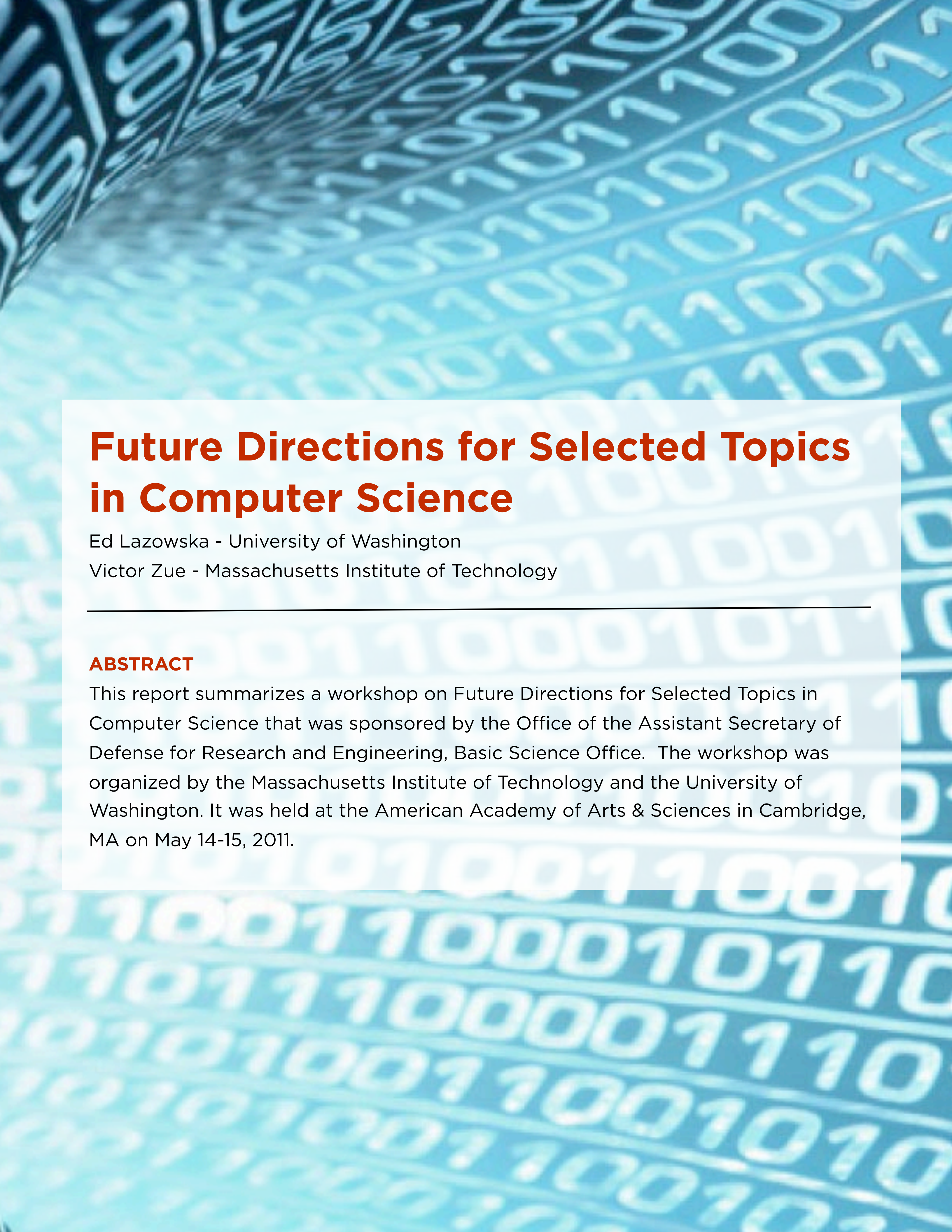

Future Directions for Selected Topics in Computer Science

Ed Lazowska - University of Washington
Victor Zue - Massachusetts Institute of Technology

ABSTRACT

This report summarizes a workshop on Future Directions for Selected Topics in Computer Science that was sponsored by the Office of the Assistant Secretary of Defense for Research and Engineering, Basic Science Office. The workshop was organized by the Massachusetts Institute of Technology and the University of Washington. It was held at the American Academy of Arts & Sciences in Cambridge, MA on May 14-15, 2011.

Summary of the ASDRE Workshop on Future Directions for Selected Topics in Computer Science

Cambridge MA, May 14-15 2011

Ed Lazowska
University of Washington

Victor Zue
Massachusetts Institute of Technology

Table of Contents

1. Executive Summary

A two-day workshop, sponsored by the ASDR&E and organized by MIT and the University of Washington, occurred on May 14-15, 2011 at the American Academy of Arts & Sciences in Cambridge, MA. The goal of the workshop was to provide perspectives on barriers to advancement and potential breakthroughs in the growing and rapidly evolving field of computer science. Twenty-three academic and industrial professionals from twelve institutions, covering diverse research areas in artificial intelligence, systems, and theory, attended the workshop, as did six government participants serving as observers.

