowledge also needs to be audited and monitored to ensure its completeness and correctness. Given the critical role of traceability knowledge in acceptance testing (to examine whether a software system meets customer requirements and if so, how), traceability audit should be part of configuration audits. Traceability knowledge complements information stored in SCM tools while conducting audits by showing the origins of each product and process of development.

## CONCLUSION

Our framework for integrating traceability and SCM offers several benefits. The key implications we draw from our study involve aspects of SCM-traceability integration and adapting traceability and SCM processes.

*SCM-traceability integration.* The guidelines we have provided here include two types of integration of SCM and traceability that project managers should understand and implement: process integration and tool integration. Project managers should recognize the extent to which SCM and traceability processes are intertwined. This should be reflected in the organizationwide quality-guideline documents prepared by quality-assurance groups. When the development processes are tailored, tight coupling between SCM and traceability practices should be maintained. Regarding tool integration, software development organizations should consider the development of homegrown applications that can act as middleware between SCM and traceability systems. Such in-house augmentation is necessary due to the limited ability of traceability and SCM tools to interoperate.

*Adapting traceability and SCM processes.* Though we suggest the inclusion of a traceability framework as part of comprehensive process frameworks like RUP, it should be adapted to suit the needs of the organization and the project. Using a generic traceability process in all projects may be counterproductive if it is

implemented without consideration of project characteristics like complexity, size, and regulatory guidelines. While mission-critical projects may adopt sophisticated traceability processes, agile software development situations may find that these may constrain their development agility. Therefore, organizations involved in agile development need to create lightweight versions of the general process framework to maintain speed of development without significantly sacrificing product quality. The accompanying table summarizes how an integrated approach provides a synergistic change-management practice when compared to SCM and traceability.

Although we have emphasized the importance of integrating SCM and traceability, we recognize significant challenges in achieving this integration. For example, it imposes a considerable amount of overhead. Also, while the benefits of traceability are typically delivered downstream, the costs involved are incurred upstream in the development life cycle. Therefore, incentive schemes that recognize this issue should be developed to motivate the contributors and consumers of traceability knowledge. Further, the adoption of an integrated practice that deviates significantly from current practices is likely to encounter resistance unless appropriate incentives, tool, and process support are provided. C

## References


1. ANSI/IEEE-828-2005 IEEE S*tandard for Software Configuration Management Plans*. IEEE Standards Association, 2005.
2. Chu-Carroll, M.C., Wright, J., and Shields, D. Supporting aggregation in fine-grained software configuration management. In *Proceedings of the 10th ACM SIGSOFT Symposium on Foundations of Software Engineering* (Charleston, SC, 2002), ACM Press, NY, 99–108.
3. Conradi, R. and Westfechtel, B. Version models for software configuration management. *ACM Computing Surveys 30*, 2 (1998), 232–282.
4. Hass, A.M.J. *Configuration Management Principles and Practice*. Addison-Wesley, 2002.
5. Palmer, J.D. Traceability. In Thayer, R.H. and Dorfman, M. eds. *Software Requirements Engineering*. IEEE Computer Society Press, Los Alamitos, CA, 1997, 364–374.
6. Ramesh, B. and Dhar, V. Representing and maintaining process knowledge for large-scale systems development. *IEEE Expert 9*, 2 (1994), 54–59.
7. Ramesh, B. and Jarke, M. Towards reference models for requirements traceability. *IEEE Transactions on Software Engineering 27*, 1 (2001), 58–93.
8. Rational Unified Process. IBM, June 15, 2005; www-306.ibm.com/software/awdtools/rup/.



**Kannan Mohan** (kannan_mohan@baruch.cuny.edu) is an assistant professor of computer information systems at the Zicklin School of Business, Baruch College, City University of New York, NY. **Peng Xu** (Peng.Xu@umb.edu) is an assistant professor of management science and information systems at the College of Management, University of Massachusetts in Boston. **Balasubramaniam Ramesh** (bramesh@gsu.edu) is a Board of Directors Professor of computer information systems at the J. Mack Robinson College of Business, Georgia State University in Atlanta.





DOI: 10.1145/1342327.1342339


# CALL FOR PARTICIPATION

## 2008 ACM International Conference on Computing Frontiers

*May 5 - 7, 2008, Ischia, Italy*

**Sponsored by ACM - SIGMICRO**



The increasing needs of present and future computation-intensive applications have stimulated research in new and innovative approaches to the design and implementation of high-performance computing systems. These challenging boundaries between state of the art and innovation constitute the computing frontiers, which must push forward and provide the computational support required for the advancement of all science domains and applications. This conference focuses on a wide spectrum of advanced technologies and radically new solutions, and is designed to foster communication between the various scientific areas and disciplines involved.

The technical program will include 3 keynote talks, a poster session on innovative computing and technical sessions covering all areas of innovative computing systems that extend the current frontiers of computer science and engineering and that will provide advanced systems for current and future applications. The conference also includes two workshops and a special session on FP7 Computing Systems Projects.

### Keynote Talks

*Keshav Pingali*, UT Austin, Data-parallel Abstractions for Irregular Programs

*Valentina Salapura*, IBM Watson, Scaling Up Next Generation Supercomputers

*Raffaele Tripiccione*, U Ferrara, JANUS: Reconfigurable High-Performance Computing for Physics

### Technical Session Topics

- Innovative Computing Platforms (Cell, GPGPU, FPGAs)
- Programming Models
- High-Performance Computing
- Microarchitecture for Low Power, High Reliability and New Application Domains
- Innovative Design
- Innovative Compilation Techniques
- Innovative Systems
- Poster Session

### Workshops and Special Session

*Workshop on "Memory Access on Future Processors: A Solved Problem?"*

*Workshop on "Radiation Effects and Fault Tolerance in Nanometer Technologies"*

A *Special Session on the European Union's Seventh Framework FP7 Computing Systems Projects* at Computing Frontiers 2008 will serve as the first public forum to present all the accepted proposals in the first FP7 call.

### Conference Venue

The conference will be held on the Mediterranean island of Ischia in the Gulf of Naples. The conference hotel is the Hotel Continental Terme in the heart of Ischia and features one of the largest and best equipped spas of the island.

For more information visit the conference web site at http://www.computingfrontiers.org

BY RUI CHEN, RAJ SHARMAN, H. RAGHAV RAO,
AND SHAMBHU J. UPADHYAYA

# COORDINATION IN EMERGENCY RESPONSE MANAGEMENT

*Developing a framework to analyze coordination patterns occurring in the emergency response life cycle.*

Emergency Response Management (ERM) enables and supports emergency response operations across organizational, jurisdictional, and geographical boundaries. Recognizing the growing importance of ERM in countering both natural and manmade hazards, the U.S. government ordered (via Homeland Security Presidential Directive-5) the establishment, at the federal level, of a National Incident Management System (NIMS) [4]. The NIMS prescribes institutional response guidelines that help in establishing rule structures and developing a normative environment with defined tasks regarding what should be done during a response. Howitt and Leonard [6] point out that while

ILLUSTRATION BY JON KRAUSE

NIMS does include a unified approach to incident management and incorporates standard command and management structures and aids coordination, it has certain limitations. For example, NIMS is a technical system that can function effectively when its goals in a particular situation are consistent, clearly prioritized and coherent. However, when situations present complex value conflicts or trade-offs, NIMS lacks the ability to make politically legitimate decisions and to mobilize public support for subsequent action [6].

Effective coordination is an essential ingredient for ERM. The coordination of emergency response is demanding as it involves requirements typical of an emergency situation that include, for example, high uncertainty and necessity for rapid decision making and response under temporal and resource constraints. Yet, the available literature on coordination issues relating to ERM consists largely of practitioner articles, governmental reports, and testimonies to Congress. Academic research in this area, other than [2, 3, 8, 11], is scarce. Given the importance of ERM coordination, this area needs to be studied in greater detail. In this article, we propose a framework to analyze emergency response coordination patterns, based primarily on semi-structured interviews with 32 emergency response personnel, including town, city, county, and state emergency managers and Federal Emergency Management Agency (FEMA) coordinators. We also illustrate the usefulness of the framework by applying it to an actual incident.

| Challenges | Coordination Support Activities |
|---|---|
| High uncertainty, sudden and unexpected events | - Real-time monitoring and timely alert notification for situation awareness<br>- Improvisation and rapid adaptation of predefined plans to the scenario<br>- Periodic evaluation and update on existing coordination practices |
| Risk and possible mass casualty | - Threat, vulnerability, and risk assessment and countermeasure<br>- Risk-sharing policy among parties involved<br>- Operational sustainability management |
| Increased time pressure and urgency | - Repository of related plans, procedures, policies<br>- Knowledge base and network of internal and external experts<br>- Efficient information and intelligence mining; knowledge elicitation<br>- Decision support technologies<br>- Psychological fortitude to deal with effect and behavior |
| Severe resource shortage | - Policies (such as priority list and access control) for resource use and requisition<br>- Logistic management and resource-sharing network across local, national, and international levels<br>- Self-equipped response teams<br>- Law and order, price control mechanism management |
| Large-scale impact and damage | - Joint effort of governmental, public, and private sectors across local, national, and international boundaries<br>- Broad information-, intelligence-, and resource-sharing networks<br>- An integrated public communication network to inform, guide, and reassure the general public |
| Disruption of infrastructure support | - Protection of critical infrastructures<br>- Performance monitoring of built structures<br>- Planning for infrastructure interdependencies such as proximity of foliage and civil infrastructure<br>- Control of infrastructure redundancy<br>- Management of alternative infrastructures |
| Multi-authority and massive people involvement | - Unified response command for coordination<br>- Establishment of role structures with corresponding authority, responsibility, and accountability<br>- Management of power and regulation conflict<br>- Exercise of leadership and norms<br>- Communication operability and interoperability |
| Conflict of Interest | - Understand the political, ethnic, economic, and environmental impact<br>- Shared vision and alignment of core interests<br>- Reconcile the objectives of various involved parties in a politically sensitive and emotionally charged environment<br>- Shared vision of priority among responders and the public |
| High demand for timely information | - Information gathering and provisioning<br>- Information fusion and validation<br>- Information exploitation and dissemination |

Table 1. ERM coordination at a glance.

## COORDINATION IN ERM

The coordination of emergency response is challenging because it involves factoring in exigencies typical of an emergency situation such as great uncertainty; sudden and unexpected events; the risk of possible mass casualty; high amounts of time pressure and urgency; severe resource shortages; large-scale impact and damage; and the disruption of infrastructure support necessary for coordination like electricity, telecommunications, and transportation. This is complicated by factors such as infrastructure interdependencies; multi-authority and massive personal involvement; conflict of interest; and the high demand for timely information. Table 1 elaborates some of these issues based on our conversations with emergency responders.

## LIFE CYCLE APPROACH

A life cycle approach provides a broad and systematic view of the activities relating to emergency response management [12]. Therefore, the framework we suggest is adapted to each of the stages in the life cycle. The management of emergency response can be visualized in terms of three distinct sets of activities on the time line continuum [4]. These include actions taken prior to an incident (typically concerning preparedness issues such as planning and training), during the incident, and after the incident. The cycle is completed when de-briefing has occurred and the lessons learned are framed as actionable items designed to affect future preparedness. Many of the core elements of ERM coordination (such as activities, coordination objects, and constraints) differ from stage to stage [12]. Cultural, political, regulatory, and infrastructural (civil structures, people, process, and technology) issues all have an impact on coordination patterns and outcomes. In Figure 1, we present the schema of the framework we developed, which represents not only a development of the work presented by Raghu et al. [9], but also a context modification of that work. The framework considers five basic elements that are applied to each stage of the life cycle:

• Task flow: tasks and interdependent relationships;

| Pre-incident Phase | | Coordination Issues | Coordination Goals | Coordination Mechanisms and Support |
|---|---|---|---|---|
| Task Flow | | Standard operating procedure development, planning for task flow dependencies, and so forth | Exact mapping from objectives to action-oriented checklist, task-sub task, goal decomposition for efficient implementation | Operation routine, task synchronization, task sequencing, training, exercise (table-top or drills) |
| Resource | | Local and external resource management (such as critical infrastructure utilization in neighboring counties, contra-flow of traffic, public hospital utilization), request schemes, and so forth | Prioritizing, resource readiness management, establishment of acquisition mechanisms for external resource | Resource deployment and usage priority schemes, guidelines, resource standardization, mutual aid, donor assistance, inventorying |
| Information | | Risk assessment (for example, chemical plants close to nursing homes), information on performance of civil infrastructures, information exchange, communication connectivity, risk assessment for engineered and built infrastructure, obtain information on tightly coupled infrastructure interdependencies, and so forth | Integrated and interoperable communication for agencies and communities, standardized communication instructions and report formats | Communication protocols, data vocabulary, and message standards, communication operational guidelines, public alert systems |
| Decision | | Decision structure development | Establishment of clear decision roles, compatible decision rules | Guidelines, protocols, knowledge sharing, community of practice (COP) structuring |
| Responder | | Developing an organizational structure among relevant responder groups to create a scalable organization maintaining span of control, and so forth | Clear individual authority, responsibility and accountability; ensure that responders know and understand their role within the Unified Incident Command System (ICS) structure (also ensure there will be no turf disputes among responders) | Guidelines, protocols, team meetings, mutual adjustment mechanism (for example, liaisons), training with the ICS structure, build familiarity and working relationships with local, state, and federal partners, building structure to address mental welfare of first and second responders such as social organizations and so forth |
| **During Incident Phase** | | **Coordination Issues** | **Coordination Goals** | **Coordination Mechanisms and Support** |
| Mini-Second Coordination Cycle (Reactive) | Task Flow | Detailed task assignment, task scheduling, and so forth | Fit between task requirement and personnel expertise, smooth functioning of task flow, no conflicts, overlap, or duplication of efforts | Routines, schedules, synchronization, notification, sequencing, tracking |
| | Resource | Resources utilization, reallocation, tactical use, and so forth | Efficient resource sharing among personnel, efficient resource positioning and utilization | Priority order (need based) |
| | Information | Direct information exchange between responders as well as with EOC | Good communication between responder groups, timely information flow | Face-to-face communication, information-forwarding mechanisms, interoperability standards (for example, EDXL) and devices (for example, 800Mhz), ensure communication despite the failure of technology |
| | Decision | Decision-making speed for rapid intervention and prevention | Efficient decision making in terms of time and loss of life and property; ensure proper rest and mental stability of first responders | Decision role delegation, reliance on protocol, support for decentralized decision making, provide training, make mental health practitioners available to first responders on site if possible to avoid PTSD |
| | Responder | Responder relationship, group dynamics, groupthink | High team spirit, ensuring trusting relationships in the completion of tasks | Coordinated pre-incident exercise, open interpersonal communication |
| Many-Second Coordination Cycle (Proactive, More Reflective) | Task Flow | Design and management of entire response activity, addressing task flow dependencies at macro-level, addressing scheduling of external activities, and so forth | Achievement of overarching response objectives, integration of local and external activities (State, Federal, NPO, private, public agencies) | Goal selection, task decomposition, plug-and-play teaming |
| | Resource | Local and external resource utilization, staging of logistics, and so forth | Efficient acquisition of external resource and optimized allocation of available resource | Plan of mutual aid, donor assistance, resource typing, inventorying, requesting, procuring, tracking, priority, and reallocation |
| | Information | Global information utilization | Integrated and improved information, global operational picture, complete situational awareness | Multi-source information acquisition, centralized information analysis, secure "publish/subscribe tools" and public collaborative sharing for information collection and distribution based on social networking (for example, MS Groove) |
| Note: Including the interactions between mini- and many-second coordination cycles | Decision | Decision-making quality | Ensuring proper decisions dependent on global picture and strategic needs, changing decisions based on developing needs | Knowledge sharing, protocol breaking, joint sense-making, epistemic contestation, COP structuring |
| | Responder | Multi-agency conflicts arising out of cultural differences, responder welfare issues, management of liaison with media, politically sensitive issues, providing clarity on legal matters, and so forth | Compatible goals, policies, interest, address jurisdictional policy difference (for example, cross-border issues to deal with the sick and infected); quarantine extensions in neighboring counties in case of a pandemic disaster, regional resource coordination to deal with surge in public health systems, and so forth | Team meetings, mutual adjustment, cross-boundary intervention, risk sharing, resolve turf disputes among responders as necessary |
| **Recovery Phase** | | **Coordination Issues** | **Coordination Goals** | **Coordination Mechanisms and Support** |
| Task Flow | | Response and recovery planning to return region to normalcy as soon as possible (for example, debris removal, restoration of utilities) | Achievement of overarching recovery objectives, integration of local and external activities | Goal selection, task decomposition, plug-and-play teaming, operation routine, task synchronization, task sequencing |
| Resource | | Local and external resource management | Efficient acquisition and distribution of resources, fast demobilization, quick maintenance | Resource deployment priority, guidelines, resource standardization, mutual aid, donor assistance, replenishment |
| Information | | Publication of information that has been collected about recovery, information on pricing for award of contract for post-disaster activities | Consistent information content, timely distribution of public information and education to reduce public fear and build confidence, building community spirit, and so forth | Centralized information publishing, joint network of public and private media |
| Decision | | Address contracts for debris clean-up, aid decision making with county officials to arrive at consensus about financial costs relating to recovery, and so forth | Strategic decision (short- and long-term) in terms of outcome, restoration of normalcy, removing constraints such as residency and driving bans, and so forth | Priority, guidelines, protocols, knowledge sharing |
| Responder | | Agency debriefing and evaluation, addressing first and second responders long-term issues | Free discussion, healthy relationship | Facilitator, boundary-spanner, guidelines, protocols, team meetings, mutual adjustment |

Table 2. Emergency coordination phases.

- Resource: resource utilization management and dependencies;
- Information: task-critical information collection, analysis, and distribution;
- Decision: decision roles, rules, and structures; and
- Responder: relationships, team-think, group dynamics (such as culture), organizational dynamics, and so forth.

This framework conceptualizes the "during-incident" response stage as comprising two distinct coordination patterns: On-site response coordination (Mini-Second Cycle) and Remote response coordination (Many-Second Cycle).

