GRAND RESEARCH CHALLENGES IN INFORMATION SYSTEMS

COMPUTING RESEARCH ASSOCIATION

COMPUTING RESEARCH ASSOCIATION

A Conference Series on Grand Research Challenges in Computer Science and Engineering

Report of the First Conference – Information Systems

As we begin the 21st century, it is appropriate for the research community to consider the research challenges that exist on many fronts and to clarify our priorities. With the leadership of the Computing Research Association, we have inaugurated a series of Grand Research Challenges conferences to explore priorities for information technology research. This report documents the conclusions from the first conference.

The focus of this meeting was information systems. Software and hardware link physical entities and people into networks that could not have been created just a few decades ago. We are still struggling to understand the implications of these changes and to realize their potential.

Leading computer science and engineering researchers came together in June 2002 at Airlie House in Virginia to discuss and debate the systems research challenges of the future. As participants in the technological change that is reshaping our age, we are excited by both the vitality of the intellectual community that is driving progress in information systems and the pace of discovery that has resulted. This conference represented an attempt to gain a larger perspective on our efforts, to suggest directions in which our energies and talents can be used for the betterment of society, and to underscore the long-term technological issues that must be resolved for us to be successful.

— Anita Jones, Chair of the Organizing Committee

ACKNOWLEDGMENTS:

Funding for this conference was provided by National Science Foundation Grant No. 0137943.

We gratefully acknowledge the assistance of Mr. Charlie Feigenoff in the preparation of this report.

To request additional copies, send an e-mail to: info@cra.org or call 202-234-2111.

TABLE OF CONTENTS

INTRODUCTION

Information technology has amplified our intellectual and physical abilities more than anything since the development of the written word. Engineering marvels such as the Internet, the global positioning system, and the human genome project became possible only with advances in information technology. Today there are eight billion computers in the world. Most are embedded invisibly in products, making goods and services safer, more secure, flexible, and energy-efficient, and less expensive than ever before. The tremendous advances in productivity that we have witnessed in the past decade rest on this foundation.

By one estimate, we have seen a 1,000,000,000,000,000-fold decrease in the cost of computation in the last 100 years, that is, it is 10^{15} times more cost-effective to work with a modern computer than a turn-of-the century mechanical tabulating machine. We are hard-pressed to think of change of comparable magnitude in human history.

Quantitative changes in performance and affordability have led to qualitative changes in applications of computing. The recent past has seen explosions in cellular telephony, digital photography, digital video, and electronic commerce as the costs of these computer-enabled applications decrease. Networking, in particular, has enabled new businesses and new ways to link people. We have already moved beyond stand-alone computers or components to build large, integrated, distributed information systems that are in service to society.

In the future, we can expect our computational infrastructure to offer an even more impressive range of social and economic benefits as it grows to include billions of people worldwide.

In the future, we can expect our computational infrastructure to offer an even more impressive range of social and economic benefits as it grows to include billions of people worldwide. Information technologies have the potential to reduce energy consumption, provide improved health care at lower cost, enhance security, reduce pollution, enable further creation of worldwide communities, engender new business models, and contribute to the education of people anywhere in the world. These new benefits will be facilitated by geometric advances in semiconductor and magnetic storage, as well as in electronic and optical communications.

Grand Research Challenges

In June 2002, the Computing Research Association, with financial assistance from the National Science Foundation, convened a group of about 65 researchers from the public and private sectors. During a three-day conference these researchers discussed the specific and urgent research challenges related to building the systems of the future.

As a result of the discussions, participants selected five Grand Research Challenges that will provide a focus for more directed and more immediately relevant research. They are:

1. **Create a Ubiquitous Safety.Net.** Providing a ubiquitous Safety.Net will save lives and minimize damage from disasters through timely prediction, prevention, mitigation, and response.

2. **Build a Team of Your Own.** Building cognitive partnerships of human beings with software agents and robots will enhance individual productivity and effectiveness.

3. **Provide a Teacher for Every Learner.** Tutoring each individual in a tailored, learner-centered format will enable people to more fully realize their potential.

4. **Build Systems You Can Count On.** Assuring reliable and secure systems—from the regional electric grid to an individual's heart monitor—will allow us to rely on information technology with confidence.