During the morning of Day One, participants identified a number of game-changing advances in the field of computer science over the past decade. Areas of discussion include:

- *Parallel Computing:* including many-core processors that will bring significant improvement in computational performance and energy efficiency,
- *Cloud Computing:* including scalable systems, virtualization, and dependable/ manageable systems,
- *Ubiquitous Wireless Networks and Mobile Devices:* enabling high-bandwidth creation, transmission, and access of massive amounts of multimedia data, regardless of location,
- *Big Data Computing:* systems and algorithms that can deal with massive amounts of heterogeneous data, turning data into decisions,
- *Open Source Software Development:* becoming mainstream and enabling collaborative development,
- *Machine Learning:* improved algorithms and tools, and wide-spread and "at scale" use in a myriad of applications,
- *Human-Machine Interactions:* particularly large vocabulary speech recognition and machine translation,
- *Robotics:* robots that can operate in unstructured environments (planning and visualization), as exemplified by the DARPA Grand Challenge and Urban Challenge,
- *Social Networking and Crowd-Sourcing:* of increasing importance for intelligence gathering, as demonstrated in the DARPA Balloon Challenge, and
- *Homomorphic Encryption:* enabling, for example, secure voting, and the use of cloud computing by ensuring the confidentiality of processed data.

The participants spent the remainder of the workshop formulating a set of research thrusts and scenarios that will achieve major impact for the DoD, and at the same time can serve to coalesce further advances of the underlying technologies. Some challenge problems that can demonstrate progress are also identified and discussed. These include:

- *Complex Social-Technical Systems.* We must build systems that are capable of sensing, decision-making, and acting in the context of varied social environments, must be able to collect and handle vast amounts of data, while mitigating concomitant security and privacy concerns.
- *Embodied/Autonomous Systems.* We must progress from today's world of robots under human control to a future in which autonomous robots perform alongside humans, augmenting their capabilities and effectiveness.
- *From Data to Decisions.* We must continue to develop new methods that will enable computers to use data to learn, infer and predict information with ever-greater fidelity and breadth of application. Societal trends in social networking and crowdsourcing should be exploited to engage people in the process of generating, filtering, and disseminating data as well.
- *Human-Computer Symbiosis.* We must learn to develop systems that combine the abilities of humans and computers to yield dramatically improved capabilities. Future symbiotic systems must combine aspects of humans assisting in computation and computers augmenting human function.
- *Security and Privacy.* Security – including computer security and information security, is critical in the DoD context. Future research must revolve around reactive systems that can automatically recover from attacks, and distribute and apply security patches.
- *Computing Technologies*. DoD requires significant advances in energy-efficient computing and wireless spectrum efficiency. Aggressive exploration of quantum computing is also essential.

All contributions, notes, and other remnants of the workshop are kept at a password-protected web site hosted at the MIT Computer Science and Artificial Intelligence Laboratory.

2. Preamble

The Department of Defense has been a primary sponsor of computing research since the inception of the field, and the warfighter has been a primary beneficiary of that research. America's national security is inextricably linked to advances in computing, and to our Nation's continued leadership in the field.

Computing research is important to our national defense. This is evident in the summary of the Quadrennial Defense Review to the President's Council of Advisors on Science and Technology in November 2010, which listed eight "critical capabilities" for the Department of Defense going forward. The first four of those eight capabilities were squarely the province of computing research: they were *decision support*, *autonomous systems*, *trusted systems / trusted cyber-physical systems*, and *immersive training*.

A two-day workshop, sponsored by the ASDR&E and organized by MIT and the University of Washington, occurred on May 14-15, 2011 at the American Academy of Arts & Sciences in Cambridge, MA. The goal of the workshop was to provide perspectives on barriers to advancement and potential breakthroughs in the growing and rapidly evolving field of computer science. Twenty-three academic and industrial professionals from twelve institutions, covering diverse research areas in artificial intelligence, systems, and theory, attended the workshop, as did six government participants serving as observers. Section 5 provides a list of the attendees.

During the morning of Day One, participants identified recent game-changing advances from their subfields, and suggested future topics of potential great impact. For the remainder of the meeting, we aggregated these topics into broad areas of future importance, and broke out into smaller groups to elaborate, refine, and specify research directions in these areas. All contributions, notes, and other remnants of the workshop are kept at a password-protected web site hosted at the MIT Computer Science and Artificial Intelligence Laboratory.

3. Technological Advances and Challenges

Participants identified a number of game-changing advances in the field of computer science over the past decade. In each case, technical challenges facing us in the coming decade were also described, as summarized below.

3.1 Parallel Computing

From the study of climate change to the mapping of the human genome, from consumer devices such as smartphones to social networking supported by cloud computing, breakthroughs in computing performance have transformed disruptively what is possible. But serial computing performance is at a virtual standstill due to power and complexity limits of single processors, so that all increases in computing performance will have to come from parallelism.

The computing industry successfully leveraged decades of research in parallel computing hardware to make a generational shift to multicore processor chips, which can take advantage of multi-tasking between jobs and virtualization in servers to provide high throughput. However, staggering challenges threaten further progress. The major challenges include the power efficiency challenge, which leads to massively parallel manycore chips,[1] and the resultant programming and algorithmic challenges. The gap between processor and memory performance continues to grow, along with the gap in energy use between computing on data and moving it through the system. Chip density and system scale raise resilience challenges, adding to the correctness problems that arise in complex parallel applications. Research will be required in all areas of computing including architecture, software systems and algorithms in order to overcome these challenges. Examples of research include architectures and algorithms for data movement avoidance and optimization, self-aware computational models that are able to optimize for power, resilience and programmability, and high bandwidth and low power memory systems and interconnect.