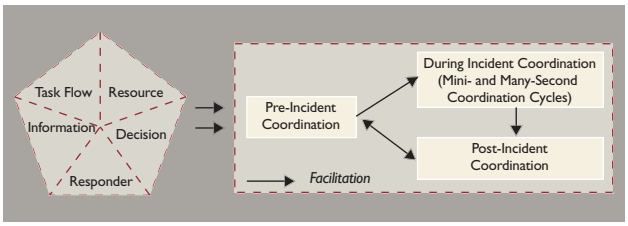



Figure 1. Emergency response coordination life cycle.

## COORDINATION LIFE CYCLE

Here, we discuss coordination patterns along the entire coordination life cycle, based on our proposed framework.

**Coordination in Pre-Incident Response.** Pre-event coordination establishes the level of operational capacity and overall readiness for resilience during emergency response. A typical disaster includes several invariants, defined as those factors that remain unaffected by the changing conditions of the emergency, such as: creating emergency shelters in appropriate places; dealing with a surge in hospital admissions; working with degraded capacities; maintaining law and order; arranging evacuation across geographic boundaries; and other factors. These issues are addressed during coordinated planning and training exercises involving the stakeholders and results in the development of Standard Operating Procedures (SOP). Planning also addresses issues such as setting up contractual agreements with business entities for providing supplies during an incident and creating infrastructure to deal with first and second responder issues (including effect and behavior). During major emergencies, the limits of local capability are soon reached and multiple agencies are involved in supporting additional response efforts. This typically requires both spatial and temporal coordination with organizations and personnel who follow different norms and practices. Training and exercise help in establishing necessary understanding between different players (whether from the same agency or from different ones) and catalyzes smoother interaction between them during an actual incident. Setting up such training activities and table–top exercises also requires coordination. Therefore, coordination is a key issue in pre-incident activities. Table 2 includes the application of the framework to pre-incident activities.

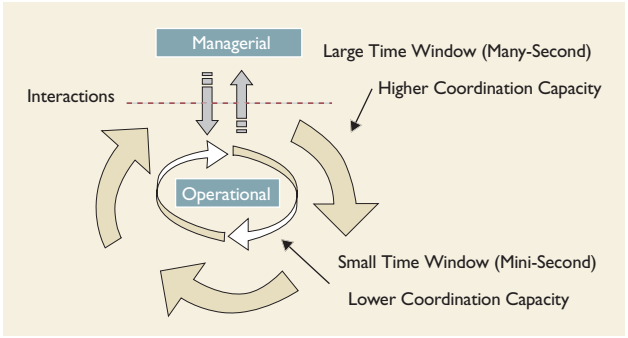



Figure 2. Mini-second and many-second coordination cycles.

**Coordination During Incident Response.** Coordination during an incident impacts both short-term and long-term outcomes. A plan-based approach to emergency response relies heavily on pre-incident preparedness and this sometimes leads to response inflexibility in the face of unexpected events. Variants in a disaster originate from hazard uncertainty; uncertainty as to the course of incident development; informational uncertainty; task flow uncertainty (whether sequential, consequential, or cascading); organizational structure uncertainty; and environmental uncertainty. Uncertainties are managed by improvisations, prioritization, and dynamic sourcing of capacities from other communities and external agencies, such as neighboring counties, state and federal agencies [5]. The variant or situation-dependent layers of knowledge create a context from which one can then understand the Incident Commander's intent. These layers may indeed serve as temporal agents during mitigation.

To support fast response during complex incidents, responders must make rapid coordination decisions, which pose constraints on their capabilities to analyze coordination problems and explore the solution domain. Response to disasters can be viewed as consisting of an onsite response coordinating entity and a remote management entity such as an emergency operations center (EOC). Onsite response is usually reactive and the time window for coordination is small. We characterize this as the "Mini-Second Coordination Cycle." It is typically characterized by working with the local picture stemming from the local scenario.

Without a proper understanding of the global pic-

**Table 3. Application of framework to CSX train derailment and tunnel fire incident, Baltimore, Maryland, 2001.**

| During-Incident Phase | | Example Coordination Issues | Example Coordination Goals | Example Coordination Mechanisms and Support |
|---|---|---|---|---|
| Mini-Second Coordination Cycle | Task Flow | Interior fire attack operations on the fire in the tunnel, support operations dealing with water supply, security, medical services, and so forth | Task division and scheduling such as:<br>- Interior attack teams laid hose lines and entered the hazard area to find and extinguish the fire<br>- During the interior attacks, the "second in" engines functioned as water supply to support interior attacks<br>- The police secured and blocked the surrounding area and ensured that only the response resources went through<br>- While the 5"-diameter hose was lowered into the tunnel through the manhole, firefighters entered the tunnel and attached their hose lines to the 5"-water supply and started firefighting<br>- Individual cars from the end of the train were removed to enable firefighters to attack the fires in the other cars | Standard Operation Procedure (SOP), tasks assignments were made and assumed as pre-determined plans with minimal specific direction |
| | Resource | Resources utilization of personnel, heavy and light rescue equipment, water supply, and communication frequency from Baltimore City Fire Department, City Police Department, City Emergency Management, Baltimore County Fire Department | Efficient and effective use of resources<br>- Allocation of a specific radio frequency for ground operations, maintenance, and management functions<br>- Allocation of water resource, supplied through a 5" water supply manifold through a manhole among interior attack teams<br>- Efficient use (distribution and refill) of personal protective equipment such as Self Contained Breathing Apparatus (SCBA) among firefighter shifts who entered the tunnel<br>- Efficient use of rehabilitation services, provided by EMS, for the firefighters<br>- Allocation of crews (engines, trucks, and battalions) among the Camden sector (north end), Mt. Royal sector (south end), and Tower sector | Resources were deployed as per predetermined plans, and priority-based cause-effect analysis |
| | Information | Communication between the IC, the fire attack teams and the supportive teams (such as police, EMS, and Haz-Mat) | Real-time information exchange among the interior attack teams for safety awareness and assistance, fast information sharing between the interior attack teams and incident commander for situation report | Baltimore Fire Department 800MHz radio systems, hand signals, narratives |
| | Decision | Decision making for ground operations such as immediate fire suppression and containment | Effective interpretation of operational challenges (sizing up the situation and making resource assessment) to develop rapid intervention schemes with appropriate level of personal and operational safety | Incident priorities (life safety, property conservation and extinguishment), Standard Operation Procedures (SOP), Haz-Mat precautions (flame and chemical contact protection, supplied air respiration, the hazards of Boiling Liquid Expanding Vapor Explosion), building safety guidance (in the face of weakening structural integrity of the tunnel and immediate surrounding areas) |
| | Responder | Responders and helpers learn to work together | Confident and trustful relationships among the responders and among the helpers, clear accountability, accurate expectations of team members' actions and capabilities | The opportunity of response partners to exercise together in pre-incident training and drills, to instill personal and team confidence |
| Many-Second Coordination Cycle | Task Flow | Supervise and support the ongoing tunnel firefighting and public welfare | Timely and necessary operation interventions to ensure the personal safety (responders and the public) and task progress of the entire response effort<br>- Appropriate interruptions of the attack on the fire to avoid both exposure to risks and waste of resource<br>- During the response, the U.S. Coast Guard deployed a series of floating booms to protect the Inner Harbor against contamination and potential hazardous runoff from the derailment site<br>- While fire was occurring, citizens of Mount Royal Station were offered the choice to leave for the "shelter-in-place" | Goal selection, task decomposition, plug-and-play teaming, coordination by expertise |
| | Resource | Management of additional response resources from the City of Baltimore Department of Public Works (DPW); from the Representatives of the South Baltimore Industrial Mutual Aid Plan, CSX Transportation, the Maryland Department of Environment; the National Transportation Safety Board and the U.S. Coast Guard | Efficient acquisition of additional resources<br>- Apply mutual aid resources to support fire scene operations and stand-by services in event of additional alarm calls<br>- Allocating and maintaining the resource staging area for rapid distribution and demobilization of resources | Need-based request (multiple alarm calls), plan of mutual aid |
| | Information | Global information collection, analysis, and dissemination | Efficient management of task-critical information among stakeholders<br>- The on-site command team kept in close contact with the CSX Transportation companies for train- and cargo-related information<br>- Hazard analyses of smoke, air, and water were quickly delivered to the on-site command team from the Maryland Department of Environment (MDE), U.S. Environmental Protection Agency (EPA), U.S. Coast Guard and CSX contractor continually during the response process<br>- Public announcements (including civil defense sirens) were made over radio and television to alert citizens and to initiate a "shelter-in-place" advisory | Computer Aided Dispatch (CAD) system, Baltimore City Communication Center, dispatch service, public and private networks |
| | Decision | Effectiveness of the action plans (response progress and resources capability), develop a strategic response with critical stakeholders involved (such as city management) to control and limit the damage to property, the environment and the welfare of the citizens of Baltimore | Effective evaluation of the response situation and timely transitions of overall response strategies for higher effectiveness<br>- From initial aggressive attacks (initial direct attacks through both ends of the tunnel) to passive attack (waiting for the fire to burn out) to defensive attack (by taking advantage of the rupture of the water main) to aggressive attack (attack through manhole on the street)<br>- With the reported water main rupture on the street right above the tunnel, the on-site management team negotiated with the Baltimore DPW to keep the water flowing from the ruptured water main into the tunnel for two hours. This decision effectively cooled the tunnel and prevented a Boiling Liquid Expanding Vapor Explosion. | Protocol breaking, joint decision making, cost-benefit analysis |
| | Responder | Management of relationship with the external agencies | The on-site management team placed confidence on the information provided by the MDE, Coast Guard, CSX contractor, EPA while issuing orders to the interior fire teams. The onsite management team and the city management trusted each other to utilize the water breakout for firefighting at the cost of massive impact on the city's water supply and risk of necessitating reconfigurations of the city water systems. | Working experience with related agencies through training; develop strong partnership with agencies responsible for responding to incidents; institute predefined plans |

[1] James Guy (Ex-Fire Chief and now Chief of Environmental Affairs at the State University of New York at Buffalo) and Dave Humbert (Fire Chief – North Bailey, Amherst, New York) provided valuable insight into the Baltimore train derailment incident.

ture, actions are motivated as a reaction to incidents from the immediate scene. Good coordination in civilian structures is better motivated by fostering common understanding and this is accomplished by creating a common operating global view that lays out the commander's intent and strategies. Efficient communication is an essential ingredient in the development and spread of common understanding and buy-in. A supervisory structure such as EOC deals with more strategic issues and works with a global picture, leveraging external resources to help on-site response. The actions of the EOC emanate based on a more reflective and proactive posture and the EOC commanders typically operate with a large time window. We have therefore classified such coordination efforts as "Many-Second Coordination Cycle." This concept (see Figure 2) is an adaptation from the work by Lewandowski et al. [7] in the area of survivable autonomic response architecture.

effected, to learn from the incident so as to positively impact the building of resiliency to better deal with future incidents. It is also a time to replenish the consumable supplies and to return the response capacity back to readiness against new incidents in the future. Unless properly coordinated, the recovery may introduce new "disasters" for the incident victims and tangibly impact the budget.

**Framework.** In Table 2, we present a framework to analyze the coordination effort for managing response to an emergency. We apply the framework to all the three phases of the emergency life cycle.

## APPLICATION OF FRAMEWORK

Here, we demonstrate the real-world application of the coordination framework presented in the previous section to the "during incident" management of an actual incident.

To support fast response during complex incidents, *responders must make rapid coordination decisions,* which pose constraints on their capabilities to analyze coordination problems and explore the solution domain.

The concepts of mini-second and many-second coordination cycle relate to distinct coordination tasks (operation- vs. managerial-level); constraints (small vs. large time window, information/intelligence and capability); and outcome quality (poor vs. good). Mini-second coordination addresses immediate response coordination needs while many-second coordination oversees and supports the former, for instance with resources and information.

This division of coordination tasks and responsibility allows better matches between coordinator expertise and task requirements [1, 10]. Frontline response teams are trained to excel on domain-specific tasks (like firefighting and rescue) and the coordination of these tasks. Remote commanders focus on global issues such as inter-agency coordination, overall logistics, and regulation compliance.

**Coordination in Post-Incident Response.** Effective response and recovery is vital to the economic health of the affected region and also to the mental health of its citizens. Recovery focuses on the return to normalcy of the impacted region and people. It is also a phase for debriefing and pondering the details of the response

At 3:07 P.M. on Wednesday, July 18, 2001, a CSX Transportation train derailed in the Howard Street Tunnel under the streets of downtown Baltimore, MD (see www.usfa.dhs.gov/downloads/pdf/publications/tr-140.pdf). The train was carrying a variety of freight and hazardous materials, with three locomotives pulling 60 cars. Complicating the scenario was the subsequent rupture of a 40-inch water main that ran directly above the tunnel. The flooding hampered extinguishing efforts, caused several city streets to collapse, knocked out electricity to approximately 1,200 customers, and flooded nearby buildings. The derailment also interrupted a major communications line associated with the Internet and an MCI fiber-optic telephone cable.

During the two-day response, five alarms were requested with 17 engines, eight trucks, and three battalions, in addition to the HazMat, EMS, and rescue teams; 150 firefighters were on the scene, working to extinguish the fire. The fire-extinguishing operations were performed from both ends of the tunnel as well as through manholes located at Howard and Lombard Streets. The city of Baltimore

activated the civil defense sirens at 5:45 P.M. to warn citizens of impending danger from the fire and hazardous materials. On the night of the derailment, city officials closed entrances to the city from all major highways and cancelled major public events.

NIMS suggests that the Incident Command Post (ICP) perform an EOC-like function in small-scale incidents [4]. In this response, the on-site management teams assumed the overall management support and supervision role (the many-second coordination cycle) and the individual divisions responded at the operational level (the mini-second coordination cycle). Table 3 shows the application of our framework to this real incident.

## CONCLUSION

Coordination in the context of emergency response is an understudied research issue. It is an important problem, as it impacts life and property in the affected area. We have proposed a framework to analyze coordination patterns along the emergency response life cycle. This framework may be further utilized by researchers and practitioners to: depict emergency coordination practices along focal dimensions elaborated in the framework; understand the overarching requirements for coordination design and implementation; and identify coordination ineffectiveness and analyze the alternatives for optimal solutions. This article has also applied the framework to a real-life emergency incident as a proof of concept of its relevance and usability. The case application demonstrates not only the applicability of the framework during disasters but also serves as a reminder template of the number of things to consider while countering emergencies and disasters.

It is important to point out that a number of new technologies have emerged in recent years to enable better emergency response coordination. Example solutions include wireless mesh networks (CalMesh; calmesh.calit2.net), sensor networks (ASPECT; www.epa.gov/naturalevents/flyinglab.htm), knowledge management systems (RKBP; www.rkb.mipt.org), geographic information systems (CATS; cats.saic.com), communication standards (CAP; www.incident.com/cap), incident forecast and analysis programs (SLOSH; www.fema.gov/plan/prevent/nhp/ slosh_link.shtm), peer-to-peer communication platforms (Microsoft Groove; www.groove.net), collaborative work systems (E-Team; www.eteam.com), and command and control systems (DisasterLAN; www.disasterlan.com). These technology elements address parts of the puzzle and have to be leveraged to improve coordination. However, the discussion of these technologies in the context of emergency response management and coordination systems is beyond the scope of this article will be taken up in future research. C

## REFERENCES


1. Backhouse, J. and Dhillon, G. Structures of responsibility and security of information systems. *European Journal of Information Systems 5*, 1 (1996), 2–9.
2. Comfort, L.K., Ko, K. and Zagorecki, A. Coordination in rapidly evolving disaster response systems. *American Behavioral Scientist 48*, 3 (2004), 295–313.
3. Dawes, S.S., Birkland, T., Tayi, G.K. and Schneider, C.A. *Information, Technology, and Coordination: Lessons from the World Trade Center Response*. Center for Technology in Government, 2004.
4. U.S. Department of Homeland Security. *National Incident Management System*, 2004.
5. Faraj, S. and Xiao, Y. Coordination in fast-response organizations. *Management Science 52*, 8 (2006), 1155–1189.
6. Howitt, A.M. and Leonard, H.B. A command system for all agencies. *Crisis/Response Journal 1*, 2 (Jan./Mar. 2006), 40–42.
7. Lewandowski, S.M. et al. SARA: Survivable Autonomic Response Architecture. In *Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX II'01)* (2001).
8. Petrescu-Prahova, M. and Butts, C.T. *Emergent Coordination in the World Trade Center Disaster*. Institute for Mathematical Behavioral Sciences, 2005.
9. Raghu, T.S., Jayaraman, B. and Rao, H.R. Toward an integration of agent- and activity-centric approaches in organizational process modeling: Incorporating incentive mechanisms. *Information Systems Research 15*, 4 (2004).
10. Rao, H.R., Chaudhury, A. and Chakka, M. Modeling team processes: Issues and a specific example. *Information Systems Research 16*, 3 (1995), 255–285.
11. Shen, S.Y. and Shaw, M.J. Using information technology for effective emergency response. In *Proceedings of the International Business Research Forum 2005*, (Philadelphia, PA, 2005).
12. Yuan, Y. and Detlor, B. Intelligent mobile crisis response systems. *Commun. ACM 48*, 2 (Feb. 2005), 95–98.



**RUI CHEN** (ruichen@buffalo.edu) is a Ph.D. candidate of Management Science and Systems at the State University of New York at Buffalo.

**RAJ SHARMAN** (rsharman@buffalo.edu) is an assistant professor of Management Science and Systems at the State University of New York at Buffalo.

**H. RAGHAV RAO** (mgmtrao@buffalo.edu) is a professor of Management Science and Systems and adjunct professor of Computer Science and Engineering at the State University of New York at Buffalo.

**SHAMBHU J. UPADHYAYA** (shambhu@cse.buffalo.edu) is an associate professor of Computer Science and Engineering at the State University of New York at Buffalo.



This research has been funded by NSF under grant 0705292. Many emergency response coordinators from FEMA and State of New York (especially Western New York) provided valuable insights toward the development of this article (for details see www.som.buffalo.edu/isinterface/ack.html).