5. **Conquer System Complexity.** Building predictable and robust systems with billions of parts will enable broader and more powerful applications of information technology.

These challenges are deliberately monumental. Although these Grand Research Challenges incorporate problems that are already the focus of research, our challenges set goals that are broader than most typical research agendas and that move far beyond what has been demonstrated by today's research. In most cases, it is precisely the scale of the vision that makes it challenging. Each challenge will require at least a decade of concentrated research in order to make substantive progress. Each is deserving of considerable investment by government and private funders because each challenge, if successfully met, will materially advance the capabilities and the civilized conduct of society.

Participants decided that there are advantages to a framework for systems research that emphasizes full-system solutions to daunting problems. Aspiring to build full systems will drive progress in the field at a more rapid rate than otherwise. Common investment and shared vision will encourage the development of a consistent set of standards that will lead to a more efficient technical and societal integration of new technologies.

Rather than propose a single Grand Research Challenge, we propose an interrelated set of challenges. This approach has enabled us to identify a set of deep technical problems that apply to several of the Grand Research Challenges. Solving similar problems in different application contexts will encourage solutions that are far-reaching and durable. Our economy will benefit from the useful innovation that results from a more unified approach.

It is worthwhile remembering that while individual components may be elegantly and effectively modeled by mathematics, complex systems that serve society are inherently messy. They must adapt to human habits and procedures, understand verbal and nonverbal cues, and present an interface that seems intuitive to human users, among other criteria. In other words, they must be efficient for human use, even if not computationally efficient. Accordingly, meeting our Grand Research Challenges requires the collaboration of experts from many fields, such as cognitive psychology, biomedical engineering, mathematics, and linguistics. We currently lack a powerful, generalizing framework that provides an overriding context for coordinating and motivating this needed research effort. We believe that our set of Grand Research Challenges provides such a framework.

In this document, we discuss each chosen Grand Research Challenge, revealing the rationale behind our choice and the series of component challenges that it encompasses. Each of the first three challenges directly addresses a fundamental societal issue. The fourth and fifth challenges are more technical in nature and address information system reliability and complexity in all their aspects.

Conclusion

Meeting the challenges posed here will require us to dramatically extend the frontiers of knowledge in science and engineering. But they will also cause us to think more coherently about the proper role for business and government, the development of public/private partnerships, and the reconciliation of different priorities and points of view. Thus, they have the potential to serve as a focal point for discussion, helping decision-makers to craft coherent policies on such issues as intellectual property, privacy, and access to information in normal and emergency situations. Addressing such policy issues is essential if we are to realize the full promise that information systems can deliver.

GRAND CHALLENGE 1.

Create a Ubiquitous Safety.Net

Disaster Prediction, Prevention, Mitigation, and Response Through Ubiquitous Computing

Introduction

Earthquakes and strokes. Terrorist attacks and deforestation. Bridge failures and floods. Disasters take many forms. They can devastate one person or one million. They can occur in the blink of an eye or over hundreds of years. They can involve manmade objects or natural forces. But no matter what their scale, ubiquitous computing networks that combine sensors and actuators with computing and communication technology hold the promise of preventing or at least mitigating the impact of disasters, improving the speed and quality of our response, and guiding recovery efforts. One of the grand challenges facing us as we move forward is to weave this ubiquitous safety net into our environment.

The vital importance of creating such a safety net is cast in high relief by low-probability, high-impact events, whether they are natural disasters like earthquakes or manmade like terrorist attacks.

The vital importance of creating such a safety net is cast in high relief by low-probability, high-impact events, whether they are natural disasters like earthquakes or manmade like terrorist attacks. After a disaster occurs, the incident command often needs 10 to 12 hours to access the hot zone adequately and to plan a coordinated, safe response. In effect, 25 percent of the first critical life-saving 48 hours is spent in preparation, trying to locate key information (such as building plans and street and utility maps) and necessary equipment, assembling experts, and examining and analyzing the site. Engineers and other specialists must determine the structural integrity of surviving buildings and check for hazardous gases or conditions. People must be evacuated, and mechanisms put in place to distribute information and directions to large groups of people including the press, victims' families, and arriving rescue teams.