3.2 Cloud Computing

Over the past decade, many applications have moved from local computers and servers to data centers, and have scaled to use tens of thousands of machines each.[2] This has required significant advances in distributed systems, hardware design, management tools to administer systems at such large scale, and algorithms. There are several upcoming challenges for the next decade. First, programming such large-scale applications is a challenge, especially as applications are spanning across data centers, and sharing

---

[1] A. Agarwal and M. Levy, "Going Multicore Presents Challenges and Opportunities," *Embedded Systems Design*, Vol. 20, No. 4, April 2007

[2] M. Berezecki, E. Frachtenberg, M. Paleczny, and K. Steele, "Many-core key-value store." *International Green Computing Conference (IGCC),* July 2011

data.[3,4,5] Second, as increasingly more sensitive data is being processed in data centers, and as data from many users is being aggregated together in new ways, security and privacy are becoming even more important. Supporting these applications is going to require high-performance and low-energy networking and hardware designs, which will be a challenge in itself as well.

3.3 Ubiquitous Wireless Networks and Mobile Devices

The wireless networking and mobile computing fields have seen unprecedented growth over the past decade. This trend will accelerate over the next decade. The last few years of research have shown that major reliability and throughput gains can be achieved in many cases with integrated inter-disciplinary *cross-layer* designs that cut across the traditional protocol stack. However, the wireless industry is generally divided according to the layers of network stack. Closer partnerships between the research community and industrial players are essential. The result will be new markets in cross-layer products that offer substantial benefits over the status quo and enable new mobile applications. Major technical challenges include:

- *Spectrum Utilization.* Research is required to enable communication-intensive applications across large numbers of wireless and mobile devices. These solutions include advanced interference management techniques, better coding and transmission schemes that increase spectral efficiency (bits/s/Hz), novel approaches to dynamic spectrum access, scalable protocols for heterogeneous radio networks, protocols to handle "truly mobile" networks and devices (smartphones and handhelds), and novel protocol and software architectures, particularly *cross-layer* approaches that work across the traditional network layers.
- *Mobile Applications.* Research on distributed application designs that work across mobile devices and compute/storage clouds is important because that model is emerging as a common one. Unfortunately, it is a complicated programming and development model. Hence, research in new programming models and abstractions for "mobile-meets-cloud" applications is important.
- *Security.* Key topics include significantly simplifying the security configurations of wireless devices, and improving the security of embedded wireless systems. Promising directions include combining physical layer techniques and modern cryptographic protocols in real systems.
- *Energy.* Energy continues to remain a scarce resource in the mobile world. Continued research in hardware and software designs for energy-efficient mobile systems, spanning low-power handsets, energy-efficient radios, and energy-scavenging sensors, is important.

---

[3] P. Alvaro, N. Conway, J. Hellerstein, and W. Marczak, "Consistency Analysis in Bloom: a CALM and Collected Approach," *Proc. Conf. innovative Data Systems Rsrch (CIDR)*, January 2011

[4] S. Bykov, A. Geller, G. Kliot, J. Larus, R. Pandya, and J. Thelin, "Orleans: Cloud Computing for Everyone," *ACM Symposium on Cloud Computing (SOCC 2011),* October 2011

[5] J. Field, M. Marinescu, and C. Stefansen, "Reactors: A Data-Oriented Synchronous/Asynchronous Programming Model for Distributed Applications," *Theor. Comput. Sci.*, 410(2-3): 168-201, 2009

3.4 Big Data Computing

The recent decade has seen several technological advances that create a tremendous area of opportunity in the area of Big Data.[6] The opportunity arises out of a confluence of a number of factors. First, there has been massive spike in the quantity of data that is available, driven by the Internet and sensors. Second, there is a perceived need for being "data-driven" in a number of fields, including business, medicine, science, and national security. Third, technologies (like MapReduce, and column-oriented relational databases) to store and process this data in parallel on massive clusters of commodity hardware have become mature.[7] Fourth, the rise of statistical machine learning has led to the wide use of sophisticated algorithms to make sense out of this data by a variety of analysts in a variety of fields.

Taken together, these advantages have the potential to revolutionize many fields (see "From Data to Decisions" in Section 4.3) by allowing sophisticated new techniques to be applied at scale to new, massive feeds of data from a variety of different heterogeneous and widely distributed data sources. The challenge is that the technological advances listed above have been driven by different sub-communities, and are not integrated or packaged in a useable way. In many cases, it is not obvious how these technologies even integrate, and new advances are needed to improve algorithms to achieve the level integration needed achieve their aggregated potential. Additionally, decision-making tools are in great need of better provenance features, so that users can understand why a particular conclusion has been reached; conflicts of "ground truth" and unexpected inputs are inevitable, and surfacing these to users in a sensible way is important if we are to integrate Big Data systems into decision making environments.

3.5 Open Source Software Development

In the past decade, open source software has gone from a hobbyist niche to mainstream infrastructure. Many commercial products, from servers to embedded systems, rely on open-source software such as Linux, MySQL, Postgres, Apache, etc. One surprising result has been that a large number of developers have been able to collaborate over the Internet to produce high-quality software. Open-source software development presents a great opportunity for research to influence future systems, by making it easy to contribute new ideas and code. At the same time, this raises important issues related to the management of the "software supply chain" in which a large number of domestic and foreign players will be involved.