DOI: 10.1145/1342327.1342340

# ICL2008

## Conference chair

**M. E. Auer**
(CTI Villach, Austria)

## Program committee

### Senior Members

**A. Y. Al-Zoubi,**
PSUT Amman, Jordan,
**Peter Baumgartner,**
Danube University Krems, Austria
**Christos. Bouras**
University and CTI Patras, Greece
**Rhena Delport**, University of Pretoria, South Africa
**Christian Dorninger,**
BMBWK Vienna, Austria
**Arthur Edwards**, Universidad de Colima, Mexico
**David Guralnick,** Kaleidoscope Learning New York, USA
**Martin Hitz,**
University of Klagenfurt, Austria
**Josef Hvorecky,**
VSM Bratislava, Slovakia
**Trayan Iliev,** University of Sofia, Bulgaria
**George Ioannidis,**
University of Patras, Greece
**Göran Karlsson,**
KTH Stockholm, Sweden
**Barbara Kerr,** Concordia University Montreal, Canada
**Andreas Pester,** Carinthia Tech Institute Villach, Austria
**Serge Ravet**, European Institute for E-Learning, France
**Rob Reilly,** MIT Media Lab, USA
**Cornel Samoila**, University Transylvania Brasov, Romania
**Wolfgang Scharl,**
Technikum Vienna, Austria
**Jeanne Schreurs**, University of Hasselt, Belgium
**Thomas Schmidt,**
HAW Hamburg, Germany
**Richard Straub,**
IBM, France and ELIG
**Andras Szucs,**
EDEN, TU Budapest, Hungary
**Linmi Tao**, Tsinghua University Peking, China
**Doru Ursutiu,** University Transylvania Brasov, Romania

## More information

http://www.icl-conference.org
info@icl-conference.org
phone: +43-4242-90500-2115

## Registration

http://www.conftool.com/icl-conference/

# FIRST CALL FOR PAPERS

International Conference
Villach/Austria
Interactive Computer Aided Learning
ICL

## 11th International Conference on Interactive Computer aided Learning

## www.icl-conference.org

### September 24 – 26, 2008, Villach, Austria

This interdisciplinary conference aims to focus on the exchange of relevant trends and research results as well as the presentation of practical experiences in interactive computer aided learning. Therefore pilot projects, applications and products will also be welcome.

This conference will be organized by the **Carinthia University of Applied Sciences, Villach/Austria** in cooperation with:

- Federal Ministries for Science, Education and Culture of Austria,
- IEEE Education Society,
- European Distance and E-learning Network (EDEN)
- International E-Learning Association (IELA)
- International Society of Engineering Education (IGIP)
- International Association of Online Engineering (IAOE)
- European Inst. of E-Learning (EifEL)
- Austrian Computer Society (OCG)
- IT Campus Carinthia

## Topics of Interest

**ICL2008 will have a focus on Educational MashUps, Collaborative Learning Environments as well as E-Portfolios.**

- Web based learning (WBL)
- Life long learning
- Intelligent content
- Adaptive and intuitive environments
- Responsive environments
- Mobile learning environments and applications
- Computer aided language learning (CALL)
- Platforms and authoring tools
- Educational MashUps
- Networks/Grids for learning
- Knowledge management and learning
- Educational Virtual Environments
- Collaborative learning
- Applications of the Semantic Web
- E-Portfolios
- Standards and style-guides
- Remote and virtual laboratories
- Multimedia applications and virtual reality
- Pedagogical and psychological issues
- Evaluation and outcomes assessment
- New learning models and applications
- Cost-effectiveness
- Real world experiences
- Pilot projects / Products / Applications
- … and others

## Types of Contributions

- **Full Papers** (20 minutes presentation followed by a panel discussion)
- **Short Papers** (15 minutes presentation)
- **Interactive Demonstrations** (15 minutes, also on-line demonstrations)
- **Posters**

All submissions are subject to a double blind reviewing process.

Other opportunities to participate:

- **Run a workshop or tutorial.**
- **Organize a thematic session.** Proposals should include a minimum of three papers.
- **Exhibit** at ICL (products and developments of e-learning technology)

## Important Dates

| | |
|---|---|
| **May 16, 2008** | **Submission of extended abstracts** (two pages) |
| **June 20, 2008** | **Notification of acceptance** |
| **Sept. 10, 2008** | **Camera-ready due** |
| **Sept. 24-26, 2008** | **Conference ICL2008** |

## Proceedings

All accepted submissions will be published in the conference proceedings (ISBN 978-3-89958- 353-3).

# TAMING *Heterogeneous Agent* ARCHITECTURES

*Using aspect-oriented techniques to construct high-quality multi-agent systems.*

BY ALESSANDRO GARCIA AND CARLOS LUCENA

The recent advances in network-based software applications and the advent of ubiquitous computing are pushing us inevitably toward a world of autonomous software architectures. This trend has spurred the revitalization of agent technology as a complement to the object paradigm for a variety of modern application domains, including e-commerce, software development environments, and personal assistants. There is explicit evidence indicating the penetration of software agents is also high for systems used in military and government contexts [2, 8]. Agents, like objects, provide services to their clients, but are recognizably different from objects as seen from an architectural point of view [3, 5, 8]. Unlike objects, an agent is an autonomous entity that takes the initiative to achieve system goals and represent software users [2, 3].

As a result, architects of software agents are faced with basic concerns, such as the agent services that are made available to the clients, and a number of additional concerns on top of the basic concerns. The internal architecture of a single agent encompasses multiple properties, including autonomy, interaction, adaptation, collaboration, learning, and mobility. Hence, architects of agent-based systems are also concerned with issues such as making an agent interact appropriately, handling the agent's adaptive behavior, structuring the agent's autonomous behavior, designing the agent roles and protocols for inter-agent collaboration purposes, and incorporating learning mechanisms into the agent's structure in a modular manner.

Not surprisingly, separation of concerns is at the core of the development of agent-based software systems [5, 12]. The reuse and maintenance of agent elements depend largely on the ability of used architectural abstractions to support the separate handling of agent-specific concerns since an early state of design. The applied architectural styles must enable the modularization of each agent concern and their proper composition, so that the achieved segregation significantly limits the impact of a change and improves the chances for architecture reuse in other software projects. This separation of concerns needs to be guaranteed throughout the different development phases, especially from the architectural to the implementation phase. The architectural separation will be lost if the implementation abstractions are not able to preserve it. In fact, the sole use of existing well-known agent platforms, such as JADE and JACK [12], do not provide advanced mechanisms to achieve separation of concerns, which are restricted to exploit the object-oriented composition and decomposition mechanisms [5]. On the other hand, even though such advanced implementation mechanisms are delivered, the benefits of separation of agent concerns would be hindered if existing architectural abstractions do not allow the achievement of proper system modularization from the design outset.

## HETEROGENEITY IN AGENT ARCHITECTURES

Although separation of concerns is critical to architects of agent-based software, it is often difficult to achieve in realistic systems for several reasons. First, these systems typically encompass heterogeneous types of agents [2]. The internal architecture of distinct agent types differs widely from each other since they incorporate distinct properties [8]. Second, each property is orthogonal and interacts with the agent's basic functionality and often with other agent properties. These properties typically crosscut several modules of an agent architecture, independently from the adopted internal model, such as the constraint-oriented model or the BDI (Belief-Desire-Intention) model [10], and from the agent's cognitive level, such as reactive agents, deliberative agents, or even hybrid agents. Third, these crosscutting agent-specific concerns are related in dramatically different ways that depend on the agents' types and the application requirements [6]. For example, in a given agent-based application, the mobility behavior of an agent may only directly affect the basic functionality of the agent, while it may also crosscut the collaboration and interaction concerns in a second application.

As a consequence, agent-based applications require an architectural approach that is flexible enough to support adjustable composition of agent concerns and the construction of heterogeneous agent architectures according to the application demands. This flexibility requirement is even more stringent in open agent-based systems due to their adaptive and open nature. In these contexts, the agent architecture needs to be modular enough to support the dynamic reconfiguration of its internal elements. For example, roles potentially need to be changed as the agent moves to new environments. The degree of autonomy and the learning strategies may also need to be adapted or disabled according to the dynamic execution contexts. While part of the system-level reconfiguration facilities are typically provided by specific middleware implementations, many of the agent concerns and their composition are essentially application-dependent. Software architects need to prepare and conceive at the design

Agent-based applications require *an architectural approach that is flexible enough to support adjustable composition of agent concerns* and the construction of heterogeneous agent architectures according to the application demands.

outset modular agent architectures in order to cope with these heterogeneity issues.

## SHORTCOMINGS OF CLASSICAL AGENT ARCHITECTURES

Agent-oriented software engineering has been studied from different perspectives, including agent-oriented methodologies and languages for higher-level development phases, conceptual modeling, and implementation frameworks [2, 3]. Although separation of concerns is widely recognized as crucial to the development of maintainable multi-agent software [5, 6, 12], existing approaches do not scale up to support the separation of agent properties in heterogeneous architectures. Developers must rely on traditional architectural patterns, such as the Mediator pattern [4] and the Layers pattern [4], in order to build their systems. As illustrated in Figure 1, these solutions define architectural abstractions, such as mediators and layers, and composition rules to support the isolation of agent concerns and their further composition.

However, these solutions impose rigid connections on the architectural components, which make the construction of heterogeneous agent types difficult and not scalable to cope with the complexity of multiple interactive agent concerns. These existing architectural proposals are applicable to simple agent architectures and with few agenthood properties. The constraints on the composition rules imposed by those solutions do not scale up to master the intrinsic crosscutting character of agent properties in heterogeneous contexts. Figure 1 illustrates this problem in terms of a layered agent architecture (Figure 1a) and a mediator-based agent architecture (Figure 1b), represented with a simplified UML notation.

The layered architecture imposes a bidirectional communication only between adjacent layers; that is, the architectural components of a software agent. In order to make proper internal decisions and adapt the agent knowledge accordingly, the autonomy and adaptation components must be aware of received messages and external stimulus coming from the surrounding environment. As illustrated in Figure 1a, this requirement forces the software architects to add interaction-related interfaces into the other layers so that the autonomy and adaptation components have access to the perceived stimulus and messages. As a result, the interaction concern (represented in blue) crosscuts the modularity of the kernel and adaptation components.

The situation is not different in the mediator-based solution. This architectural style requires all the inter-component communications being intermediated by a central component, which is also in charge of encapsulating the agent's basic functionality. As illustrated in Figure 1b, this architectural approach also fails to isolate the interaction concern and leads to the intermingling of the agent kernel functionality and the message-processing functionality. Note that a similar problem occurs with the mobility concern as the communication with the autonomy component needs to be settled by the agent kernel. Hence, the inability of traditional architectural styles and respective composition rules in capturing the crosscutting feature of agent concerns in heterogeneous architectures leads to several negative consequences, such as architectural tangling and bloated interfaces, which are described in detail in the table here.

**Shortcomings of existing architectural styles.**

| Architectural Problem | Description |
|---|---|
| Architectural scattering | The occurrence of architectural elements, such as interfaces, that belong to one agent concern in architectural components encapsulating other agent concerns. For example, the interaction-related interfaces are scattered over the agent architectures in Figure 1. |
| Architectural tangling | The mix of multiple agent concerns together in the same module. For instance, tangling is evident in the agent kernel of Figures 1a and 1b as both are implementing interfaces associated with different agent concerns. |
| Hindering of modular reasoning | The inability of the architect being able to reason and make decisions about an agent architecture's module while looking only at its functionality and its interfaces. For example, the architects treating the collaboration and interaction concerns in Figure 1a need to consult the definitions and interfaces of other modules, leading to an expanded or global reasoning rather than a modular reasoning. |
| High architectural coupling and low cohesion | The presence of an overly strong connection between architecture components (high coupling), and the lack of closeness between the internal functionalities and interfaces of a specific component (low cohesion). In Figure 1b, for example, the system coupling is increased, and the cohesion of each component (for example, the Kernel) is decreased. |
| Architectural interface bloat | The evidence of complexity increase in the component interfaces. For example, the number of interfaces of several components in Figures 1a and 1b is augmented due to the crosscutting nature of the interaction concern. |
| Poor traceability from requirements to architecture | The inability of directly mapping in the architecture description the separation of agent concerns achieved in the requirements description. For example, the mobility-specific requirement cannot be directly traced to the Mobility component (Figure 1b) as it also affects the boundaries of the Kernel component. |
| Poor modularization of design decisions and rationale | The incapacity of the architectural decisions and rationale associated with one agent concern being documented in a localized manner. The rationale and design decisions related to the interaction concern are typically tangled and scattered over the architecture artifacts in the same way as interaction-specific interfaces and functionalities. |
| Inflexible design | The rigidness in the composition rules to support alternative compositions between agent components in order to smoothly produce heterogeneous architectures. This inflexibility is visible both in the mediator and layered solutions. |
| Limited evolvability and maintainability | The adversities associated with including, changing, or removing the design elements associated with an agent crosscutting concern as they crosscut several architectural components. For instance, this problem would occur recurrently in changing or replacing the collaboration and interaction components in Figure 1a. In fact, in a previous study [6] we have observed a poor stability of a typical mediator-based agent architecture implementation in the presence of maintenance and evolution scenarios. |

## SPREADING THE CROSSCUTTING TO SUBSEQUENT ARTIFACTS

The problems associated with heterogeneous agent architectures are typically disseminated from the agent architecture specifications to artifacts generated later in the software life cycle. Hence, the crosscutting of agent concerns is inevitably spanned over the resulting artifacts in the detailed design level and implementation. It does not matter what kind of decomposition and abstractions the agent-based software developers are relying on; the problem is the same if you are using an agent-oriented design language [2], or an OO modeling language, such as UML.

For example, the architectural abstractions and composition rules are not directly supported in OO design and programming languages. They are not aligned with the composition and decomposition mechanisms of the object paradigm, which makes it difficult to handle the heterogeneous agent types in a single system [6]. For instance, the inheritance mechanism usually leads to large inheritance trees with replication of code and an explosion of the number of classes [5, 6, 12]. Even in some simple agent architectures, bringing them into the detailed design and code also raises similar problems (examples are shown in the table), such as scattering and tangling.

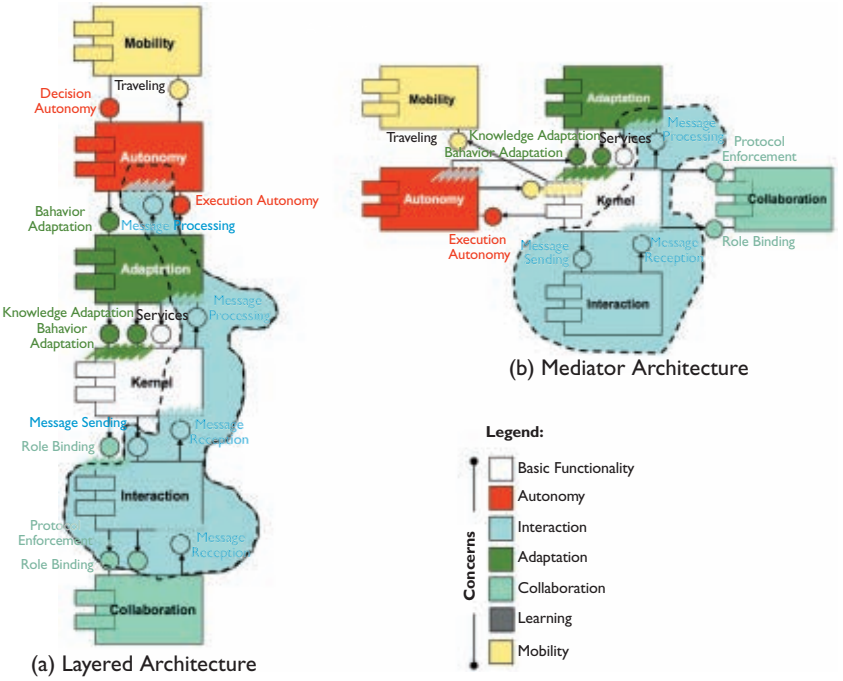


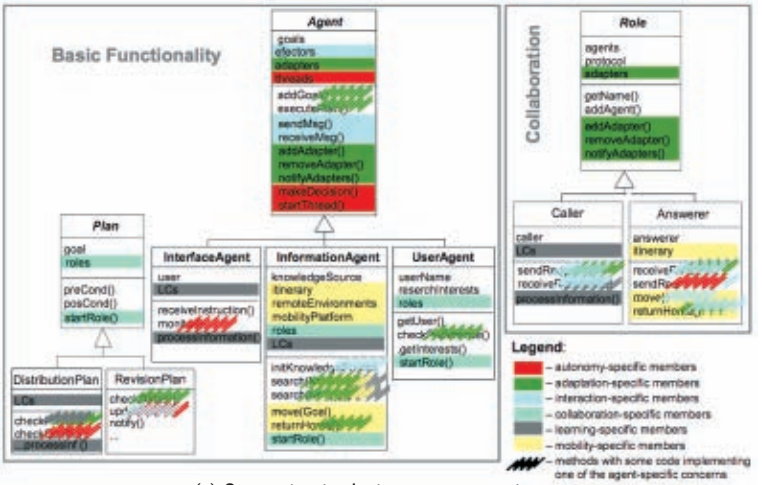

**Figure 1. Crosscutting concerns in heterogeneous agent architectures.**

Figure 2a illustrates how the reification of heterogeneous agent architectures in the design tends to be scattered over many classes of the system design. It shows a partial representation of an agent-based system [6] composed of three agent types, each with a different internal architecture. Each set of classes, surrounded by a gray rectangle, has the main purpose of modularizing a specific agent concern, namely the agents' basic functionality and collaboration. The figure shows that agent-specific concerns crosscut the classes implementing those surrounded classes, such as "Information Agent" and "Role." The use of OO mechanisms breaks not only the class-level modularity, but also the operation-level modularity. Figure 2b shows the partial implementation of a searcher agent, which consists of a class and respective operations with a confusing tangle of lines of code for different concerns. On the left, there is the code for a class as a non-agent entity. On the right, there is an equivalent class with code for implementing agent-specific concerns. After analyzing the code on the right, it is clear that such code has lost the functional encapsulation of the non-agent version (left side). Moreover, it is a confusing intermingling of lines of code for different agent concerns, and the modularization of several agent-specific concerns is lost across the system classes.