Even after the basic facts are known and a plan is established, rescue and recovery operations proceed fitfully. Operational information is usually disseminated slowly and, in some cases, unevenly. Rescuers conduct technical searches and rescues with low-tech equipment, usually drawing maps by hand to describe where they have been and reporting results by radio at the end of shifts. The lack of integrated weather information and accurate positioning data means that rescuers may fail to search areas of the hot zone or may expose themselves to toxic plumes.

To make matters worse, all members of the response team experience cognitive disability; stress, fatigue, and emotions reduce decision-making capabilities. Due to uncertainty about the environment, responders in the field must sometimes work inside personal protective equipment that is not only distracting and heavy, but may also limit their visibility and human performance.

The results of delay and inevitable error during the first hours of a disaster can be tragic. In the case of structural collapse, survivors deep within the structure may wait 4 to 10 hours before they can be removed and given medical care, and many die before they can be reached.

Ubiquitous computing systems could significantly alleviate these problems. The Grand Research Challenge is to create a ubiquitous Safety.Net. Linking ubiquitous networks of sensors and actuators with existing information systems would create a platform for performing data mining, fusion, and management. Such a system would improve both the quality of information available to planners and the speed with which they receive it. It would provide the intelligent assistance critical for decision-making, rapid planning, resource allocation, and incident command response. The result would be a revolutionary enhancement of the efficiency and effectiveness of the current 9-1-1 emergency response system.

Research Required by the Challenge

Creation of a Safety.Net system is impeded by unsolved research problems in a number of areas. They can be grouped into four broad, interdependent categories:

1. Embedded multimodal sensor/actuator nodes
2. Systems architecture and communication services
3. Data fusion and analysis
4. Coordinated response

There is considerable overlap in this taxonomy, and there are direct linkages to research required by other Grand Research Challenges in the areas of complexity and augmented cognition.

1. **Embedded multimodal sensor/actuator nodes**
   The key to creating an effective Safety.Net is up-close, highly capable in situ sensing and actuation, with the potential for dense deployment of such nodes. The sensing components must be multimodal, combining a wide range of physical, chemical, and biological inputs. Similarly, actuation may deploy smart materials or chemical neutralizing agents.

In addition to sensing and actuation within physical structures such as buildings and roads, these systems must also integrate biocompatible devices that can operate in and on humans for medical assessment and delivery.

Dense in situ deployment will also determine the characteristics of these nodes. Where the number of measurement points within an area is very large, there will be a need for on-board processing of sensed data. The sensors and actuators will be part of a network that is established through on-board wireless devices, which in turn require local computing capability. A critical element in developing such nodes is minimizing their energy demand and developing a low-power energy source to meet that demand.

A great deal of the sensing/actuation infrastructure must be put in place prior to any emergency effort. Such an investment may be worthwhile because it could also be used to increase the efficiency of the everyday operation of an environment. For example, motion detectors useful in emergencies may be used to create a more energy-efficient building. Ideally, these preexisting nodes could create a heterogeneous, evolvable, and long-lived infrastructure. During an emergency, they would be supplemented by rapid deployment of a variety of new and different nodes specialized for operation in a harsh, emergency environment. Some nodes will need to be mobile, collecting data opportunistically or explicitly when carried by people, aircraft, or robots. In situ nodes and new ones, stationary and mobile, must be capable of rapid self-configuration and of maintaining coherence in a dynamic environment.

In situ nodes and new ones, stationary and mobile, must be capable of rapid self-configuration and of maintaining coherence in a dynamic environment.

2. **Systems architecture and communication services**
   A primary research challenge is to create a purposeful system that connects and integrates highly capable nodes consisting of sensors, actuators, communication devices, power sources, and computers. The system must form a coherent, adaptive whole that processes the gathered data to produce useful information on a priority basis for a multitude of different users who have different objectives. The data should be combined synergistically to produce a rich set of system services. Redundant detections must be suppressed; complementary detections combined; and distributed patterns of interest, such as plume characterization, must be identified and characterized. With the development of highly accurate mapping and localization services, these system services can help map a geographic area and identify and locate victims, rescuers, resources, and hazards.

   We need the ability to process data and derive information locally, as well as to deploy highly adaptive communication capabilities through which otherwise separate entities can share information through coordinated transmissions. While we know how to build large, complex communication systems such as the Internet, the systems described here require fundamental breakthroughs in order to operate under extreme conditions and constraints that may include structural failure, fire, and even intentional disruption.