---

[6] D. DeWitt and J. Gray, "Parallel database systems: the future of high performance database systems," *Communications of the ACM*, 1992

[7] J. Cohen, B. Dolan, M. Dunlap, J. Hellerstein, and C. Welton, "MAD Skills: New Analysis Practices for Big Data, *Proc. VLDB*, 2009

3.6 Machine Learning

Machine learning techniques have been one of the great successes of computer science in the past decade. Specific technical advances have included: (a) machine learning formalisms for sequential,[8] relational, and structured data including Conditional Random Fields,[9] Markov Logic Networks, Probabilistic Relational Models,[10] and structured Support Vector Machines, (b) non-parametric probabilistic models based on hierarchical Dirichlet processes,[11] Kingman's coalescent, and related structures, (c) non-parametric probabilistic models based on boosted regression trees,[12] (d) methods for dealing with large amounts of unlabeled data including semi-supervised learning,[13] active learning, and transfer learning, and (e) matrix-based learning methods for collaborative filtering,[14] multi-task learning,[15] and coupled feature selection.[16] Machine Learning methods have been applied widely and at large scale on problems such as object recognition, speech recognition, natural language understanding, mining Internet traffic, optimizing Internet advertising, email and web spam detection, network intrusion detection, dynamic customization of eCommerce web sites, adaptive data structures in operating systems, and so on.

There are many challenges for the future: (a) transforming existing algorithms and developing novel algorithms that support ultra-large-scale machine learning on parallel and multi-core hardware, (b) new algorithms that can jointly learn thousands or millions of related "concepts" (e.g., learning to recognize all of the objects in the Amazon product catalog, learning to recognize all of the world's plants and animals), including methods for rapidly learning a new concept from a handful of examples by leveraging thousands of previously-learned concepts, (c) algorithms for learning under non-stationary and adversarial conditions or from data collected from diverse sources of highly-variable quality, (d) methods for overcoming/automating the process of representation design and feature engineering for machine learning systems to make it easier for non-experts to create machine learning applications, and (e) integrated development environments and software engineering tools to support the design, debugging, and maintenance of deployed machine learning systems so that machine learning software can be integrated more easily with traditional software.

---

[8] P. Domingos and M. Richardson, "Markov Logic Networks," *Machine Learning*, Vol. 62, 107-136, 2006

[9] J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," *Proc.s of the 18th International Conference on Machine Learning*, 282-289, 2001

[10] D. Koller, D., "Probabilistic relational models", *Inductive Logic Programming*, 3--13, 1999

[11] Y. W. Teh, M. I. Jordan, M. J. Beal and D. M. Blei, "Hierarchical Dirichlet Processes," *Journal of the American Statistical Association*, 101, 1566-1581, 2006

[12] J. Friedman, "Greedy Function Approximation: {A} Gradient Boosting Machine." *Annals of Statistics*, Vol. 29, 1180, 2001

[13] X. Zhu, X., Goldberg, A., "Introduction to Semi-Supervised Learning," *Synthesis Lectures on Artificial Intelligence and Machine Learning,* 3, 1-130, 2009

[14] J. Rennie and N. Srebro, "Fast maximum margin matrix factorization for collaborative prediction," *Proceedings of the 22nd international conference on Machine learning*, 713--719, 2005

[15] R. Caruana, "Multitask learning," *Machine Learning*, 28, 1, 41-75, 1997

[16] R. Bell, Y. Koren, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *IEEE Computer,* 42-49, 2009

3.7 Human-Machine Interactions

We have witnessed remarkable progress in areas such as speech recognition, machine translation,[17,18] information extraction and summarization from human language,[19] and human-machine interfaces based on human language. The success is brought on largely by the widespread use of statistical modeling techniques such as Hidden Markov Modeling, and the availability of large corpora to help the system learn the model parameters. Today, our lives are constantly touched by systems that can allow us to dial phone numbers, issue verbal commands, perform transactions, or perhaps dictate a letter, all using human language.

But we are a long way from developing systems that will enable people speaking different languages to communicate freely among themselves and with machines. To get there, we will need to broaden the underlying human language technologies from recognition to understanding, to include discourse and dialogue modeling, and to incorporate multiple modalities. These systems must also be able to exploit contextual knowledge, offer opinions as well as facts, learn from mistakes, and heal themselves when facing catastrophic failures. The need to develop *organic* computing capabilities is a recurring theme in several areas of computing research.

3.8 Robotics

The military has made significant investments into developing light combat ships (LCS), high altitude long endurance (HALE) vehicles, tactical UAVs, and unmanned underwater vehicles (UUVs) for maritime surveillance. However, despite today's dynamic battlespace, human operators are consumed by the direct control, monitoring and supervision of autonomous vehicles, constantly "tweaking'" the system to obtain the required information. It is necessary to rapidly allocate and re-task different assets to support time-critical intelligence needs for persistent surveillance, and to respond to the detection of threats or changes in missions. To leverage the military's investments in the unmanned vehicle space, it is necessary to make dramatic breakthroughs in decentralized planning and control, perception, autonomy and human interaction. The necessary breakthroughs in computational intelligence require combined expertise in artificial intelligence, control theory, formal methods, machine learning, robotics, and perception, to create a new science base for computational intelligence that will enable intelligent individual agents to operate in complex environment to interpret and reason about these

---

[17] P. Koehn, *Statistical Machine Translation*, Cambridge University Press, 2010

[18] S. Ravi and K. Knight, "Deciphering Foreign Language," *Proc. of the Annual Meeting of the Association for Computational Linguistics*, 2011

[19] G. Tur and R. De Mori, *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*, John Wiley & Sons, 2011

environments, and groups of agents that can coordinate with other agents, and with humans, to achieve tasks of larger scope than single agent tasks.[20]

3.9 Social Networking and Crowd-Sourcing

By connecting billions of humans, the Internet has given rise to fundamentally new capabilities. Over the past decade, we have seen many instances where human collaboration has happened at a fundamentally new scale and speed.[21] For example, search engines collect knowledge generated by millions of people creating and linking web pages and then use this knowledge to answer questions instantaneously. In Wikipedia, thousands of people around the world have collectively created a massive and high quality encyclopedia with almost no centralized control. These instances are only the beginning of a fundamental shift in which we, as humans, interact with each other – a shift that is likely to revolutionize what we can accomplish.[22]

Research in this area is focusing on how to use these new capabilities to solve problems at a massive scale. Some of the questions include: How do we encourage people to participate and contribute?[23] How do we split up complex tasks so that they can be solved by combining millions of small human contributions?[24] How do we route the sub-tasks to the appropriate individuals? How can we ensure the correctness of the outcome?[25]

Being able to coordinate millions of people through the Internet could have applications to national security, health care and government. If 400,000 people put a man on the moon, what can we do with 100 million?