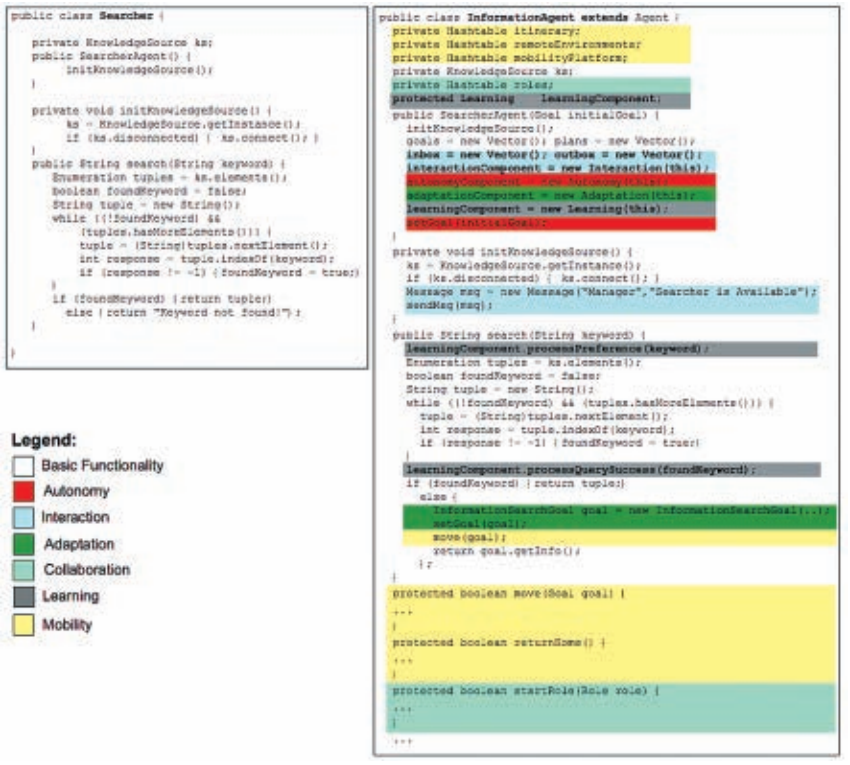
## ARCHITECTING SOFTWARE AGENTS WITH ASPECTS

Aspect-oriented software development [9] is an evolving paradigm to modularize concerns that existing software engineering abstractions and mechanisms are not able to capture explicitly. The notion of aspects encourages modular descriptions of complex software by providing support through new composition mechanisms for cleanly separating the system functionality from its crosscutting concerns. However, existing aspect-oriented approaches have not been explored in the context of heterogeneous agent architectures; they have focused on the context of classical crosscutting concerns, such as distribution [11], persistence [11], and design patterns [7]. These concerns have been mostly studied in an isolated way and from the implementation point of view.

Aspects and its new composition possibilities can be exploited at the architectural level to capture the multiple interacting agent concerns that are difficult to




(a) Scattering in design representation.




(b) Tangling in a class representing an agent.

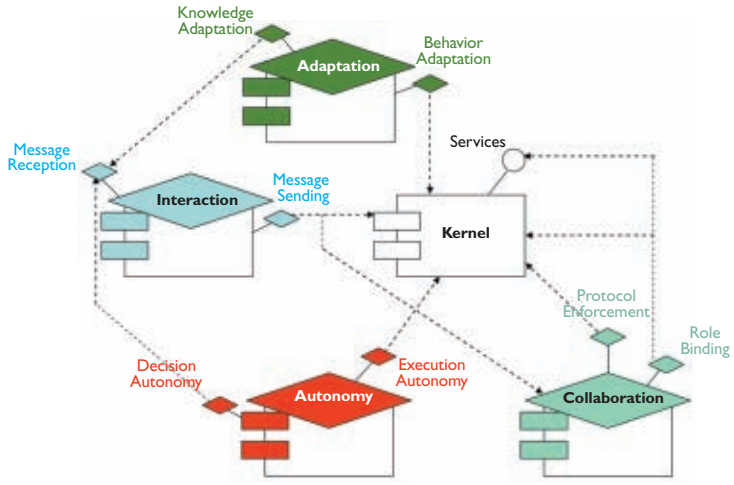
Figure 2. Heterogeneous agent architectures: Crosscutting the life-cycle artifacts.

modularize with existing architectural abstractions. An aspect-oriented architectural style brings to the software architects a new abstraction, the notion of architectural aspects, and new composition means for handling each crosscutting agent concern as an individual component at an early stage of design, as illustrated in Figure 3a. Each architectural aspect modularizes a typical crosscutting agent property and separates it from the agent kernel. The "aspectization" of agent architectures allows the association of agenthood properties with the basic functionality in a way that is transparent to the agent kernel.

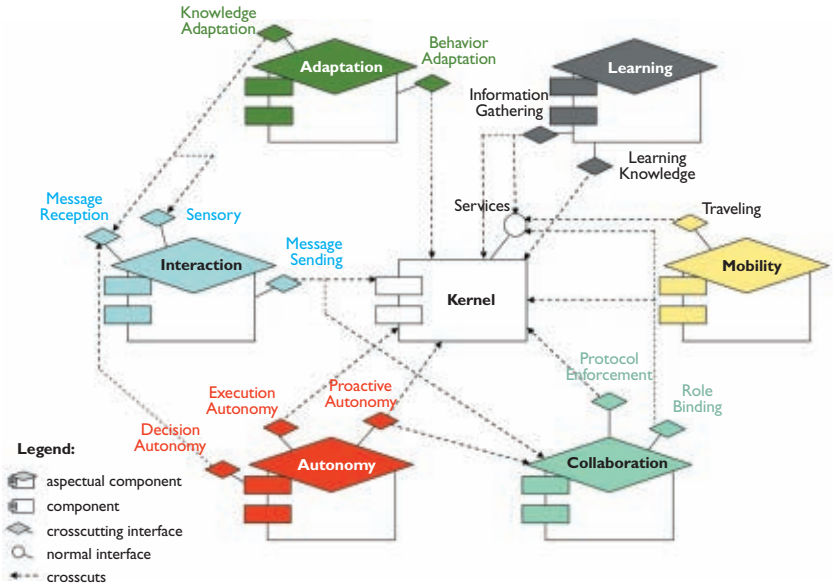
The key idea to enable adjustable compositions is the notion of crosscutting interfaces, which are modularity abstractions attached to the architectural aspects. A crosscutting interface is different from a module interface in the sense that the latter essentially provides services to other components. Crosscutting interfaces provide services to the system, but also specify when and how an aspect affects other architectural components. Contrary to interfaces in traditional architecture styles, they flexibly determine which external components and interfaces the architectural aspect of a software agent will be connected. With this dependency inversion, crosscutting interfaces overcome the problems associated with the rigidness implicit in traditional architectural styles (see Figure 1). Each agent's architectural aspect can be more flexibly composed with the agent kernel and with any agent aspects depending on the requirements of a specific agent architecture.

Each of the architectural aspects is related to more than one component, representing the crosscutting nature of agent properties in complex architectures. An agent's architectural aspect can realize more than one crosscutting interface since it can crosscut multiple agent components in different ways. The interface of an architectural aspect can crosscut the Kernel component and other architectural aspects. An aspect interface crosscuts either the internals of an agent component or elements of other interfaces. The first case means the architectural aspect affects the internal structure or dynamic behavior of an agent component. The second case means the aspect affects elements of an interface.

Instances of aspect-oriented agent architectures are illustrated for a user agent (see Figure 3a) and an information agent (see Figure 3b). Extensions of UML are used to represent the architectural aspects (compo-




(a) Aspect-oriented architecture of a user agent.




(b) Aspect-oriented architecture of an information agent.

Figure 3. Aspectizing heterogeneous agent architectures: Enabling flexible composition.

agent roles (collaboration component) in order to make decisions.

Figure 3 shows that the aspect-oriented style easily supports both architectural configurations. A partial architectural design is shown and some non-aspectual components, such as the ones for agent knowledge representation and message assembling, are left out for simplicity purposes. In the first case (Figure 3a), the autonomy aspect does not have an interface for enabling proactive behavior. It has two crosscutting interfaces that affect the kernel and interaction components. In the second case (Figure 3b), the autonomy component includes a third crosscutting interface that supports the observation of the agent kernel and the collaboration component in order to trigger the proactive behavior of the information agent. The architectures have other important differences in their configurations and compositions. For example, the architecture of the information agent also includes the mobility and learning aspects and a third interface to enable sensory behavior. Moreover, the adaptation aspect crosscuts the sensory interface in order to adapt as new environmentally relevant events are sensed by the agent.

nents with diamonds on their top), crosscutting relationships (dotted arrow), and crosscutting interfaces (small diamonds). Their partial non-aspect-oriented detailed designs were represented in Figure 2a. The user agent is prominently reactive, while the information agent is a proactive entity. As a result, their architectures are very different. For example, the component in charge of implementing the autonomous behavior has different interfaces and relationships with each component in the agent architecture. Reactive agents make decisions only as a matter of deciding about external requests embedded in incoming messages. Proactive agents also need to inspect changes in their internal state and decide for starting new actions by their own initiative independently from requests from other agents and system users. In other words, the autonomy component also needs to observe events in the agent kernel and in the

## WHAT ARE OTHER BENEFITS?

The benefits of having new architectural design rules entailed by an aspect-oriented style are not limited to improved composability of agent architecture concerns. We have noticed a number of other positive consequences such as the ones discussed here.

**Modular Reasoning at the Architectural Level.** The aspectization of agent architectures supports modular reasoning, since the architects are able to more independently treat each agenthood property. Different from other architectural styles, the agent concerns are modularized and do not affect the definitions and interfaces of multiple architectural modules. The notion of crosscutting interfaces allows addressing tangling- and scattering-related problems typically found in the definition of heterogeneous agent architectures. As a result, software architects are able to

make decisions about an agent property while only looking at its description and its interface with other concerns.

**Smooth Transition in Software Life-Cycle Phases.** Aspect-oriented agent architectures are directly mapped to implementation abstractions using well-known aspect-oriented programming languages, such as AspectJ [1]. AspectJ, which extends the Java programming language, is the most popular aspect language. Architectural aspects are decomposed into a set of AspectJ aspects and classes. The crosscutting interfaces are realized as pointcuts, advice, and inter-type declarations, because they define different ways an aspect affects other design and implementation modules. Join points are well-defined points in the dynamic execution of the system components. Examples of join points are method calls and method executions. Pointcuts have a name and are collections of join points. Advice is a special method-like construct attached to pointcuts. Inter-type declarations introduce attributes, methods, and interface implementation declarations into the components to which the crosscutting interface is attached. We have implemented an AspectJ framework that supports this aspect-oriented architectural style at the implementation level [5]. It helps to guarantee a smooth transition from the specification of heterogeneous agent architectures to their detailed design and implementation. We have also developed several other aspect-oriented techniques to cope with agent aspects in different software development phases in order to facilitate the traceability of the software engineering artifacts.

**Coping with Dynamic Adaptability and Customizability.** Architecting software agents with improved separation of concerns is of paramount importance in open, dynamic agent-based systems. Dynamic reconfiguration of agent roles and collaboration protocols are often required as the agents move to different environments. In addition, with the growing number of applications for pervasive computing, the selection of learning and coordination strategies may depend on the context in which the agent is being executed. As such, agent architectures need to be designed properly, and the dependency inversion gained with aspect-oriented agent architectures is a key factor to allow the agent adaptation and customization.

## CONCLUSION

Numerous types of agent architectures are prevalent [8] and must be developed in a way that meets the stringent modern requirements of evolvability, reusability, and dynamic reconfigurability. Existing architectural styles are rigid by their very nature and unable to cope with the crosscutting nature of agent properties as well as the complexity of heterogeneous agent architectures, which are often required in realistic modern systems. On the other hand, aspect-oriented software development is gaining wide attention both in research environments and in industry as a paradigm to promote improved modularity of complex software systems. The exploration of aspect-oriented techniques clearly seems a promising step forward to allow the construction of more flexible agent architectures and to foster enhanced quality of realistic multi-agent systems. C

## REFERENCES


1. AspectJ Web site. *The AspectJ Guide*; eclipse.org/aspectj/.
2. Bergenti, F., Gleizes, M.-P., and Zambonelli, F., Eds. *Methodologies and Software Engineering for Agent Systems: The Agent-Oriented Software Engineering Handbook, Volume 11*. Springer-Verlag, 2004.
3. Bordini, R., Dastani, M., Dix, J., and Seghrouchni, A., Eds. *Multi-Agent Programming: Languages, Platforms, and Applications*. Springer, NY, 2005.
4. Buschmann, F. et al. *Pattern-Oriented Software Architecture: A System of Patterns*. Wiley, NY, 1996.
5. Garcia, A., Lucena, C., and Cowan, D. Agents in object-oriented software engineering. *Software: Practice and Experience 34*, 5 (Apr. 2004), 489–521.
6. Garcia, A. et al. Separation of concerns in multi-agent systems: An empirical study. In C. Lucena et al., Eds., *Software Engineering for Multi-Agent Systems II*, Springer-Verlag, LNCS 2940 (Mar. 2004), 49–72.
7. Garcia, A. et al. Modularizing design patterns with aspects: A quantitative study. In *Proceedings of the 4th ACM International Conference on Aspect-Oriented Software Development (AOSD'05)*, 1–11.
8. Jennings, N. An agent-based approach for building complex software systems. *Commun. ACM 44*, 4 (Apr. 2001), 35–41.
9. Kiczales, G. et al. Aspect-oriented programming. In *Proceedings of the European Conference on Object-Oriented Programming-ECOOP'97*, LNCS 1241, 220–242.
10. Rao, A. and Georgeff, M. BDI agents: From theory to practice. In V. Lesser and L. Gasser, Eds., *Proceedings of the 1st International Conference on Multi-Agent Systems (ICMAS-95)*, 312–319.
11. Soares, S, Laureano, E., and Borba, P. Implementing distribution and persistence aspects with AspectJ. In *Proceedings of the ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA'02)*, 174–190.
12. Ubayashi, N. and Tamai, T. Separation of concerns in mobile agent applications. In *Proceedings of the 3rd International Conference Reflection 2001*, LNCS 2192, 89–109.



**ALESSANDRO GARCIA** (garciaa@comp.lancs.ac.uk) is a lecturer of computer science at Lancaster University, U.K.

**CARLOS LUCENA** (lucena@inf.puc-rio.br) is a professor of computer science at Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil.





DOI: 10.1145/1342327.1342341

# Call for Nominations

## The ACM Doctoral Dissertation Competition

### RULES OF THE COMPETITION

ACM established the Doctoral Dissertation Award program to recognize and encourage superior research and writing by doctoral candidates in computer science and engineering. These awards are presented annually at the ACM Awards Banquet.

**SUBMISSIONS:** Nominations are limited to one per university or college, from any country, unless more than 10 Ph.D.'s are granted in one year, in which case two may be nominated.

**DEADLINE:** Submissions must be received at ACM headquarters by September 30, 2008 to qualify for consideration. This date is not negotiable.

**ELIGIBILITY:** Each nominated dissertation must have been accepted by the department within a 12-month period prior to September 30, 2008. Exceptional dissertations completed in September 2007, but too late for submission last year will be considered. Only English language versions will be accepted.
Please send a copy of the thesis in PDF format to emily.eng@acm.org. If a PDF is not possible, send a copy to the address listed here.

**SPONSORSHIP:** Each nomination shall be forwarded by the thesis advisor and must include the endorsement of the department head. A one-page summary of the significance of the dissertation written by the advisor must accompany the transmittal.

**PUBLICATION RIGHTS:** Each nomination must be accompanied by an assignment to ACM by the author of exclusive publication rights.

**PUBLICATION:** Winning dissertations will be published by Springer.

**SELECTION PROCEDURE:** Each nominated dissertation will be screened by three to six specialists in the relevant fields to provide a preliminary evaluation.

On the basis of these evaluations, five finalists will be chosen for submission to a selection committee of six persons who serve staggered terms of three years.

The selection committee will select the winning dissertation by December 15.

**AWARD:** The Doctoral Dissertation Award is accompanied by a prize of $20,000 and the Honorable Mention Award is accompanied by a prize of $10,000. Financial sponsorship of the award is provided by Google.

Doctoral Dissertation Coordinator
ACM
2 Penn Plaza, Suite 701
New York, NY 10121

TECHNOLOGY-FACILITATED

# 'GIVE ACCORDING TO YOUR ABILITIES, RECEIVE ACCORDING TO YOUR NEEDS'

*Software, hardware, computer networks, and online content all require mechanisms to prevent free riding, overuse of common resources, use of resources without consent, and blocking of resource access.*

Information technology, especially the Internet, facilitates the production, distribution, and consumption of products and services by increasingly following the principle of "Give according to your abilities, receive according to your needs," or GARN. Users contribute nonmonetary resources (such as programming skills, computing power, and network access) to a resource pool and draw similar resources from it. All this is done in return for no monetary reward, and giving or receiving can be at a level of zero; that is, users receive without giving or give without receiving. The GARN phenomenon is manifest in at least four contexts:

*Software.* Volunteer-based open-source software (OSS) initiatives (such as Linux) where code is created by volunteers and made available to the public;

*Hardware.* Organizational, interorganizational, and organizational/private grid computing initiatives (such as SETI@home, setiathome.berkeley.edu);

*Networks.* Wireless access points for wireless devices (primarily in urban areas); and

*Content.* Asset-based content contribution (such as P2P file-sharing systems) and knowledge-based content contribution (such as Wikipedia).

By Oded Nov and Bharat Rao

## Like other nonexcludable goods, content lends itself to FREE RIDING.

Here, we compare these contexts using the GARN prism, identify the threats associated with the GARN phenomenon, and discuss ways to counter the threats. We provide a unified framework to phenomena that were previously viewed in isolation. Our aim is to unravel the common thread among Internet-facilitated contexts and consequently examine how lessons from one context may be used in others. Before exploring specific GARN phenomena, we turn to the economics literature and present several types of goods, as they are useful in our analysis.

| | | Nonrival | |
|---|---|---|---|
| | | Low | High |
| Nonexcludable | High | Commons good | Public good |
| | Low | Private good | Collective good |

Table 1. Types of goods [3].

Goods can be classified into four categories across the dimensions of nonexcludability and nonrivalry (see Table 1) [3]. For example, it is difficult to prevent nonexcludable goods (such as public parks) from being consumed by nonpayers; a nonrival (or nonconsumable) good (such as a TV show) can be consumed without diminishing its availability for future consumption. Unlike physical goods, information goods (such as software and content) are typically nonrival [11].