The core research challenge is to develop an overall architecture, along with a variety of algorithms for distributed computation, secure networking, real-time systems, dynamic reconfiguration, and richly connected networks, while compensating for substantial node failure.

Another key ingredient involves resource mobilization and actuation in real-time. For example, Safety.Net communication services should support real-time video from robots and secure commands being sent to robots. Many other nodes in the system may also have actuators that can be controlled in a coordinated manner.

Finally, the system architecture must allow for application-specific optimization and for interoperation and independent evolution that comes from layered, modular structures. The configuration and the set of services required for each emergency will differ. A Safety.Net, by its very nature, exhibits tremendous complexity, and yet the resulting behavior must be predictable and trustable. Modeling the performance and complexity of such systems will be a significant challenge.

3. **Data fusion and analysis**
   Processing data within and among nodes is just one element of Safety.Net. The ability to fuse related data from multiple sources to derive useful information is critical. The Safety.Net architecture must therefore enable large-scale collection, archiving, and processing of live streams of sensor data for on-line and off-line analyses, decision-making, and dissemination. Key technical challenges include the security of the infrastructure, privacy of individuals and related data, integrity and authenticity of data, and well-defined semantics for the behavior of the information infrastructure under different types of failures.

   To facilitate higher-level inferences, there is also a need for application-specific data fusion in the presence of insufficient and error-prone data. The data-fusion process should include a risk analysis of confidence in the data and of making inferences from faulty and insufficient data. Analysis and interpretation of collected data in conjunction with information databases (that could often be out of date and/or have partial or incorrect information) are essential in enabling sound decision-making. Other essential components of the information architecture include novel data models, segmentation, and organizational methods for storing and mining collected sensor data in conjunction with existing databases.

4. **Coordinated response**
   Above all, the Safety.Net system must serve the rescue workers. The incident command must be able to track personnel on the scene, as well as the logistical resources needed to coordinate a rescue. To accomplish this goal, the system must produce coherent and eas-

ily understandable descriptions of the overall status of the environment and the systems and personnel within it, based on aggregate information from the sensor network as well as external data sources. Further, it must be able to do this in the presence of dynamic reconfiguration of devices—both in response to failures and on-demand reorganization–while maintaining near real-time interactivity. The system must also allow rescue personnel to reconfigure the states and locations of sensors, actuators, and mobile elements easily as they respond to the evolving emergency.

To address this set of issues, research must be conducted in developing context-appropriate accounts of the state of the environment, autonomous and semi-autonomous robotics, knowledge systems that are adaptable to unreliable and partial information, efficient algorithms for real-time analysis, and the larger distributed command and control of personnel and resources. These situations are themselves inherently complex, and the systems that will manage them will likewise be highly complex. We must produce systems that can be trusted even in the face of this complexity. (The types of cognitive systems proposed in Grand Research Challenge 2 are relevant here.)

The large, complex, and real information system applications that are the heart of Safety.Net require more than technical solutions. They must be developed with sensitivity to social and cultural norms.

We must also make this information easily accessible to people who can use it. They will need to access and manipulate the system, use it to deploy resources and personnel, and do so quickly and with low cognitive overhead in a high-stress environment. For some rescuers, user interfaces must be multimodal, since they may be used in environments that are alternatively noisy or require hands-free interaction. Subsets of the information in the system must be made available to parties other than the rescuers themselves. These parties include victims and potential victims, the media, both on-site and off-site government officials, hospitals, and so on. This range of constituencies means that the systems must support tailorable presentation, adaptable to different needs, contexts, privacy situations, and languages. The core research areas include multimodal user interfaces, language translation, visualization, and security.

Conclusion

The large, complex, and real information system applications that are the heart of Safety.Net require more than technical solutions. They must be developed with sensitivity to social and cultural norms. For instance, Safety.Net raises the issue of individual privacy. A Safety.Net system designed to gather and process information in the environment could be considered an invasion of privacy. Appropriate safeguards, strict guidelines, and user education must be employed to respect public concerns, both in benign times and during an emergency. Creating such safeguards may entail the formation of new public policy and new laws.