3.10 Homomorphic Encryption

Ever since its beginning in the modern era, cryptographers have wondered if encrypted data can be manipulated in desired ways.[26] Note this is not always a goal. Sometimes, e.g, in implementing sealed bid auctions, it is desired that no one in possession of encrypted data should be able to produce encryptions of related data without having the power (the secrets needed) to decrypt. Such an encryption scheme is termed “non-malleable” and various cryptosystems have been proposed over the years with this

---

[20] “50 Years of Robotics,” Special Issue of the *IEEE Robotics and Automation Magazine*, Vol. 17, No. 3, September 2010

[21] L. von Ahn, B. Maurer, C. McMillen, D. Abraham and M. Blum, “reCAPTCHA: Human-Based Character Recognition via Web Security Measures.” *Science*, 1465-1468, September 12, 2008

[22] G. Pickard, W. Pan, I. Rahwan, M. Cebrian, R. Crane, A. Madan, A. Pentland, “Time-Critical Social Mobilization,” *Science*, 509-512, October 28 2011

[23] L. von Ahn and L. Dabbish. “General Techniques for Designing Games with a Purpose,” *Communications of the ACM*, 58-67, August 2008

[24] G. Little, L. Chilton, M. Goldman, and R. Miller, “Exploring Iterative and Parallel Human Computation Processes,” *Second Annual Human Computation Workshop*, HCOMP 2010

[25] V. Ambati, S.Vogel, and Jaime Carbonell, “Towards Task Recommendation in Micro-Task Markets,” *Third Annual Human Computation Workshop*, HCOMP 2011

[26] R. Rivest and L. Adleman and M. Dertouzos, "On Data Banks and Privacy Homomorphisms,” in *Foundations of Secure Computation*, ed. R. DeMillo, D. Dobkin, A. Jones, and R. Lipton, Academic Press, New York, 169-180, 1978

property. But the other extreme, where every possible manipulation of the data can be carried out under the seal of encryption, while still learning nothing about the data itself, had so far been out of reach. A recent breakthrough[27] devised such a scheme, building on some very clever but intricate ideas. The feasibility of such fully homomorphic encryption opens the door to a vast class of privacy-enabled data analysis. It is in principle possible for vast amounts of data containing information about individuals to be released by the people owning the data, with each piece of the data being encrypted separately by the owners, and then an aggregating/analysis algorithm could process all this data under the seal of encryption and get an encrypted version of the conclusions which could then be decrypted and released.

The possibilities of such schemes are vast. However, much work is required to move this research into use. One issue concerns its inherent security. We know now that homomorphic encryption is possible, but the initially proposed scheme does not conform to classical crypto-assumptions and there is less confidence in its robustness. A second issue concerns performance. Current homomorphic schemes are much slower than classical encryption, and computation on the encrypted data is much slower as well. But as the research continues in this direction, this status is sure to change.

---

[27] G. Gentry, “Fully homomorphic encryption using ideal lattices,” *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)*, 169-178, 2009

4. Broad Areas of Future Importance

The participants spent the remainder of the workshop formulating a set of research thrusts and scenarios that will achieve major impact for the DoD, and at the same time can serve to coalesce further advances of the underlying technologies. Some challenge problems that can demonstrate progress are also identified and discussed.

4.1 Complex Social-Technical Systems

Modern society is composed of a number of large and complex socio-technical systems – healthcare, crisis response, energy, and transportation – that involve experts, amateurs, and computers working together. These systems are sometimes long-lived national infrastructure projects and sometimes ad-hoc emergency deployments. In both cases, there is great need to improve our collective responsiveness and capability by developing technological solutions targeted at social systems.

We must build systems that are capable of sensing, decision-making, and acting in the context of varied social environments. These systems must be able to collect and handle vast amounts of data, while mitigating concomitant security and privacy concerns. They must support a variety of decision structures, from decentralized operations to local government. The wealth of predictive modeling capabilities developed over the past few decades must be integrated into these systems so that non-specialists may better make data-driven decisions.

Challenge problems include:

- *Eliminate search from search and rescue.* For example, know the location of any individual in need, authenticated crisis alerting.
- *Create a rapid-start kit for social-tech systems,* including broadcast, logistics support, a person finder, data analysis, and data integration tools out of the box.
- *Model large populations with small number of people.* Enable what-if simulators or games that enable experimentation and synthetic study of socio-technical systems.
- *Opportunistically adapt emerging consumer tools.* Make better use of social networks; harness user communities and technologies; and develop an ongoing research program for adaptation of each development.
- *Facilitate collaboration between domain experts and lay people.* Build systems which prioritize limited interventions, can translate knowledge from expert jargon into common speech, and can provide cultural awareness for domain experts.
- *Micro-targeting for social goals.* Develop strategies and technologies for communicating/influencing targeted sub-populations – e.g., those in compliance with their medical plans, those needing to evacuate their homes in a disaster
- *Develop effective and appropriate instrumentation of humans* for personal health alerts and situational crisis alerts.