OSS development is a growing phenomenon [12], evident in the number of users registered at Sourceforge.net (a host for open-source projects, www.sourceforge.net) that increased from around 500,000 in 2003 [6] to more than 1.8 million in 2008. A growing variety of software applications are developed in the open-source model; more than 170,000 are registered at Sourceforge. Contributions are made by individuals independently deciding to donate time and effort to produce software that is freely available to all (a pure GARN) or by software companies actively supporting OSS development.

The availability of high-speed networks allows the sharing of computer resources via grid computing, enabling better allocation and use of computing resources within and among organizations. Grids may be closed (to members of a certain community) or open (accessible to all); they may also be one-way (contributors lack access to the pooled resource) or two-way (contributors are also receivers). A well-known example of a grid-computing project is SETI@home, a volunteer-based grid comparable to the strongest computer available today, for searching for "extraterrestrial intelligence." Computing grids can be viewed as grids of computers and primarily represent an architectural issue. However, our focus here is the wider issue of grids of computer users and owners, including their social and organizational implications.

Increasing use of wireless broadband computing is accompanied by the emergence of wireless commons, whereby participants provide Internet access to other participants through their WLAN access points [2]. Like computing grids, wireless commons may be either closed or open to all [2]. If the latter, which represents a pure GARN mechanism, individuals must not password-protect their access points.

Content GARN often manifests itself as contribution of content produced by others (such as media files shared via P2P systems). However, in many P2P networks, contributing and consuming are technically almost inseparable for users. Therefore, we won't analyze P2P networks here further. A clearer example of content GARN is knowledge-based, self-created content, such as Wikipedia, a free online encyclopedia with 1.4 million entries in English alone, written and edited by volunteers [7].

The GARN mechanism carries with it a number of threats that can be categorized using the goods typology (see Table 2); these are analyzed in the next sections. Given the common GARN mechanism

underlying the four contexts, each threat is followed by a discussion of the possible countermeasures, along with countermeasures used in other contexts that may be relevant.

**Threat 1. Free riding, or asymmetry between giving and receiving.** Asymmetry manifests itself as a lack of contributed resources for maintaining and improving the common pool of resources [2].

*Contexts. OSS.* OSS lends itself to free riding, as no contribution is requested in return for the use of freely available resources [12]. Countering the threat involves motivating potential contributors. This can be done by highlighting contributor benefits: learning from the expertise of peers in a technical community [6, 12]; being able to tailor software to one's personal needs [12]; having fun [9]; having the opportunity to signal status [5], potentially enhancing employment prospects [6, 10]; having the opportunity to support self identification [4]; and having the opportunity to help others.

*Computing grids.* In two-way grids, asymmetry of resource contributions can be countered through agreements among users to align receipt levels with contribution levels. In one-way grids, motivation to contribute is required in order to reduce free riding. The literature is limited on grid contribution motivations, but according to a SETI@home survey, the main one is "the good of humanity" (59% of respondents). Other motivations may include the performance of competing contributor groups. Motivations revealed in other contexts could also support contribution, including the public display of one's abilities, the prospect of self identification, and status signaling.

| | | Nonrival | |
|---|---|---|---|
| | | **Low** | **High** |
| Nonexcludable | High | Wireless LAN<br>-Commons good | OSS, wiki-based content<br>-Public good |
| | Low | Private good | Computing grids<br>-Collective good |

Table 2. Typology of GARN-facilitated goods.

*Wireless networks.* Free riding manifests itself when a user password-protects his or her own access point or shuts it down when not using the commons [2]. Here, too, an effort to encourage the motivations discussed in the other contexts, including the prospect of self-identification and status signaling, helps reduce free riding.

*Content.* Like other nonexcludable goods, content lends itself to free riding. To mitigate against the limitations of online cooperation and collective action, online networks require ongoing interaction, identity persistence, and knowledge of previous interactions [5]. Other potential countermeasures discussed in the literature include: making people publicly commit to contribution; increasing social validity by demonstrating to potential contributors that many people like themselves have benefited from contributions; and creating contribution ratings by participants who rank contributors highly (the Slashdot model) [1]. In addition, countermeasures used in the OSS context may be useful, including incentives, such as the opportunity to display one's abilities and the opportunity to signal status and help others.

**Threat 2. Use of others' resources without their knowledge or consent.**

*Contexts. OSS.* Because OSS is nonexcludable, project hijacking is a potential threat, such as when an OSS is packaged with proprietary code in order to take it private [8]. Countermeasures include: adopting software that restricts proprietary appropriation; encouraging compliance with licensing terms through normative and legal sanctions; incorporation as a way to protect individual contributors from liability; transferring individual property rights to nonprofit corporations; trademarking a project's brands and logos; trademarking a foundation; and protecting a project's brand [8].

*Computing grids.* The threat is not significant in the grid context as long as contributing computers are protected from unauthorized access (excludability). Failure to provide this protection can result in others using their processing power or storage capacity. Ways to prevent the unwarranted use of resources includes: monitoring traffic; building firewalls; and establishing legal agreements among grid parties.

*Wireless networks.* Taking over contributors' devices is a significant threat; access is nonexcludable, making this commons good susceptible to the "tragedy of the commons." This threat can be countered through various security products [2]. In closed commons, it can be countered by monitoring the network for bandwidth use and removing devices that continuously overuse bandwidth [2].

*Content.* The threat here is the use of knowledge-based content produced without acknowledging the source. As both knowledge-based content and OSS are public goods, this threat can be countered in ways similar to the open-source software tactics identified in [8]. For example, Wikipedia is copyrighted, and proprietary appropriation is prohibited; Wikipedia is incorporated, and its contributors are protected from liability; individual property rights are transferred to a nonprofit foundation; and the Wikipedia brand is trademarked.

**Threat 3. Overuse of resources.** Also referred to as "overgrazing" in the literature [2] or free riding, it involves overuse without concern for the harm it causes others [10].

*Contexts. OSS.* The overuse of resources is not a significant threat to OSS; software, like many other information goods, is nonrival, and the cost of producing an additional unit of existing software is negligible.

*Computing grids.* In the case of two-way grids, the potential threat of overuse depends on agreements among the participants. In many cases, this threat is resolved by allowing a contribution only when the contributor computers are idle at the time of the contribution (so the resource becomes a nonrival good). In the case of one-way grids, the threat of overuse is mitigated against in a similar way, as resources are usually contributed only when the computer is idle.

| | OSS | Computing grids (users) | WLAN | Content |
|---|---|---|---|---|
| Giving/ receiving asymmetry | Increase motivation - benefits for contributors: learning, become lead users, tailor software to needs, fun, signal status, signal ability, support self identification, help others | Increase motivation: public display of abilities, the prospect of self identification, and status signaling | Increase motivation: self identification, status signaling | Make people publicly commit to contribution, increase social validity, create contribution rating by other participants; increase motivation: display abilities, signal status, help others. |
| Use of others' resources without knowledge or consent | Adopt software that restricts such proprietary appropriation, encourage compliance with licensing terms through normative and legal sanctions, incorporate as a way to protect individual contributors, transfer property rights to nonprofit corporations; trademark brands and logos, trademark a foundation;, protect brand | Monitor traffic, firewall, establish legal agreements between parties involved in the grid | Use security products, monitor the network for bandwidth usage, and remove devices that overuse bandwidth | Prohibit proprietary appropriation, incorporate as a way to protect individual contributors, transfer property rights to nonprofit corporations trademark brands and logos, trademark a foundation; protect brand |
| Overuse of resources | Not relevant | Contribute only when resources are idle | Limit the number of receivers, limit access to the resource for a certain period | Not relevant |
| Preventing others from using resources | Not relevant | Oversight body to monitor and periodically analyze user complaints | Oversight body to monitor and analyze user complaints and removes offenders from the commons | Not relevant |
| Damaging a particular resource | Review before release | Not relevant | Vaccinate own devices, avoid use of hotspots with nonstandard equipment, quarantine infected devices, examine new access points | Use "change history" function to restore content. use alerts to monitor content changes, block user accounts, anonymous IP addresses, and IP ranges |

Table 3. Threats and countermeasures.

*Wireless networks.* Here, the threat is significant, as wireless access is rival and nonexcludable; overuse of resources could result in their insufficient availability. Potential countermeasures include limiting the number of benefiting devices based on resource availability and limiting access to the resource for a certain period [2].

*Content.* In the case of knowledge contribution via Wikipedia and Wikipedia-like initiatives, the threat of overuse is low, as content, like software, is also a nonrival good.

**Threat 4. Preventing others from using resources in order to increase one's own availability of resources.**

*Contexts. OSS.* The threat of preventing others from using resources is not significant, as code is a nonrival good; preventing others' use of the resource provides little value.

*Computing grids.* The threat of being blocked from access is not significant in grids. However, a countermeasure would be like the one suggested in the case of wireless networks: having the oversight body monitor and periodically analyze user complaints.

*Wireless networks.* This threat manifests itself in the prevention of others' device operations so the remaining devices have more available bandwidth. It can be countered by assigning an oversight body to monitor and analyze user complaints, identify offenders, and remove the offender from the commons [2]. However, this method is possible only in closed commons where excludability is possible.

*Content.* As in OSS, this is not a major threat for content assets, as a content asset is a nonrival good.

**Threat 5: Damaging resources.**

*Contexts OSS.* Damage to resources is not a significant threat in the case of OSS, as software is reviewed and quality-controlled before it is released. Thus, the threat associated with nonexcludability is averted.

*Computing grids.* This is not a significant threat, due mainly to agreements among parties involved in computing grids and the ability to monitor members' behavior.

*Wireless networks.* As open commons are nonexcludable, this threat manifests itself in terms of someone adding a device with viruses to wireless commons or adding a device or access point that does not comply with 802.11X standards, possibly disrupting signals [2]. Ways to counter this threat in open commons include regularly vaccinating one's devices and avoiding the use of hotspots with nonstandard equipment. In closed commons, one should quaran-

The type of good largely determines the type and LEVEL OF VULNERABILITY to the threats associated with GARN.

tine infected devices and examine new access points for usability and standard compliance [2].

*Content.* Vandalism and corruption of files is relatively easy, as these are nonexcludable goods. A useful countermeasure against potential vandalism in Wikipedia is the "change history" function, which makes it easy to restore content. Moreover, the platform alerts users whenever a specific page is changed. In addition, potential vandalism stemming from differences over content may be thwarted by the fact that it is easy to create a new entry. In severe cases of vandalism, Wikipedia is able to block user accounts, anonymous IP addresses, and IP ranges (see Table 3).

## CONCLUSION

The type of good largely determines the type and level of vulnerability to the threats associated with GARN. Volunteer-based OSS and knowledge-based content—both public goods—are susceptible to free riding, which in turn may be a major threat to their existence. However, being nonrival, software and content are not susceptible to threats stemming from nonexcludability. Computing grids are more easily protected from free riding due to their excludability and the agreements among participants in two-way grids. Motivational issues may help mitigate the associated threats. Wireless networks—another example of a common good—are susceptible to the threat of free riding.

The GARN phenomenon has implications for both business and policy making. One way for companies to take advantage of the GARN phenomenon is to operate as intermediaries between givers and receivers; for example, a company may create an exchange for such resources as content, network access, or computing power. Other potential ways to capitalize on GARN may be for companies to contribute resources and use this contribution as a vehicle for marketing (such as contributing network access and using it to display the company's logo and other messages).

In terms of policy making, GARN seems to offer a more efficient use of resources, allowing better access to products and services for people who otherwise could not afford it. At the same time, GARN involves legal implications (such as ownership of IP and responsibility for actions through GARN-available resources).

Further research is needed to better understand and utilize the GARN phenomenon. While we have applied a descriptive, qualitative focus here, it may focus on such issues as the motivational and behavioral factors underlying GARN, the economic analysis of this model, and its legal implications.

## REFERENCES


1. Cheng, R. and Vassileva, J. User motivation and persuasion strategy for peer-to-peer communities. In *Proceedings of the Hawaii International Conference on System Sciences* 2005 (Mini-track on Online Communities in the Digital Economy/Emerging Technologies) (Hawaii, Jan. 3–6, 2005).
2. Damsgaard, J., Parikh, M., and Rao, B. Wireless commons: The tragedy and its causes. *Commun. ACM 49,* 2 (Feb. 2006), 105–109.
3. Frank, R. and Bernanke, B. *Principles of Economics.* Irwin/McGraw-Hill, New York, 2001.
4. Hertel, G., Niedner, S., and Herrmann, S. Motivation of software developers in open source projects: An Internet-based survey of contributors to the Linux kernel. *Research Policy 32* (2003), 1159–1177.
5. Kollock, P. Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology 24* (1998), 183–214.
6. Lakhani, K., Spaeth, S., and von Krogh, G. Community, joining, and specialization in open source software innovation: A case study. *Research Policy 32* (2003), 1217–1241.
7. Nov, O. What motivates Wikipedians? *Commun. ACM 50,* 11 (Nov. 2007), 60–64.
8. O'Mahony, S. Guarding the commons: How community-managed software projects protect their work. *Research Policy 32* (2003), 1179–1198.
9. Osterloh, M., Rota, S., and Kuster, B. *Open Source Software Production: Climbing on the Shoulders of Giants. Working Paper,* 2003; opensource.mit.edu/papers/osterlohrotakuster.pdf.
10. Ostrom, E., Burger, J., Field, C., Norgaard, R., and Policansky, D. Revisiting the commons: Local lessons, global challenges. *Science 284* (1999), 278–282.
11. \Varian, H. Markets for information goods. In *Monetary Policy in a World of Knowledge-Based Growth, Quality Change, and Uncertain Measurement,* K. Okina and T. Inoue, Eds. Palgrave, New York, 2001.
12. von Hippel, E. and von Krogh, G. Open-source software and the private-collective innovation model: Issues for organization science. *Organization Science 14,* 2 (2003), 208–223.

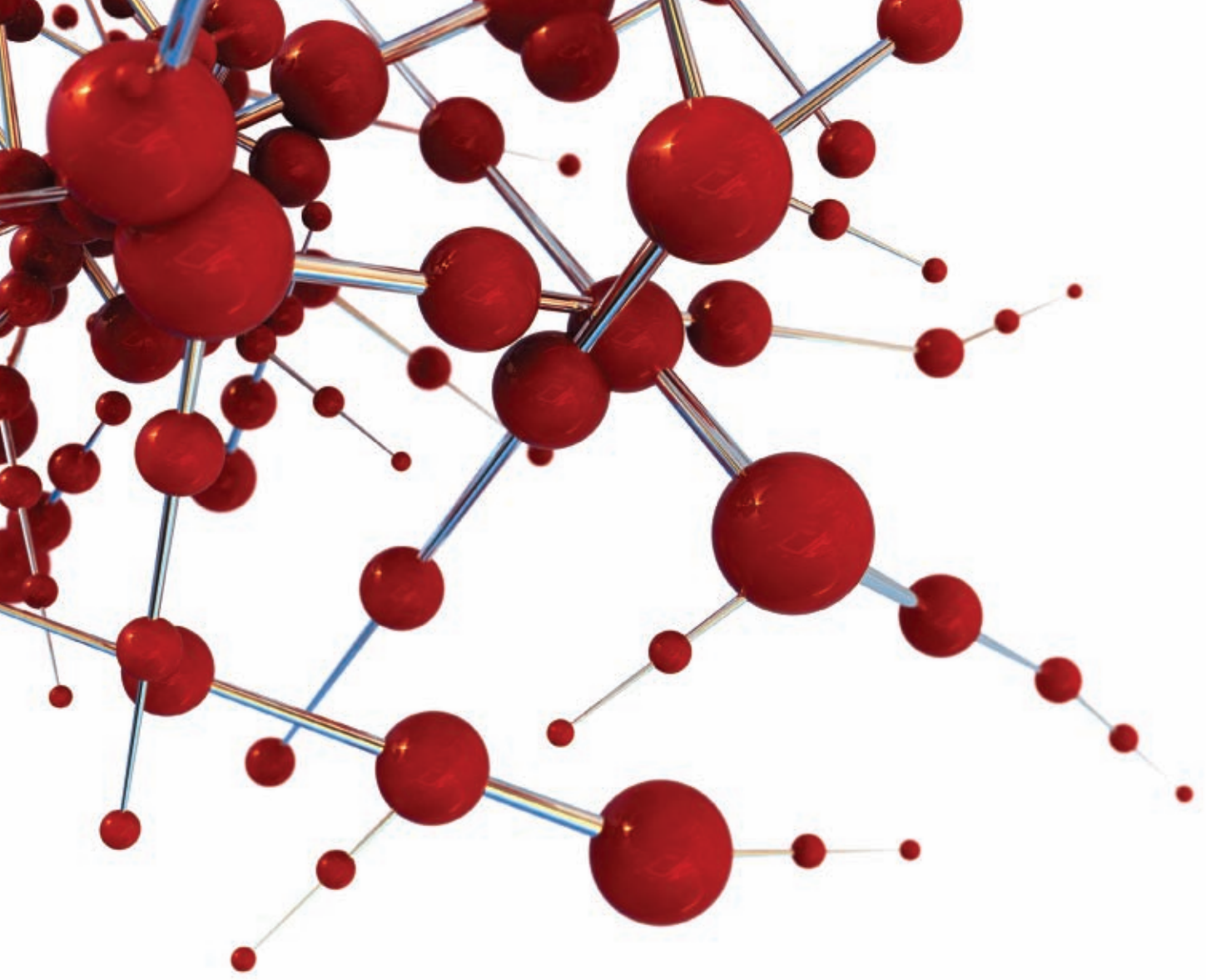


**ODED NOV** (onov@poly.edu) is an assistant professor in the Department of Management, Polytechnic University, New York.
**BHARAT RAO** (brao@poly.edu) is an associate professor in the Department of Management, Polytechnic University, New York.





DOI: 10.1145/1342327.1342342



By Bezalel Gavish and Christopher L. Tucci

*When it comes to online auctions, "caveat emptor" is an understatement.*

# REDUCING INTERNET AUCTION FRAUD

Fraud on the Internet is developing into a major issue for consumers, businesses, and governments [1, 6, 10]. The *Financial Times* in 2003 called online fraud "an epidemic of huge and rapidly growing proportions" and noted the incidence of fraud was 20 times higher online than offline [12]. The complaints of online fraud registered at the IC3 Web site—which is jointly sponsored by the U.S. Federal Bureau of Investigation and the U.S. National White Collar Crime Center—have grown from around 20,000 in 2000 to around 200,000 in 2007, which represents a compound annual growth rate of 39% [7]. At the same time, the dollar value of losses has skyrocketed at an annual rate of 50% from less than $18 million in 2001 to over $200 million in 2007 [7].