At the same time, work on Safety.Net would have significant and immediate social benefits. As mentioned earlier, significant elements of a Safety.Net system should be in place before disaster strikes. Such nodes would have multiple functions. Multimodal nodes linked by the Safety.Net system architecture could be included in structures like bridges and elevators; lighting, heating, and cooling systems in large buildings; and many other aspects of the environment and infrastructure where they could detect deterioration and monitor usage trends. In this way, they could extend the lifespans and increase the efficiency of manmade structures. At the same time, they could be used proactively to predict incipient failures and prevent or mitigate disasters. They could also be used to furnish data needed to simulate and model potential disasters and to train responders.

The ability of these ubiquitous Safety.Nets to predict, prevent, mitigate, and respond to emergencies makes them ideal for human applications, especially when coupled with research in augmented cognition. To provide one example, Safety.Net, utilizing some of the cognitive technologies discussed in Grand Research Challenge 2, could be adapted to enable people to remain in their own homes as they age. With Safety.Net, seniors could live in an environment that would respond to their needs. It would make it easier for them to perform necessary daily activities safely, while caregivers could monitor their medical and social needs. It could also be used to detect life-threatening situations and generate the appropriate emergency response.

GRAND CHALLENGE 2.

Build a Team of Your Own

Enhance Individual Productivity and Capabilities Through Cognitive Partnerships

Introduction

For centuries, a partnership between humans and machines has enhanced the productivity of individuals by augmenting their physical capabilities, allowing them to employ more force, move more quickly, work at smaller and larger scales, and operate in otherwise inhospitable environments such as water, air, and space. Within the past century, computers and the programs that run on them have extended this partnership by augmenting individuals' capabilities for routine activities. This new partnership has automated machine control and offloaded rote human activities, yielding faster performance and improved outcomes with reduced human effort.

Cognitive systems are capable of mastery of the tasks on which they work, as well as extracting contextual information from the environment in which these tasks are posed.

The next logical step in the evolution of this partnership is to augment human cognitive capabilities by developing machines capable of offloading human thought processes, yielding cognitive systems capable of actively supporting individuals in pursuing their goals. Cognitive systems are capable of mastery of the tasks on which they work, as well as extracting contextual information from the environment in which these tasks are posed. They solve problems as they arise and plan for the future. They communicate appropriately with others about themselves and their activities in order to work effectively in close collaboration. They also adapt their understanding and skills as they and the world around them evolve.

Cognitive systems consist of some combination of hardware—sensors, actuators, communicators, and computers—and, most importantly, constituent software. Cognitive systems capable of interacting with the world through sensors and actuators are referred to as robots, while cognitive systems operating in the online environment are known as agents. The Grand Research Challenge is to build a team of your own, for each and every individual, composed of robots and agents functioning as active partners. These teams effectively pool the capabilities of the robots and agents in service of the individual, simultaneously reducing the individual's supervisory burden by taking an active role in determining their own behavior. They amplify the individual's physical abilities by providing new sensory and motor capabilities and by enabling the individual to exercise these capabilities at a distance. They enhance the individual's cognitive abilities by supplementing memory and problem-solving capabilities and by providing direct access to relevant data, expertise, guidance, and instruction. They work towards shared goals while understanding enough about the task, the individual, and each other to assist, mentor, cooperate, and monitor as needed. And they reduce the individual's performance degradation by offloading activities and by anticipating the kinds of errors that tend to occur in stressful situations.

Cognitive Partnership Scenarios

The impact of teaming individuals with agents and robots will affect many areas of human endeavor. These partners will, for example, provide assistance when needed, mentor individuals when they are in unfamiliar territory, facilitate lifelong learning (e.g., via the "Learning in Context/Just-in-Time Learning" capabilities of Grand Research Challenge 3, Provide a Teacher for Every Learner), use advanced sensing and scanning technologies to monitor and adapt to the individual's physical and mental status, help maintain the individual and his or her environment, act as facilitators in communications with other individuals and cognitive systems, and serve as foils in brainstorming sessions.

For seniors, teams of agents and robots will provide an evolving "antidote" to their growing physical and mental frailty to enable them to live longer on their own.

For knowledge workers, teams of agents will help locate and assemble the information they require; transform it into the forms necessary to facilitate human and machine understanding; organize it into categories; suggest how to respond, act on, or process the information; and manage the dissemination of this processed information.