4.2 Embodied/Autonomous Systems

Robots can give us eyes where we cannot see, hands where we cannot reach, strength where we cannot function, and productivity where we have the need. We must progress from today's world of robots under human control to a future in which autonomous robots perform alongside humans, augmenting their capabilities and effectiveness.

The technology that will enable robots to act *with* humans, rather than simply for them, presents a number of significant challenges. Improvements in multimodal sensing and situational understanding will be required for safety when machines comingle with humans in densely populated environments (such as a sidewalk or store). In addition to safety, advancements in "common sense processing" are required for embodied machines to behave naturally in human environments such that their presence is accepted.

The ability to formulate and revise plans is another broad area rich with challenges. Interpreting high-level natural language objectives and translating these into low-level representations and control signals is an unsolved area with much potential payoff. Revising and evaluating multiple such plans is a natural capability whose development would follow. The ability to translate these plans into multi-agent actions (i.e., plan to enact a goal with other robots or humans) is another challenge area. It is one thing for a robot to pick up a large object and move it, but quite a more sophisticated thing for that robot to pick up that object with a human and move it together: such a cooperative task requires complex spatial, linguistic, and control reasoning.

Challenge problems include:

- *Demonstrate robotic mastery in aquatic environments.* De-fuse a dirty bomb in deep sea. Or create a controllable underwater robot that can explore a wide variety of environments.
- *Urban-scale mobility on demand.* Create a service such as ZipCar that can drive itself and interact with clients using natural language.
- *Construction work alongside humans.* Assemble a command post with a team of robots and humans to supervise them.
- *Remote sensing.* Develop a robot "Avatar" with all senses.
- *Match or exceed nature's best controllers.* For example, build a robot that can move through a real environment as a dragonfly, bird, or frog does.
- *Learn to learn from mistakes.* Develop life-long (persistent) performance – construct robots that are able to translate their experiences into improvements in operation over time.

Tackling these challenges will require fundamental advances in areas such as:

- *Physical intelligence* (perception, control).
- *Social intelligence* (understanding intent, human-robot coordination, robot-robot coordination, co-learning).
- *True multi-sensory telepresence.*

4.3 From Data to Decisions

Data has taken center stage in recent technological and sociological trends. *Wired Magazine* recently described modern science as entering "The Petabyte Age." *Fortune Magazine* highlighted "Data Science" as the emerging hot job in today's economy. *The Economist* ran a cover article entitled "The Data Deluge" in which they surveyed the effects of data on society more broadly, referencing an "Industrial Revolution of Data." The importance of "Big Data" has entered common culture.

Within computing itself, Big Data has also taken center stage across the field. Scalable data analysis has become a critical theme for a wide range of computing techniques, from cyber-security to natural language processing to datacenter management. New methods from machine learning and computational statistics have enabled computers to use data to learn, infer and predict information with ever-greater fidelity and breadth of application. Societal trends in social networking and crowdsourcing have engaged people in the process of generating, filtering, and disseminating data as well.

This diversity of techniques and applications differentiates today's Big Data challenges from traditional, business and web-oriented transactional data management problems that drove the development of relational database technology over the previous few decades. Important new requirements include:

- *Data acquisition and discovery.* Data come from many resources, not just databases, files, and webservers. It is often not obvious which places might be relevant to a given task, or how to ensure that relevant data is acquired in future.
- *Integration.* Data from multiple sources need to be meaningfully combined to produce composite answers. This remains a time-consuming and surprisingly manual task even for pre-structured data.
- *Extraction.* Raw source data, e.g., from ambient sensors, unstructured text, imagery, audio, natural language, and so on require feature and structure extraction before it can be processed.
- *Scale.* Data volumes and rates in many Big Data applications easily exceed tens of terabytes a day, while just a few years ago a terabyte of data was considered massive. Though tools have been developed for parallel computation over data, and processor core counts and storage capacities continue to grow, new techniques to cope with these volumes are needed.
- *Statistical/Logical views of data.* Increasingly, analysis involves imposing statistical models or approximation methods on the data (e.g., regression, classification, clustering, sketches, etc.), and then using those "views" of the data to support querying, inference or prediction via probabilistic methods. Such analyses need to leverage both traditional logical approaches to structured data, and statistical models and probabilistic methods.
- *Programmability.* Typical large data sets today span hundreds or thousands of computers. Using traditional programming models in this context, programmers are forced to specify methods to move computation to the data, and make sure the disparate computations coordinate and parallelize efficiently. Declarative and functional approaches to programming have helped here (e.g., SQL, MapReduce),

but today's languages still provide rather crude building blocks for parallel algorithms on core mathematical structures like graphs and matrices that underlie the powerful methods in science and statistics.

- *Visualization and interfaces.* As data complexity grows, tools to visualize and interact with data, to spot outliers and trends, and to highlight features are needed, and need to be better integrated into programmers' toolboxes.