One area that is particularly interesting is auctions. Auction fraud reported to the Federal Trade Commission has likewise grown tremendously from 106 in 1997 to around 24,000 in 2007 [9] and continues to climb dramatically. Internet auction fraud is one of the main sources of overall Internet fraud, with estimates of incidence from 64% to 87% of all Internet fraud [4, 7]. However, major Internet auction sites estimate that fraud is involved only once in approximately 10,000 auctions [9], which appears to contradict these observations.

Understanding fraud with respect to Internet auctions is especially important because the oft-cited "network externality" in which having large numbers of buyers leads to large numbers of sellers, which leads to even more buyers, and so on. It is based upon the knowledge that the winner of the auction will receive what he or she was expecting. If more than a handful of buyers perceive that the system does not work in a fair and neutral manner, the entire network effect may start unraveling.

If the claim of such a low level of fraud were true, an auction site could encourage a higher level of buyer activity (and prices bid) by charging a minimal surcharge on the final winning bid to provide automatic insurance to buyers. Analyzing one auction site, we found that a one-dollar increase in the final winning bid translates into over $700 million in annual dollar volume. A surcharge of one promille per dollar of the final price above the additional transaction costs such as bookkeeping, investigation of claims, and issuing of compensation should be enough to provide insurance and to leave ample additional profit for the auction site operator. The fact that insurers do not offer such policies calls into question industry's estimates of fraud incidences.

This procedure is not as simple as it sounds, as having automatic insurance would have an effect on the behavior of both buyers and sellers. Buyers will be willing to tolerate a higher level of risk and bid higher even for unknown sellers, or sellers with lower levels of reputation. Another complication is that it creates incentives for crooked sellers and buyers to form coalitions in which they swindle the auction operator or the insurance entity. The advantage of an auction site operator compared to the occasional buyer or seller is that the operator has the resources and experience to develop methodologies that detect and reduce such swindling operations to a minimal level. By providing automatic insurance, the auction site operator would benefit from the higher level of volume in bidding activity.

We concentrate on buyers being swindled. In our study we discovered that sellers face a similar problem of cheating; however, auction houses have devoted extensive resources toward protecting sellers, but have invested limited effort into protecting buyers. Why? We contend that if sellers do not put items up for auction, the site cannot survive, while bidders/buyers on their own do not justify the existence of an auction site.

Another factor involves the fact that most sellers use the auction site multiple times (in many cases, thousands of times). As a result they develop experience and methods that protect them. Bidders, on the other hand, have limited experience with the auction process and do not have the expertise to protect themselves.[1] For example, in most auction sites, sellers do not ship a product to the buyer until the buyer has paid for the product.

## RESEARCH SUMMARY

We employed a variety of methods to undertake an exploratory investigation of Internet auction fraud. We performed a literature review of prior empirical studies on Internet fraud and auction fraud. This led to a preliminary and exploratory survey conducted in 2003. Our survey results also helped guide our participant-observation exploration: We bought and sold on major auction sites from 2003–2005 to better understand how swindlers work and what actions might be taken by the auction houses and by buyers. We also interviewed bidders and sellers that contacted us through our own buying and selling.

In our exploratory pilot study (see [5]), we surveyed 1,298 winners of Internet auctions at a major auction site to see whether they received what they were expecting. The respondents came from 14 different item categories and a full range of prices. We asked primarily two questions and collected other information from the auction site itself.

1. Did you receive any item after you won the auction in question?
2. If you did receive an item, was it what you were expecting?

We interviewed willing respondents to our survey. It was difficult to gather the data using automatic means, so we did everything manually. Generally speaking, auction houses put many obstacles to prevent automatic data gathering for such a research investigation. They limit the number of interactions one can have with other members per day. Thus, it is not obvious how to do such a study.

Staying within the rules set up by the auction houses, we were prevented from using mechanical methods such as Web crawlers; otherwise, it would have been much easier to contact a large number of auction winners. Further, the auction houses are willing to give researchers information, but what they offer is mostly meaningless for such a study. For example, looking simply at feedback does not convey much information. It carries a value of 0, neutral, or 1, so there is no magnitude. For example, a "winner" in Athens, Greece, whom we talked to lost close to

[1]This can be ascertained by observing the distribution of the number of transactions or the number of feedback ratings for sellers and buyers.

Auction houses have devoted extensive resources toward protecting sellers, but have invested limited effort into protecting buyers. Why?

$10,000 and felt very strongly about his loss. His situation and (negative) feedback for the seller carried the same weight as someone who lost one dollar in an auction.

## RESULTS AND ANALYSIS

The results of our preliminary survey are as follows: We received 98 responses from the auction "winners," 21 of which either did not receive any item at all, or did not receive what they were expecting (mostly because the item was damaged or otherwise in worse shape than described). Further, eight out of 98 received absolutely nothing at all, which represents 0.62% *minimum* (assuming everyone else, including non-respondents, received something, as we will discuss). This incidence would be 62 times higher than official estimates of fraud, and this is clearly a subset of all fraud.

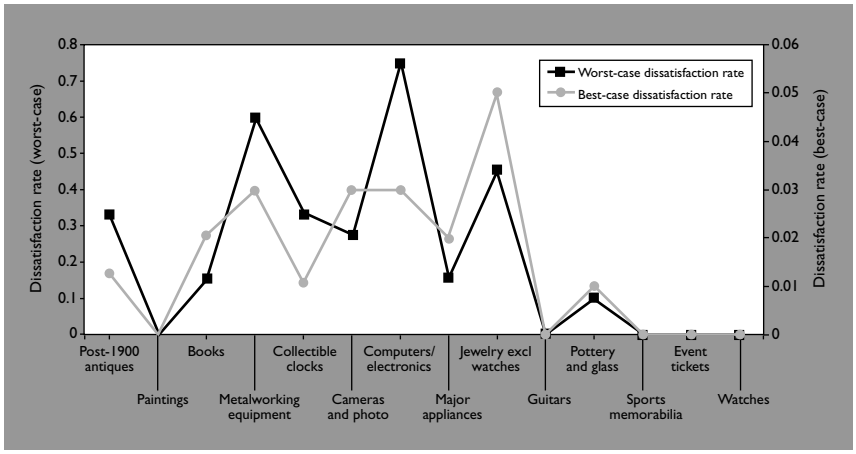



**Figure 1. Dissatisfaction rates for different categories.**

If either the buyer received nothing or did not receive what was expected, we call that "swindled" (recognizing, of course, that the buyer may or may not have been intentionally swindled and that this is definitely a superset of purely fraudulent activity). *Still, knowing the incidence of "swindling" is important because it most certainly is an input into buyers' perceptions of the fairness of auctions.*

We calculate the worst-case rate of negative response (by dividing the total negatives by total responses) and the best-case rate (by dividing the total negatives by total contacted). It depends on how one views the representativeness of our sample to know which is more "accurate." If one feels that the respondents are representative of the population at large, then the worst-case estimate would be closer. If one feels that our respondents are "disgruntled," then the best-case estimate would be closer.

In addition, we have two opposing factors. For those who are embarrassed about being cheated, they will be underrepresented. They do not want to report that they were cheated because it would make them look foolish. On the other hand, there are people who were cheated on one auction and are so mad they want to tell their story, even if we asked them about a different auction, and that will lead to overestimation. The true rate would most likely fall somewhere in between the best- and worst-case estimates.

The percentage of negative responses varied for different categories of auctions (see Figure 1). Computers and electronics had the highest worst-case rate of negative responses: three-quarters of respondents either did not receive their computer or it arrived damaged. In terms of the best-case rates, the worst category was Jewelry (excluding watches), in which 5% of all auction winners contacted either did not receive anything at all or did not receive what they were expecting. Four categories had no negative responses (Paintings; Guitars; Event Tickets; and Watches). One surprise was that Jewelry, excluding watches, had one of the worst rates of dissatisfaction, while Watches had one of the best. In one category,

Buyers afraid of being swindled will stop participating, thus fewer buyers, meaning fewer sellers, and so on.

Sports memorabilia, we received no responses at all. Of course, these category measures should be taken with a grain of salt, as the number of responses within each category is quite low. Still, the overall response indicates a worst-case estimate of 21.4% negative, with a best-case (and conservative) estimate of 1.6% negative.

We also found the following contingencies:

- Price does not appear to be a very good predictor of whether the buyer was swindled;
- 25% of the respondents who won auctions with a selling price below an insurability cutoff of $500 were "swindled," versus only 5% of those winners whose items sold for more than the cutoff (we will discuss "moral hazard" later);
- The location of the seller is not related to whether the item arrived intact in our sample;[2]
- Having a photo is not associated with whether anything was received; and
- 26% of respondents whose auctions had a photo did not receive the item intact vs. only 11% of auction winners where there was no photo.

Perhaps one lesson to be learned from this breakdown is that buyers appear to be more careful in situations in which the dangers are more obvious.

Based on a survey of literature (for example, [2, 3, 10]), plus our interviews, plus our own participation in auctions, we derived some generalities in the swindling area. First, there are the methods used to actually execute the swindle. Second, there are methods used to avoid appearing fraudulent. The main methods used to actually execute the swindle are shown in Figure 2, while the methods used by swindlers to appear as a legitimate seller to potential bidders are shown in the accompanying table.[3]

| | |
|---|---|
| Selling cheap items | To establish a good track record, they put up many low cost (a few dollars) items for sale and provide excellent service. The feedback from the buyers is highly positive. Once an excellent track record has been established, they go for the kill by putting up for bid expensive items that are never delivered. Once discovered, they repeat the same process under a different identity. |
| Taking advantage of pooled buyer-seller feedback | Some auction houses provide rating for sellers and buyers, but they do not distinguish between the seller's past selling and buying feedback. Smart swindlers buy and sell many items for a penny or so. They provide excellent feedback to the sellers. Once they have established a long positive record, they put up items for auction. |
| Changing seller ID | Using several IDs so there is no track record. Auction houses try to overcome such practices by requiring sellers to provide legitimate credit card information in order to receive a seller ID. Unfortunately it is easy for swindlers to receive several credit cards and use them to open accounts on the auction site as sellers. |
| Changing terms | Changing the payment method. For example claiming that they accept credit card payment but after the auction, insisting on non-credit card payments. |
| Changing location | Use Internet broadband phone service to establish a U.S. phone number. Use a fake U.S. shipping address while actually being based overseas. |
| Phishing | Phishing can provide sellers with fake IDs and other information that permits them to take over an established high rating user account. |

**Methods swindlers user to appear legitimate.**

## WHAT CAN BE DONE TO REDUCE FRAUD?

Our findings, although preliminary, could spell trouble for bidders in online auction sites, and ultimately the sites themselves. Even though the number of buyers and sellers has been rising rapidly over the last decade, any sudden shift in perception could reverse the cycle [10]: Buyers afraid of being swindled will stop participating, thus fewer buyers, meaning fewer sellers, and so on.

We have several broad categories of recommendations based on our study. In this section, we review each category and give specific recommendations to

[2]Although if we think that swindlers use fake U.S. addresses, we would not expect the location to be associated with swindling. Incidentally, broadband Internet phone service—wherein non-U.S. swindlers can obtain U.S. phone numbers and answer them anywhere in the world—while providing a very valuable service to legitimate customers, is actually exacerbating this problem.

[3]Buying and selling on the Internet is a very dynamic system that is evolving over time. For every fraud detection procedure that auction houses institute, swindlers adapt and deploy countermeasures that overcome the new defenses. We realize that we might not have covered every method; if a reader knows of anything else, please send it to us so we can continue with our larger follow-up study.

reduce auction fraud along with the advantages and disadvantages of doing so. The broad categories are the following:

1. Increase the information shared by auction houses on buyers and sellers.
2. Make the use of legitimate escrow services extremely easy, and possibly mandatory.
3. Charge the seller an amount equal to the amount of the sale (temporarily).
4. Develop sensible insurance policies.
5. Institute regulatory control over the auction houses.
6. Encourage buyers to protect themselves.

Note that the first category is probably the most important, and the first and last are the easiest to implement.

**Information on sellers.** Generally speaking, the more information is disclosed on sellers (and buyers) the easier it is for buyers and sellers to verify that they are dealing with a reputable entity, leading to a higher degree of confidence in the system. The problem now is that the system is based on self-reporting by both buyers and sellers. The auction house cannot guarantee that the information they provide is accurate.[4] It is still the responsibility of the buyer and seller to verify the data.

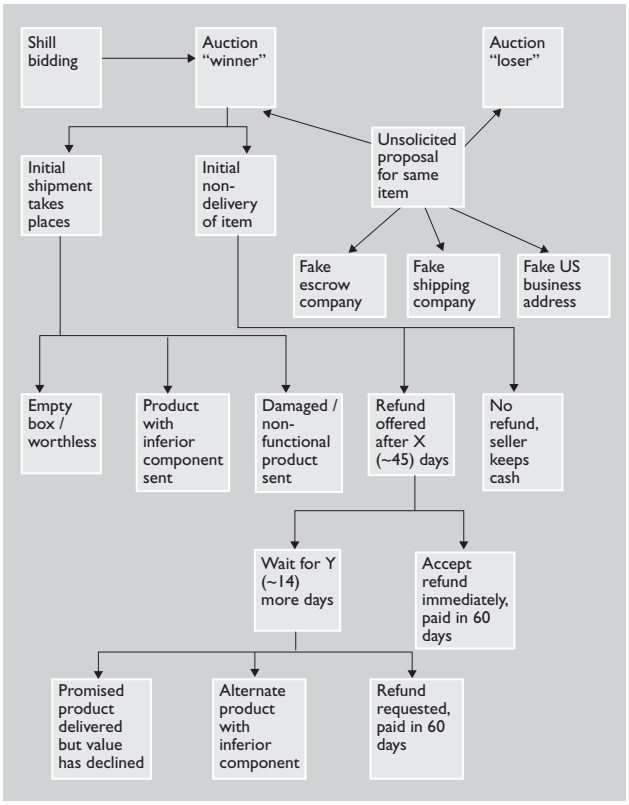



**Figure 2. Possible Internet auction swindling methods.**

Auction houses collect a significant amount of data. Making more of it public increases the chances that a transaction will be valid. Here is some data that should be posted for all reputation scores:

- The *percentage* of positive responses of the total number of transactions for sellers (some sites began doing this in 2003) rather than the absolute number. In many cases, the auction houses only reveal the number of positive scores, thus if someone sells 1,000 items, of which 100 are positive, she or he would have a higher score than someone with 99 positives out of 99 auctions.
- As mentioned in the table and discussed in Chua and Wareham [2], some sellers sell (or buy) a large number of cheaper items to establish their reputation. To address this, we propose posting the average selling price (or a distribution of selling prices) of all the seller's previous auctions. It would be even better to divide it into months and also by item category.
- Separate statistics for selling vs. purchasing. This reduces the chances that a seller can establish a positive reputation by simply buying many low-cost items, which he subsequently sells thus doubling his positive score.
- Seller activity history (not score) as a function of time. The pattern of activity and dollar value can provide clues to participants.
- The number of distinct users (buyers) that participated in creating a reputation history for a seller.
- Statistics on the number and percentage of cases in which the seller was not paid, or buyer did not pay (indicates a shill bidder running up the price if it happens too often).
- Information on the seller, for example, listing the physical (as opposed to electronic) address for payment at the end of the auction. Such physical information can help determine if this is a legitimate seller or a swindler. In addition, there are Internet-based tools that could be used to disinguish real from purely virtual sellers. Auction houses could accomplish this by sending a physical letter to the postal address of the registered user, much as a bank sends a PIN under separate cover to a postal address.

[4]Using tools available on the Net it is possible to verify a significant part of the information needed to establish the person or company identity, for example, does a person with that name exist in his address? Is the phone number associated with him? Is such a company declared and filed with his local authorities? What is his educational background? Does he have a family? With whom did he deal in the past? In some cases, what is his credit score? The amount of useful data that can be collected on the Internet pertaining to an individual is surprising.

- Some auction houses limit the history available on auction results. For example, eBay gives out details (details about individual item auctions) for only 90 days after activity. A crook makes sure she or he goes 90 days back and gets good comments selling cheap items, then does not sell too many items for 90 days. The auction houses can do better by keeping the history longer. (This would raise the cost of storage but we do not feel that it would be a very large expense amortized over many auctions; further it would increase confidence and liquidity.)
- Give out actual email addresses or at least release actual email addresses and other information to the other party under certain circumstances. For example, auction winners could demand it before final payment takes place. Auction houses could inform the seller automatically that their information has been given to the winning buyer. Auction houses now avoid this practice to prevent buyers and sellers from taking their transactions "offline" and thus circumventing some of the fees charged by the auction house.
- Flags based on statistical analyses of past behavior by sellers (for example, using quality control methods). For example, if price skyrockets or volume jumps dramatically for this seller, the auction would be "flagged" for all bidders to see, thus raising awareness of the potential for swindling. The seller could have the flag removed through some certification process, for example, proving that he or she actually has the items. Another alternative is the computer gives the item the flag and the seller has the choice of continuing or not.
- Develop efficient methods for bidders to alert the auction houses about swindling sellers. Swindlers can discover the response times of auction houses to swindling alerts, for example how quickly do they remove a false auction, or announce the existence of a swindler to the bidding community? Based on the expected response time information, swindlers design the timing of their selling activity.
- Mechanism for alerting other users about sellers. This one must be treated very carefully. The advantage is that by quickly getting the word out, one avoids others from falling victim to the swindler. On the other hand, it is difficult to adjudicate and decide when someone should be "blacklisted." It could be that a buyer hates a seller and uses it for revenge; or purely for strategic reasons, a buyer gives out some alert that reduces the price so the buyer himself can get a better price! Or, a seller can use such a mechanism to remove (temporarily) other sellers of competing items. It is understandable why auction houses have shied away from this one, but perhaps some kind of "Amber Alert" for auction swindlers could be developed, perhaps by holding accusers responsible for issuing the alert.
- Divide sellers into classes based on their past performance and rating by buyers. Use different flags to identify reliable sellers from less reliable ones (for example, classes of power users on eBay).