Challenge problems include:

- *Expert finding.* Here the goal is to identify a person in an organization or community who has information about or is relevant to a particular topic, e.g., when critical infrastructure malfunctions, identify the people who can best assist repair—this may include subject-matter experts as well as (possibly former) employees or contractors. The challenge is to integrate data from multiple sources (e.g., public records, websites, emails, images, financial records, sensor data, etc.) to understand what a person knows, where they are, how to contact them, etc. This kind of problem appears in a number of domains, from homeland security to question-and-answer websites to corporate search.
- *Public Health outbreaks.* Google (via its flu trends project) has shown that mining users' search streams can be an effective tracker of the spread of infectious diseases like the flu, as users in different locations search for information about symptoms and treatment when they become sick. The "Big Data" generalization of this problem involves combining data from web searches, blogs (e.g., Twitter), drug sales records, and hospital and doctor's office reports to automate the discovery and spread of disease, and to differentiate known sicknesses from unknown ones. This can be generalized to tracking the spread of malware on the Internet, using a similar collection of data sets (e.g., downloads of system repair tools, sales of anti-virus software, Twitter, etc.).
- *Urban planning and the smart city.* Here the challenge is to create models and visualizations of activity in urban areas that help individual users, city planners, shop owners, and government to understand phenomena in the city. Examples include the flow and behavior of traffic, the use of public transit, the general flux of people through different areas of the city at different times of day (e.g., areas people avoid at night, or on weekends, or places that are seeing more than an expected amount of activity), levels of pollution, locations of crimes, etc. This is again enabled through a complex fusion of several vast and diverse data sets, including cellular location feeds, call logs, transit sales, embedded road sensors, police reports, tweets and blog posts, and so on.
- *Personalized medicine and genomics.* The dramatic drop in the cost of sequence an individual genome creates a compelling "Big Data" problem, especially when genomes are combined with each other and medical data about symptoms and treatment effectiveness. By comparing genomes of microbes and diseased patients with historical records of treatment outcomes (and the genomes of previously treated patients), a new kind of "data driven" approach to medicine is enabled. For example, it becomes possible to create targeted treatments specialized to both the exact variant of a given disease as well as the genetic predisposition of sick people to certain treatments or progressions of illness.

4.4 Human-Computer Symbiosis

There have been remarkable recent advances in the design of symbiotic human-computer systems: systems that combine the abilities of humans and computers to yield dramatically improved capabilities. These symbiotic systems can be largely cast into two groups: humans assisting in computation and computers augmenting human function. Examples from the former category include crowdsourced intelligence and data collection systems that employ game mechanics to entice participation. Examples from the latter group include language translation, citizen policing, health maintenance, and the management of distributed and ad hoc organizations (FEMA, nation rebuilding).

The line dividing these two categories is increasingly blurring, such as with language translation systems. Human-quality language translation is becoming a reality within some language families, augmenting our ability to communicate. At the same time, the algorithms that make this possible rely on human corrections and feedback to improve their performance.

Challenge problems include:

- *A universal translator* that allows people of different languages to communicate effectively. Such a device would not only need to perform machine translation, but would also need minimize its invasiveness in the conversation – requesting help or restatements at times, providing approximate translations at other times. Accurate, automatic translation between 50 human languages would be a good specific goal: the input may be text or speech; the machine should know if it is failing and be able to gracefully degrade; when translating foreign text into English, the translator should provide automatically-generated footnotes with extra background information needed for typical English speakers to properly understand the text. Achieving these goals will require deeper modeling of language structure in addition to new training algorithms.
- *Immersive learning techniques* that would enable learning about a new environment without the human operator having to be in that environment.
- *Conversational information systems* that allow natural language inquiry over the course of a conversation. Having a conversation with a web search engine, or medical database, are examples.
- *Understanding human communication,* including multimodal aspects.
- *Large-scale collaboration.* How do you get one million people to collaborate to achieve some high level goal (that requires coordination)? How can we create systems that utilize components contributed by thousands of individuals?
- *Large-scale knowledge collection.* How do you organize the collection of large data sets, such as medical knowledge, computer software repair techniques, or problem solutions?
- *Computer-assisted self-governance.* How can we build systems that guide the process of self-governance? At the individual level, such a system could be a

personal health coach. At the community level, we could build systems that distribute tasks such as policing and community maintenance.
- *Crowdsourced intelligence and idea markets.* How do we employ crowdsourced information to help identify terrorists or identify vulnerabilities?

Tackling these challenges will require fundamental advances in areas such as:
- *Intelligent systems techniques.* One-shot learning, scalable algorithms that combine statistics and logic, understanding semantics / sarcasm / cultural influence.
- *Crowd programming.* Who to ask for what? How to best use people as a part of algorithms? How do you use crowdsourcing to train experts? How do you use a crowd to train an algorithm, or an algorithm to train people? How do you program the crowd?
- *Sensing.* Behavior imaging (social and sleep patterns, etc.), brain-computer interface.
- *Human-centered computing.* How to motivate people (e.g., to participate in the crowd), interface design, studies of how people learn a new environment, how to present information effectively and in an actionable way, functional models of human behavior.

4.5 Security and Privacy

Security is critical in the DoD context. This includes computer security (mobile devices, desktop devices, servers, cloud platforms, etc.) and information security (aggregation/ processing of sensitive data, disclosure of private information, provenance of information).

The topic is particularly challenging due to the many facets that security represents, and the often-reactive nature of the problem. Most researchers have given up on preventing all security attacks due to (for example) the role of humans as the weak link. Thus, future research must revolve around reactive systems that can automatically recover from attacks, and distribute and apply security patches.

The ideal system can automate workflows such as: 1) *identify* that the system is not running properly and isolate the problem, 2) *search* the network for other instances of the problem and (possibly partial) solutions, 3) *try and choose* solutions that work, and 4) and proactively *prevent* the problem from arising in the future. This requires Big Data technology that logs all state, actions and context, *and* analyzes the data efficiently.

Challenge problems include:
- *Eliminate or mitigate large-scale botnets*. We want to disable their command and control infrastructure, achieve safe military applications despite the existence of malware, and ultimately achieve complete malware resistance.
- *Complete data recovery*. We must be able to recover from attacks on distributed apps (possibly months after compromise) with no data loss

- *Internet-scale data provenance system*. It must span devices (e.g., phones, sensors, laptops) and be able to answer questions such as: Where did this data come from? What are the legal policies for data dissemination and use? Who has copies of this data? How is it used? Where did this information come from? Such a system must also be able to be managed, so that users can revoke/delete copies of private data.
- *Accurate predictions* of the possibly negative consequences of computational actions.

Tackling these challenges will require fundamental advances in two major categories:

- *Information security.* Data-centric security, including system-wide data provenance, understanding data lifecycles (the lifetime, auditing, deletion), and performing computation of secured (e.g., encrypted) data (e.g., fully-homomorphic encryption).
- *Software security.* Focuses on the actual systems, such as secure layering techniques (e.g., virtualization, interpreters), attestation and software verification, undoable systems, secure applications, and protocols, strong (but inexpensive) authentication, and intrusion detection (machine learning, crowd-sourcing, execution auditing).

4.6 Computing Technologies

DoD requires significant advances in energy-efficient computing and wireless spectrum efficiency. Aggressive exploration of quantum computing also is essential.

The military wants to deploy devices that span a spectrum of processing speeds and energy densities – handheld devices in the field or and powerful machines in a base. Regardless of the scenario, the performance/watt metric is still critical – in-the-field devices cannot afford to run out of energy, and neither can servers in a base afford to process intelligence inefficiently. This can be addressed by identifying more readily available energy sources, and in utilizing the energy more efficiently. In pushing the second option, future software must become energy aware, and be able to adaptively scale its energy requirements depending on the energy availability. However, pushing energy efficient software forward will likely require programmers to consider energy utilization during development and impact programmability. Ultimately, the tradeoffs between performance, energy usage, and programmability must be addressed, and is currently an open problem.

Complementary to processing performance is networking performance. In comparison to processing power, wireless networking bandwidth in the face of dense groups of users has improved at a comparatively glacial pace. For example, participants at a baseball game are unable to simultaneously upload mobile videos of the game. Although the existing available spectrum can likely support 10x improvements, two key problems prevent further advance: highly variable channel qualities and interference. For example, as more fans in an area try to upload videos, their transmissions interfere with each other and reduce the overall bandwidth. A key goal is to realize order-of-magnitude increases in wireless bandwidth while supporting the same user densities.

Finally, there is a global race to be at the forefront of quantum technology. Although we are steadily able to factor larger and larger numbers, the current trajectory falls short of running a real, practical application that highlights the utility of quantum computing. In fact, we currently have not even designed benchmarks that show quantum computing executing classically exponential algorithms.

Challenge problems include:

- *Real-time threat analysis on a hand-held.* Requires processing data from video, audio, other sensor streams on an energy limited device (5 years).
- *10 Peta-op data center than runs on a portable 50KW generator* (10 years).
- *Handheld that never needs to be charged* (10 years).
- *A high density wireless system* that exceeds the current technology trajectory by 4-8x in bits/s/Hz under real-world conditions.
- *Develop a quantum experiment* that, although noisy and non-universal, cannot be simulated efficiently on a classical computer; a small scale (e.g., 50 qubits), experiment can already start exploring properties like high temperature super-conductivity, which we do not know how to do classically.

Tackling these challenges will require fundamental advances in areas such as:

- *Low-power processors.* Rethink processor architecture to optimize for efficiency versus performance; how to use lots of lower performance processors; heterogeneous special-purpose functional units.
- *Low-power memory systems.* Both architecture and memory interfaces (3D stacking, low power SRAM architectures, sub-threshold logic.
- *Low-power interconnects.* On-chip photonics, low-swing interconnects.
- *Algorithms that minimize data movement and energy.*
- *Systems that optimize automatically.* Across languages, compilers, runtimes, OSs, and architectures for both performance and power (organic computing, auto-tuning)
- *Cross-layer protocols to improve spectrum efficiency.* Handling interference and time variations, sensor-augmented mobile network architectures, and several ideas from the convergence of networking, communications/information theory, and systems.
- *Agile and opportunistic spectrum management.* Dynamic management, use-as-needed, accommodate primary/others, and other policies.

5. List of Workshop Attendees

Anant Agarwal
MIT

Hari Balakrishnan
MIT

Bob Bonneau
Air Force

Tom Dietterich
Oregon State University

Erin Fitzgerald
ASDRE

Susan Graham
UC Berkeley

Joe Hellerstein
UC Berkeley

Purush Iyer
ARO

Frans Kaashoek
MIT

Adam Kalai
Microsoft Research

Kevin Knight
USC, ISI

Butler Lampson
Microsoft Research

Ed Lazowska
U. Washington

Pat Lincoln
SRI International

Sam Madden
MIT

Beth Mynatt
Georgia Tech

Sandy Pentland
MIT

Daniela Rus
MIT

Bob Sproull
Oracle Labs, ret.

Robin Staffin
ASDRE

Madhu Sudan
Microsoft Research

Josh Tenenbaum
MIT

Ralph Wachter
Navy

Stu Wolf
ASDRE

Wim van Dam
UC Santa Barbara

Luis von Ahn
CMU

Kathy Yelick
Lawrence Berkeley
National Lab & UCB

Nickolai Zeldovich
MIT

Victor Zue
MIT

We acknowledge the help of three MIT graduate students – Ted Benson, David Wentzlaff, and Eugene Wu, who served as scribes at the workshop.