*Advantages and disadvantages.* The more information buyers and sellers have on each other, the better off they are. We should keep in mind that this additional information can be used by swindlers to better target their messages/offers to potential victims. Auction houses, on the other hand, have interests in protecting the identity of buyers and sellers, so it is a question of balance. All of these involve disclosing information about buyers and sellers (mainly sellers). In our opinion, however, legitimate sellers should be happy to reveal their information to benefit from the price premium of being reliable.[5] This direct cost is negligible, involving hiring a few people to develop the algorithms and the software to display the new calculations.

**Escrow.** Extremely easy, seamless, and possibly mandatory escrow services should be available for all auctions. A small service charge could be added to the listing fee to facilitate the transaction. We propose the following modifications to the escrow systems currently in place:

- Traders should only use escrow services officially certified by the auction house. Auction houses should develop a list of escrow services that they have certified and provide a direct link to the sites.
- Shipping the item to a third reliable party, who inspects it, instructs the buyer to pay the seller, and once payment (or proof) has been received, transfers the item to the buyer. This incurs an extra shipping cost, which may make it prohibitive for less expensive items; still, for expensive items, the shipping cost is minor compared to the cost of the item.
- We notice that the premium to reputation is about 20% on expensive items. There is a busi-

[5]Revealing the sellers identity on the auction site could be an option that the seller selects, thus giving the indication that he is legitimate. Sellers of very expensive items (cars), list in the body of the listing information that allows potential buyers to verify that they do exist and do business (their dealership title, address, phone number, vehicle location and stock number, VIN, and many pictures of the car).

The schemes proposed here increase the knowledge of buyers about the quality (or even existence!) of the items to be sold.

ness opportunity for an "intermediary" to come in, receive the item before the auction, verify it, and then put a seal of approval/authenticity of that auction [2]. The seller would receive a better price, which would be shared with the intermediary. The fact that such a service does not exist raises some questions. Auction houses could also make a feature that enables one entity to certify another's auction. Even having an apparent certification by such an entity is not a guarantee.[6]

- A casual browsing of auction sites reveals that sellers often give blatant instructions not to check out through the official auction checkout procedure. These sorts of circumventions should be monitored and policed more effectively. It could even be done in an automated way by searching for keywords proposing to circumvent the checkout procedure.
- Some sellers offer money-back guarantees if the buyer is not happy with the items he won. What is lacking is a mechanism that actually enforces that offer when a buyer is unhappy.

*Advantages and disadvantages.* The advantages include a steep reduction in information asymmetries at the expense of swindlers. Information asymmetries exist when one party does not know everything relevant to a transaction that the other party knows. The schemes proposed here increase the knowledge of buyers about the quality (or even existence!) of the items to be sold. The disadvantages include greater transaction costs, even in the more mild versions and the possible exclusion of some buyers from the market (some buyers are willing to take a chance on being swindled for a lower price—those buyers would be priced out of a market in which escrow was mandatory). One could imagine an extension of this in which only auctions closing above a certain dollar amount were forced to adopt some of these techniques.

**Performance Bond.** To open an account as a seller or buyer, one needs a credit card. The auction house should charge the seller an amount equal to the amount of the sale (temporarily). Then, once the buyer gives the go-ahead, the auction house would reverse the charge.[7] Auction sites already do this on the buyer side. The amount charged could also be based on the seller's reputation or other characteristics of the seller (such as, how many prior items sold at the current price range). This "seller escrow" would also benefit sellers without an established reputation. If the credit card transaction is rejected, than the auction house knows that this is not a legitimate seller and will remove his or her ID and items from the auction site. The auction site would also know immediately that they are probably dealing with a swindler.

The company would need to provide a cash flow management guarantee (return your money within, for example, one day), though, to protect those legitimate sellers who make a living selling on auctions. Even so, most sellers would only take a one-day hit because after the first day, they would be applying their performance bond to future auctions.

*Advantages and disadvantages.* The performance bond is a very strong level of protection and would sharply increase trust in the system as well as sharply reduce swindling. However, it is fundamentally a little unfair as it requires sellers to be relatively well-off, with enough credit on their credit card to pay for the items they already own. If someone was auctioning off items to raise money, say for a hospital bill, and they were already in debt, they would not be able to participate as a certified seller under this system. An alternative could be some kind of certification by the auction house for unusual circumstances, but this is likely to fail for two reasons: the auction houses would like to avoid certifying individual sellers for legal reasons; and the less "automated" the process, the more

[6]Choicepoint is an example of a false and dishonest entity that appeared as a legitimate business for a number of years.

[7]Some car rental and hotel companies use a similar method to protect themselves against nonpaying customers.

likely that actual swindlers would be able to pass off as legitimate customers. A third alternative could be the encouragement of third-party certifiers (or surety bond companies) such as BuySafe, which certify sellers for a commission—paid by the seller on items sold—and promise to compensate buyers if there is a problem.

**Insurance.** Appropriate insurance policies could encourage electronic commerce activities. Premia should be reasonable, and mechanisms designed to reduce or prevent collusion. The way it would work would be that the buyer would pay a premium to the insurance company and if there is a problem, the buyer makes a claim and is reimbursed. Here are a few possible recommendations for insurance:

- Insurance by payment beforehand. At the moment the auction is consummated, the buyer is charged a premium based on the seller's reputation and the category. These are both known in advance and so could be prominently advertised within the auction itself while it is going on. The buyer would pay the premium, although there would be nothing preventing the seller from subsidizing or even paying the premium him- or herself, just as some sellers pay for shipping.
- Insurance companies should be certified by the auction house, with recommended insurers linked directly from the auction house.
- Insurance premium size should be tied to the level of fraud in the category and the past reputation/insurance claim activity related to the seller. Insurers would then have the tools to go after swindlers and to increase the cost of being a crook.
- Another option is that the seller would be required to take out the policy.

*Advantages and disadvantages.* Smarter insurance policies would make the market fairer, especially if taking out a policy were mandatory. It would be fair because buyers would be hedged against absolute fraud and would not lose their money. In addition, the market for insuring fraud would become more efficient, because there are relatively few insurers and they would be scrutinizing the antecedents of fraud and pricing the risk accordingly as their own money would then be on the line if they are wrong. The disadvantage is that on the margin, the increased transaction cost would exclude some buyers. In addition, there could be a moral hazard problem in that buyers would be participating in riskier auctions (ones they estimate might have a fraudulent outcome) but they do not care as much because their insurance coverage provides a safety net.

**Regulatory control.** State or federal governments in the U.S. and elsewhere may want to make sure that the auction houses follow the rules set in their own descriptions of the mechanisms. There is a need for an agency that confirms they follow without bias the rules they advertise. Auction houses have at their disposal the ability to take advantage of situations in which they can make extra money at the expense of sellers and/or buyers, just as a real estate agent could conceivably buy and then resell a house rather than showing it to a potential buyer. We are not claiming that auction houses intervene in the auction markets frequently, but in fact who would know if they did or did not? Are the auction houses completely neutral? Some sort of government oversight might be helpful in this area.

In return for submitting to regulatory control (and for taking some proactive steps toward reducing fraud as outlined in this article), we propose that auction houses could be shielded from some forms of legal liability in fraud prevention. It could be that auction houses have resisted any action up until now for fear of legal reprisals, in other words, if they pursue better escrow or insurance policies, that might be considered an admission of guilt and open them up to lawsuits by unhappy auction winners.

*Advantages and disadvantages.* Without any a priori knowledge of biased interventions in their own auctions, we are not certain that instituting government oversight will create more problems than it solves. Still, adding government oversight will make the entire system more transparent. Further, if auction houses could be shielded from some legal liability through an oversight process, it might give them incentives to implement anti-fraud policies rather than simply claim they are a neutral market and are not at all responsible for members' trades.

**Buyer precautions.** There are many tools available on the Internet to help either party verify the information given by a seller. Examples:

- Auction houses should ask a buyer (and seller) during the registration process to provide a phrase/code that is known only to them. Any email message coming from the auction house to him should have that phrase displayed in it. This enables the buyer (or seller) screen out fraudulent messages.
- Putting the name of a company and the word "fraud" or "scam" into a search engine like Google.
- Check with the Chamber of Commerce and Better Business Bureau regarding the company.

- Do a telephone number search (a reverse search and check the name of the party owning that number); legitimate companies should have their own number. The buyer can also call that number to make sure that a live entity is behind that number. You can also cross match it to an email address.
- Occasionally, a seller asks to send cash or equivalent. What one can do is send an email saying that "by coincidence I am passing through your city. Can I drop by, pay in cash, and pick up the item?" If the seller does not respond, or says I only ship, do not do business with that person. Smart crooks could bluff, however, so there is no guarantee.
- Through the exchange of email, one acquires much more information about the entity with which one is dealing. Phone number, name, address, tracing the routing of response messages and so forth. Then one can use the Internet to check on the entity, as discussed earlier. If the seller does not respond to your questions/email this should raise suspicion.
- Always pay with a credit card. Some sellers will not sell the item to you, but that is the price you pay for security. Some swindlers allow PayPal, but they only accept the one that comes from your bank account, not the kind of charge that comes through your credit card.
- When you receive a package, open it in front of the delivery person or company representative. You can tell your delivery company not to drop off without a signature.
- Go to an "eBay Xchange Point," or something equivalent. In Switzerland, eBay has proposed meeting points at major train stations where buyers and sellers can meet face-to-face to inspect the goods and pay for them. This is not entirely practical for goods bought far away, but for more local transactions, nothing beats face-to-face.
- Chua [2, 3] also proposes being more proactive in reporting and preventing fraud through such techniques as collective action on reporting (for example, to Traderlist), contacting other potential victims directly, or even "vigilantism" in which buyers deliberately sabotage auctions they feel are fraudulent.

*Advantages and disadvantages.* The advantages of buyer precautions are obvious, they reduce fraud and make it more difficult for crooks. They do not cost as much, if anything, to implement. The disadvantages are that not everyone would take advantage of them.

## CONCLUSION

Auction houses today appear to be at a crossroads, with many people now losing confidence in the system [8]:

"...do you have a reasonable expectation that the Mac G5 you are bidding on is going to be shipped to you? How about the Sony plasma 60-inch monitor? Or, the Sony PSP for $50 to $100 over retail? No, you don't. Now, you might say, 'Hey, moron! Why are you bidding on stuff like that on eBay? Don't you know that 95% of those auctions are fraudulent, especially the ones that only accept payment via Western Union?' Of course, I know. But what about all of the people who don't?"

Now is the time to act before the negative cycle mentioned at the beginning of this article develops in full swing. We hope that some of the recommendations listed here restore the public's confidence in the system. While it may be impossible to eliminate Internet auction fraud, at the very least we may be able to reduce it drastically and make it very expensive for those who persist. C

## REFERENCES


1. Bajari, P. and Hortaçsu, A. Economic insights from Internet auctions. *J. Economic Literature 42*, 2 (2004), 457–486.
2. Chua, C. and Wareham, J. Fighting Internet auction fraud: An assessment and proposal. *IEEE Computer 37*, 10 (Oct. 2004), 31–37.
3. Chua, C. Wareham, J. and Robey, D. The role of online trading communities in managing Internet auction fraud. *MIS Quarterly 31*, 4 (2007), 759–781.
4. eMarketer. The ePrivacy and security report. Jan. 2001. White paper published by MarketResearch.com.
5. Gavish, B. and Tucci, C.L. Fraudulent auctions on the Internet. *E. Comm. Research 6*, 2 (2006), 127–140.
6. Hafner, K. With Internet fraud up sharply, eBay attracts vigilantes. *New York Times,* (Mar. 20, 2004).
7. Internet Crime Complaint Center. 2007 Internet crime report. *National White Collar Crime Center and the Federal Bureau of Investigation* (March, 2008).
8. Palmer, S. Can eBay be saved? *OnlineSPIN*. MediaPost Publications (June 16, 2005).
9. Reuters. Read this before buying on eBay. *Wired News* (Jan. 9, 2003); www.wired.com/news/ebiz/1,57153-0.html.
10. Snyder, J.M. Online auction fraud: Are the auction houses doing all they should or could to stop online fraud? *Federal Communications Law Journal 52*, 2 (2000), 453–472.
11. Trade Commssion of the U.S. (FTC) releases list of top consumer fraud complaints in 2007 (Feb. 13, 2008); www.ftc.gov/opa/2008/02/fraud.shtm.
12. Waters, R. How fraudsters set traps and take the credit. *Financial Times IT Review* (May 21, 2003), 1.



**BEZALEL GAVISH** (gavishb2000@yahoo.com) holds the Eugene J. and Ruth F. Constantin Distinguished Chair in Business at the Cox School of Business at the Southern Methodist University, Dallas, TX. **CHRISTOPHER L. TUCCI** (christopher.tucci@epfl.ch) is a professor of Management of Technology at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, where he holds the chair in Corporate Strategy & Innovation.





DOI: 10.1145/1342327.1342343

Edgar A. Whitley and Ian R. Hosein

# Departmental Influences on Policy Design

## How the U.K. is confusing identity fraud with other policy agendas.

Problems of identity fraud are becoming common in all countries and increasingly governments are expected to be taking action to address these problems. Yet we understand little about the nature of the problem, and even less about proportionate solutions.

In many cases, identity fraud arises in relation to financial transactions, for example, when an individual's identity is used to fraudulently open a bank account or withdraw money. Occasionally, however, it is more than just financial inconvenience that results. For example, Derek Bond from Bristol, U.K. was arrested in Durban, South Africa in February 2003 for crimes committed by a Las Vegas criminal who had stolen Bond's identity documents. Bond spent three weeks of his vacation in jail at the behest of the U.S. Department of Justice before the truth was uncovered.

Individual cases capture our attention but figures are often better to focus our concerns. The scale of identity fraud is often difficult to measure, in part because a variety of definitions of identity fraud (or identity theft) exist [2] and there is no certainty that different reporting organizations are using the same definitions of identity fraud in compiling their figures. In addition, it is often unclear as to whether the reported fraud is due to problems of identity or other matters [8]. For example, in 2006 the U.K. government announced that the cost to the U.K. economy of identity fraud had risen from £1.3 billion in 2002 to £1.7 billion per annum[1] with part of this difference arising from the inclusion of approximately £400 million from sources "not included in the 2002 study." In addition, it was claimed that losses from fraudulent use of payment cards, or using a fictitious identity to obtain such a card, was £504.8 million per year. The government had attributed that figure to the U.K. Payments Association, APACS. However, when approached by the media, APACS reported that this form of identity fraud had totalled only £36.9 million in 2004, and in the first six months of 2005 they had already experienced a 16% drop in fraud, principally as a result of the introduction of chip and PIN technology for point-of-sale verification (replacing signatures), according to APACS spokesman Mark Bowerman.[2] In 2006, there was a further 3% drop in the amount of money lost to credit card fraud.[3]

Given this complexity in even identifying identity fraud, it is not immediately obvious which branch of government should be responsible for implementing measures for combating the problem. As the table here indicates, different countries place responsibility for addressing identity fraud within the scope of different government departments (see [7]). The choice of government department that designs the policy on this issue directly influences the kinds of approaches and other policy agendas enrolled in the solution. The response and emphasis of a department of consumer affairs is likely to be very different from that

[1] Cabinet Office, *Identity Fraud: A Study*, 2002; www.ips.gov.uk/identity/downloads/id-fraud-report.pdf. Home Office, *Updated Estimate of the Cost of Identity Fraud to the U.K. Economy*, 2006; www.ips.gov.uk/identity/downloads/FINAL-estimate-for-annual-cost-of-fraud-table-v1-2.pdf.

[2] McCue, A. Government ID fraud claims: Are they what they seem? Costs UK £1.7 bn a year? Figures "not an exact science"... Silicon.com 2006; www.silicon.com/publicsector/0,3800010403,39156140,00.htm.

[3] BBC News, Reduction in card fraud in 2006, 2007; news.bbc.co.uk/1/hi/business/6445409.stm.

of a department with policing responsibilities and will differ from departments responsible for trade and industry.

## IDENTITY MANAGEMENT AND IDENTITY FRAUD

The U.K. government has delegated powers for implementing identity management solutions to the Home Office (equivalent to Interior or Justice departments in other countries). As a result, the U.K.'s efforts to combat identity fraud are closely aligned with other parts of the Home Office policy agenda. These include crime, policing, passports, and immigration; the scheme they developed directly reflects this wider policy agenda of the Home Office.

| Region | Country | Government Department |
|---|---|---|
| Africa | South Africa | Department of Home Affairs |
| Americas | Canada | Office of Consumer Affairs |
| Americas | U.S. | Federal Trade Commission |
| Asia | South Korea | Ministry of Government Administration and Home Affairs |
| Australasia | Australia | Attorney General |
| Europe | U.K. | Home Office |

**Government departments responsible for combating identity fraud.**

The Home Office proposed issuing biometric identity cards, linked to a central identity register. Through a combination of extensive biometric collection (at one point including 10 fingerprints, two iris scans, and a face-recognition biometric) and a detailed, semi-automated biographical footprint check, the government intended to develop a de novo, clean database of all U.K. residents. Once issued, the biometric identity card could be verified against the National Identity Register in such a way that it would be virtually impossible, in theory, for someone to impersonate another individual. For example, every time a new bank account is opened or a credit card is issued, the bank or issuer would have to verify the card (and perhaps the biometrics of the card holder) against the national register. The lack of standards for the representation of biometric data at this time would mean that all banks and other such institutions across the country would need to have the same types of sensors to verify biometrics of their clients as were used to enroll people in the scheme, at each of their tens of thousands of branches.

A large centralized system seems almost inevitable once it is decided the policing arm of government will be responsible for combating identity theft. It is no surprise the resulting scheme has been widely criticized [5, 6] in part because the U.K. government has a relatively poor record of successfully implementing very large IT systems [3].

By choosing a high-tech solution, drawing on the state of the art in biometric technologies, the scheme is also high-risk. Few of the constituent technologies have been used on the scale envisaged by the identity cards scheme (60+ million citizens are expected to be registered once it is up and running).

Another question merits asking: Why the inclusion of fingerprints into the register? They are no more, and more likely much less, effective than iris-scanning technologies. The answer was provided in an email message from Prime Minister Tony Blair to those who had signed a petition against the introduction of identity cards: "The National Identity Register will help police bring those guilty of serious crimes to justice. They will be able, for example, to compare the fingerprints found at the scene of some 900,000 unsolved crimes against the information held on the register."[4] Thus the decision to locate measures against identity fraud in a government department that is also responsible for policing results in a scheme that seeks to address both of these policy agendas.

This centralized scheme, together with a single National Identity Registration number, has the potential to make the problem of identity fraud greater,[5] as the problems with the U.S. Social Security number and Australian Tax ID have shown [1, 4]. Though the U.K. government would argue that a government-certified high-tech solution would make it more difficult to perpetrate such fraud, it is likely the new solutions are only

[4]Tony Blair, PM's response to ID cards petition, 2007; www.pm.gov.uk/output/Page10987.asp.

[5]Young, K., Microsoft slams UK ID card database: Central database could lead to 'massive identity fraud'. VNUNet.com 2005; www.vnunet.com/vnunet/news/2144113/microsoft-slams-uk-id-card.

offering new vulnerabilities while dangerously increasing our confidence in a scheme that is advertised as the 'gold standard' for secure identity management.

Another problem faced by the Home Office in implementing identity cards is the process of enrolling the support of other government departments and industry to make use of the scheme. By linking enrollment into the Identity Cards Scheme with the voluntary renewal of passports (also managed by the Home Office), the department is able to ensure a relatively smooth rollout of the scheme over a 10-year period. However, as a consequence, for the first four or five years of the scheme, fewer than half of the eligible population will have identity cards. Until nearly all the population is enrolled in the scheme and has been issued identity cards, there will be little incentive for organizations to buy into the verification services of the scheme, affecting the cost-effectiveness of the scheme as a means of providing identity management solutions for the country [2]. This problem is heightened with the recent announcement that the rollout of identity cards to British citizens will be delayed until at least 2011 or 2012.[6] If identity fraud is indeed getting worse every year, it will get much worse before the solutions devised nearly a decade and a half earlier have any significant effect.

Moreover, by focusing on high-tech solutions, the Home Office risks downplaying other, lower-tech, solutions that might be equally effective. For example, one recent recommendation is that all consumers be given a free copy of their credit rating every year. Giving individuals access to the means of discovering whether or not they are being impersonated is one of the most powerful means of combating this form of fraud. Another solution would require banks and credit card companies to bear the risk of identity fraud and as a result the market could come to its own solution.

Other such possible measures that could help address identity fraud include:

- Working with the credit reporting industry to ensure that, on an opt-in basis, access to files involves security measures (prompt questions and so on);
- Helping industry to develop a secure means of automated notification whenever files are accessed or amended;
- Making paper shredders sales-tax exempt and tax deductible; and
- Promoting secure online account activity to reduce the amount of paper documentation in circulation.

A final recommendation, which again would be more meaningful coming from a government department with responsibility for trade or finance, would be to require public disclosure of all data losses and mass data thefts from companies and governments, following on the trend started by a number of U.S. states. When people are more aware of security risks they may be in a better position to judge the likely benefits of emerging solutions including biometric technologies and credit-managing companies. After all, a better understanding of the nature of our vulnerabilities may lead to better solutions that actually serve to solve problems that matter to people, rather than to the policy agendas of specific government departments. C

[6] See www.ips.gov.uk/identity/downloads/national-identity-scheme-delivery-2008.pdf.

## References


1. Berghel, H. Identity theft, Social Security numbers, and the Web. *Commun. ACM 43,* 2 (Feb. 2000), 17–21.
2. Crosby, J. *Challenges and Opportunities in Identity Assurance.* HM Treasury, 2008; www.hm-treasury.gov.uk/media/6/7/identity_assurance060308.pdf.
3. Dunleavy, P., Margetts, H., Bastow, S., and Tinkler, J. *Digital Era Governance: IT Corporations, the State, and E-Government.* Oxford University Press, Oxford, 2006.
4. Garfinkel, S.L. Risks of social security numbers. *Commun. ACM 38,* 10 (Oct. 1995), 146.
5. Guizzo, E. Loser: Britain's identity crisis. *IEEE Spectrum* (2006).
6. LSE Identity Project. LSE Identity Project (Main Report), London School of Economics and Political Science 2005; identityproject.lse.ac.u.k./identityreport.pdf.
7. Owen, K., Keats, G., and Gill, M. *The Fight Against Identity Fraud: A Brief Study of the EU, the U.K., France, Germany and the Netherlands.* Perpetuity Research & Consultancy International (PRCI) Ltd 2006; www.perpetuity-group.com/prci/publications.html#euid.
8. Whitley, E.A., Hosein, I.R., Angell, I.O., and Davies, S. Reflections on the academic policy analysis process and the U.K. Identity Cards Scheme. *The Information Society 23,* 1 (2007), 51–58.



**Edgar A. Whitley** (e.a.whitley@lse.ac.uk) is a Reader in Information Systems in the Department of Management at the London School of Economics and Political Science, U.K. **Ian R. (Gus) Hosein** (i.hosein@lse.ac.uk) is a Visiting Senior Fellow in the Information Systems and Innovation Group, Department of Management, London School of Economics and Political Science, U.K.




DOI: 10.1145/1342327.1342344


### Earlham College

Adjunct Professor in Computer Science

Earlham College invites applications for a potential one year full-time visiting position beginning in August of 2008 (with the possibility of renewal for a second year). The successful candidate must hold a PhD in CS or a closely related discipline (possibly ABD or, in exceptional circumstances, have equivalent experience). See http://cs.earlham.edu/search for a complete position description.

### King Abdullah University of Science and Technology (KAUST)

Faculty Openings in Computer Science and Applied Mathematics

King Abdullah University of Science and Technology (KAUST) is being established in Saudi Arabia as an international graduate-level research university dedicated to inspiring a new age of scientific achievement tha t will benefit the region and the world. As an independent and merit-based institution and one of the best endowed universities in the world, KAUST intends to become a major new contributor to the global network of collaborative research. It will enable researchers from around the globe to work together to solve challenging scientific and technological problems. The admission of students, the appointment, promotion and retention of faculty and staff, and all the educational, administrative and other activities of the University shall be conducted on the basis of equality, without regard to race, color, religion or gender.

KAUST is located on the Red Sea at Thuwal (80km north of Jeddah). Opening in September 2009, KAUST welcomes exceptional researchers, faculty and students from around the world. To be competitive, KAUST will offer very attractive base salaries and a wide range of benefits. Further information about KAUST can be found at http://www.kaust.edu.sa/.

KAUST invites applications for faculty positions at all ranks (Assistant, Associate, Full) in Applied Mathematics (with domain applications in the modeling of biological, physical, engineering, and financial systems) and Computer Science, including areas such as Computational Mathematics, High-Performance Scientific Computing, Optimization, Computer Systems, Software Engineering, Algorithms and Computing Theory, Artificial Intelligence, Graphics, Databases, Human-Computer Interaction, Computer Vision and Perception, Robotics, and Bio-Informatics (this list is not exhaustive). KAUST is also interested in applicants doing research at the interface of Computer Science and Applied Mathematics with other science and engineering disciplines. High priority will be given to the overall originality and promise of the candidate's work rather than the candidate's subarea of specialization within Applied Mathematics and Computer Science.

An earned Ph.D. in Applied Mathematics, Computer Science, Computational Mathematics, Computational Science and Engineering, or a related field, evidence of the ability to pursue a program of research, and a strong commitment to graduate teaching are required. A successful candidate will be expected to teach courses at the graduate level and to build and lead a team of graduate students in Master's and Ph.D. research.

Applications should include a curriculum vita, brief statements of research and teaching interests, and the names of at least 3 references for an Assistant Professor position, 6 references for an Associate Professor position, and 9 references for a Full Professor position. Candidates are requested to ask references to send their letters directly to the search committee. Applications and letters should be sent via electronic mail to kaust-search@cs.stanford.edu. The review of applications will begin immediately, and applicants are strongly encouraged to submit applications as soon as possible; however, applications will continue to be accepted until December 2009, or all 10 available positions have been filled.



PAUL WATSON

# Career Opportunities

In 2008 and 2009, as part of an Academic Excellence Alliance agreement between KAUST and Stanford University, the KAUST faculty search will be conducted by a committee consisting of professors from the Computer Science Department and the Institute of Computational and Mathematical Engineering at Stanford University. This committee will select the top applicants and nominate them for faculty positions at KAUST. However, KAUST will be responsible for actual recruiting decisions, appointment offers, and explanations of employment benefits. The recruited faculty will be employed by KAUST, not by Stanford. Faculty members in Applied Mathematics and Computer Science recruited by KAUST before September 2009 will be hosted at Stanford University as Visiting Fellows until KAUST opens in September 2009. At Stanford, these Visiting Fellows will conduct research with Stanford faculty and will occasionally teach courses.

### Polytechnic University
Professor

Polytechnic University invites applications for a faculty who will take on a leadership role in the area of cyber security. The position requires a strong track record of high impact funded research, along with the ability to create innovative partnerships that bring together academia, industry and government. Because Polytechnic University and New York University are in an advanced stage of merger proceedings, the successful candidate will have excellent opportunities to initiate interdisciplinary research and educational collaborations with the diverse institutes and departments of NYU.

Polytechnic is an NSA Center of Excellence in Information Assurance Education and has received two rounds of funding in the Scholarship For Service (SFS) program. Over a dozen security courses are offered regularly and an on-line graduate level cyber security certificate program is also available. Current research focus of the program at the MS and PhD level is on trusted hardware, trusted software systems, digital forensics, multimedia security, biometrics, application security, network security, etc. The ideal candidate would help us expand in some of these areas as well as expand to new areas of expertise in cyber security.

The Computer and Information Science Department (CIS) of Polytechnic University has a strong faculty with a vibrant research program and strong course offerings in a wide area of computing. Please submit a CV, Research Statement and the names of three references to:

Professor Stuart Steele
Cyber Search Committee
Polytechnic University
Six MetroTech Center
Brooklyn, NY, 11201

or by e-mail to
securitysearch@poly.edu

Polytechnic is an Equal Opportunity Employer

Polytechnic University,
Brooklyn, NY 11201

http://www.poly.edu/cis/

### Polytechnic University
Professor

Polytechnic University seeks a senior faculty member who will take a leading role in the general area of data management, including databases, data mining, information retrieval, web search and mining, and closely related areas. The position requires a strong track record of high impact publications and funded research, along with the ability to create innovative partnerships that bring together academia, industry, and government. Exceptional candidates at lower ranks will be considered.

Polytechnic University and New York University are in an advanced stage of merger proceedings, so the successful candidate will have excellent opportunities to initiate interdisciplinary research and educational collaborations with the diverse institutes and departments of NYU. Polytechnic is located in downtown Brooklyn, minutes from New York City's financial district and with easy access to New York's wide array of cultural and educational institutions.

The Computer and Information Science Department (CIS) of Polytechnic University has a strong faculty with a vibrant research program. It offers BS, MS, and PhD degrees. The ideal candidate will work in an area that complements the department's existing strength in web search technology, P2P content distribution, or information security.

Polytechnic University is an Equal Opportunity Employer

Please submit a CV, Research Statement and the names of three references by e-mail to
cissearch@poly.edu

or hard copy to

Chair, Faculty Search Committee
(Data Management)
Polytechnic University
6 Metrotech Center
Brooklyn, NY, 11201 USA

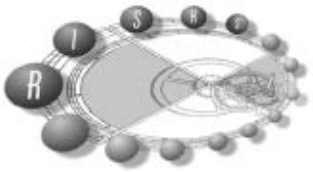
### Sam Houston State University
Department of Computer Science

The Department of Computer Science in the College of Arts and Science at Sam Houston State University announces the availability of

tenure track positions in Computer Science, Information Assurance and Computer/ Technology Education for Fall 2008.

Candidates should have a terminal degree in the appropriate field. The successful candidate will be expected to teach undergraduate and graduate levels and pursue an active research program. Candidate materials and other information are available via email to csc_pac@shsu.edu.

Sam Houston State University is an EEO/AAP employer. For additional information on this and other positions, access our website at: www.shsdu.edu/hr.





# The Physical World and the Real World

Most of us rely on the Internet for news, entertainment, research, communication with our families, friends, and colleagues, and myriad other purposes. What if it went away?

Precisely that happened to many people in early February, in the wake of the failure of several undersea cables. According to some reports, more than 80 million users were affected by the outages. Both the initial failure and the subsequent recovery have lessons to teach us.

The first lesson, of course, is that failures happen. In fact, multiple failures can happen. Simply having some redundancy may not be sufficient; one needs to have enough redundancy, and of the right types. In this case, geography and politics made life more difficult.

The geographical issue is obvious when viewing the region on a map: there aren't many good choices for an all-water route between Europe and the Persian Gulf or India. And despite this series of events, cables are generally thought to be safer on the seabed than on land. (There is a standing joke in the network operator community, the essence of which is that you should bring a length of fiber-optic cable with you when going hiking in the wilderness. If you get lost, throw it on the ground. A backhoe will soon show up to sever it; ask the driver how to get home.)

The obvious answer is to run some backup cables on land, bypassing the chokepoint of the Red Sea. Again, a glance at the map shows how few choices there are. Bypassing the Red Sea on the west would require routing through very unstable countries. An eastern bypass would require cooperation from mutually hostile countries. Neither choice is attractive.

From this perspective, it doesn't matter much just why the cables failed. Cables can be cut by ship anchors, fishing trawlers, earthquakes, hostile action, even shark bites. Regardless of the cause, when so many cables are in such a small area, the failure modes are no longer independent.

*For this problem, there are no good solutions.* Anyone whose business depends on Internet connectivity through this region must take this into account.

PAUL WATSON

The dangers aren't only physical, as several recent incidents will attest. The last few months have also shown that a 1999 National Research Council report was quite correct when it warned of the fragility of the routing system and the domain name system used for the Internet.

In one highly publicized incident, a routing mistake by a Pakistani Internet service provider knocked YouTube off the air. There was a lot of speculation that this was deliberate—the government of Pakistan had ordered YouTube banned within the country; might someone have tried to "ban" it globally?—although later analysis strongly suggests that it was an innocent mistake. An outage affecting such a popular site is very noticeable; there was a great deal of press coverage. By contrast, when a Kenyan network was inadvertently hijacked by an American Internet service provider, there was virtually no notice. Quieter, deliberate misrouting—say, to eavesdrop on traffic to or from a small site—might go completely unnoticed.

The DNS-related incidents are scarier because they do reflect deliberate actions, with the force of the U.S. legal system behind them. In one case, the Wikileaks.org Web site was briefly deleted from the DNS by court order, because a bank claimed the site contained stolen documents. (The site owners had apparently foreseen something like that, and had registered other names for the site in other countries: the .org registry is located in the U.S.) In a second incident, a U.S. government agency ordered the names of some non-U.S. sites removed from .com (again, located in the U.S.) because they violated the embargo against Cuba.

What can we learn from these incidents? The moral is simple: the Internet is a lot more fragile than it appears. Most of the time, it works—and works very well—without government interference, routing mistakes, or outages due to occasional fiber cuts. Sometimes, though, things go badly wrong. Prudence dictates that we plan for such instances. C

**Steven M. Bellovin** (smb@cs.columbia.edu) is a professor of computer science at Columbia University.

DOI: 10.1145/1342327.1342345

ifip

IFIP
WCC 2008
WORLD COMPUTER CONGRESS
Milano

AICA
Associazione Italiana per l'Informatica ed il Calcolo Automatico

# 20TH WORLD COMPUTER CONGRESS

## Milano (Italy), September 7-10, 2008

MIC - fieramilanocity

OPENING SESSION: GLOBALIZATION IN THE DIGITAL SCENARIO

HOSTED CONFERENCES: WOMEN AND TECHNOLOGY
*e*-INCLUSION
THE AICA'46TH NATIONAL CONGRESS

### TECHNICAL CONFERENCES:

- Biologically Inspired Cooperative Computing - BICC [TC10]
- Distributed and Parallel Embedded Systems - DPES [WG10.2]
- 1st IFIP Entertainment Computing Symposium - ECS [TC14]
- Learning to live in the knowledge Societies - ED_L2L [TC3]
- History of Computing and Education - HCE3 [TC9, WG9.7, TC3]
- Human Computer Interaction - HCI [TC13]
- IFIP Artificial Intelligence 2008 -IFIP AI [TC12]
- 23rd IFIP International Information Security Conference - IFIP SEC [TC11]
- Advances in Information Systems Research, Education, and Practice - ISREP [TC8]
- Knowledge Management in Action - KMIA [WG12.6]
- Open Source Systems 2008 - OSS [TC2, WG2.13]
- 5th IFIP International Conference on Theoretical Computer Science - TCS [TC1]
- Topical Session on CAI (Computer Aided Innovation) [WG5.4]

### INDUSTRY CONFERENCES:

- ICT for Sport & Fitness
- ICT for Tourism
- R&D for Textile and Fashion
- ICT for Environment
- ICT for Cultural Heritage
- Service Science
- e-Government
- ICT Professionalism & Competences
- e-Inclusion
- ICT for Health
- Web 2.0
- ICT for Education
- ICT Innovation in Finance
- Intelligent Hospital
- Smart Grids
- Intelligent Building
- Development Governance

**Information and registration: www.wcc2008.org**

in cooperation with: INTERNATIONAL TECHNOLOGY GROUP

The 1st ACM SIGGRAPH Conference and Exhibition in Asia
www.siggraph.org/asia2008

# New Horizons

## ACM SIGGRAPH launches the premiere SIGGRAPH Asia in Singapore

Programs include:

- Papers
- Sketches and Posters
- Courses
- Art Gallery
- Computer Animation Festival
- Educators Program
- Emerging Technologies
- Exhibition

***Calling all creative researchers, artists, digital innovators and animators!***

This is your opportunity to present your stellar work or attend the 1st ACM SIGGRAPH conference and Exhibition in Asia.

***Queries?***

Contact Conference and Exhibitions Management at asia2008@siggraph.org or Tel: +65 6500 6700

**SIGGRAPHASIA2008**

Conference and Exhibition on Computer Graphics and Interactive Techniques

**Singapore, 10-13 December 2008**

Held in

UNIQUELY Singapore
visitsingapore.com