For seniors, teams of agents and robots will provide an evolving "antidote" to their growing physical and mental frailty to enable them to live longer on their own (and in this capacity, could be part of the Ubiquitous Safety.Net described in Grand Research Challenge 1). Physical complementation through robotic assistance of various sorts has been under investigation for some time, but it will also be possible to provide cognitive supplements to an individual's memory and problem-solving abilities, as well as to provide appropriate monitoring.

For disaster response and military operations, partnering individuals with teams of agents and robots should enable a range of new concepts of operation, while simultaneously reducing casualties. Such partnerships will involve robots for reconnaissance (and possibly force projection), networks of autonomous sensors, and agents capable of providing information and coordination services (e.g., as discussed in Grand Research Challenge 1 to Create a Ubiquitous Safety.Net), all in service of augmenting the effectiveness of soldiers or emergency responders. The partnership will help protect them from harm and enable some of them to participate remotely from safer locations. Such partnerships necessarily provide the flexible communication and coordination support needed for dynamic command and control. In the case of disaster response, the partnership will contribute to communication and coordination across the full set of stakeholders—public and private organizations; local, state, and federal response agencies; diverse local communities; and the general public. In both disaster scenarios and in military coalition operations, new ad hoc partnership structures will be dynamically created.

In both disaster scenarios and in military coalition operations, new ad hoc partnership structures will be dynamically created.

Although these scenarios only begin to suggest the possible impacts of partnering individuals with teams of agents and robots, they should be sufficient to indicate the pervasive and deep effect that is possible.

Some progress has been made in creating teams for such scenarios, but these predominantly consist of either an individual partnered with a single agent or robot, or a small, homogeneous team of agents. For example, the U.S. military and the intelligence community partner individuals with unmanned air vehicles for military operations and for the wars on drugs and terrorism. Initial prototypes of such vehicles have moved rapidly from the research laboratory to the field as their value was recognized. Research teams with robots rushed to aid the response to the World Trade Center disaster. Prototype search and rescue robots discovered five victims in the rubble. Likewise, teams of intelligent automated pilots and commanders have been successfully partnered with higher-level human commanders in distributed battlefield simulations, while teams of office agents have worked successfully on a 24/7 basis to assist in the accomplishment of routine organizational tasks.

To provide a concrete example, consider military simulations. Simulated commanders are already reasonably competent. They can receive orders from their human commanders in a formal command language; plan the missions for their units; communicate these missions to their subordinates in the same command language; and monitor and re-plan the missions as changes in the situation dictate. Simulated pilots can accept orders from their commanders; fly their missions; coordinate with other members of their unit concerning timing and roles; report back to their commanders on what they observe; reorganize automatically to adapt to losses of teammates; engage enemy forces in a coordinated manner; and defend themselves when attacked.

But, these are small, relatively homogeneous teams that work only in relatively narrow domains, that is, for specific kinds of military missions and then only for the length of single missions (minutes to hours). Furthermore, they lack essential capabilities required of team members. For example, they are incapable of learning or adapting their behavior based on experience, dynamically forming new teams on their own, understanding or reasoning much about the context in which they operate, sharing initiative with other elements except as preplanned, communicating outside formal languages, or forming a true collaborative relationship with human beings. Accordingly, the research problems that remain are substantial.

Research Required by the Challenge

There are two primary technological challenges that must be met to achieve a team of your own. The first technical challenge is to incorporate cognitive capabilities into both programs and machines, that is, to upgrade them from the status of automation tools to that of cognitive systems. They should be capable of understanding tasks and tools as well as the context in which they are embedded. They should be capable of acquiring knowledge by learning from experience and/or instruction. They should be able to represent their acquired knowledge in visual displays and in appropriate human and formal languages. They should be able to track the evolution of their environment over time and infer critical consequences of what is known and sensed. They should be capable of acting in a goal-oriented manner, while solving problems as they arise and planning for future contingencies. And they should be capable of operating over extended periods of time and across a broad range of domains.

The second technical challenge is to effectively team people, agents, and robots. This requires progress in technologies and techniques to solve five deep problems related to the teaming of heterogeneous cognitive systems:

1. Communicating effectively via both formal languages and traditional modes of human interaction, such as speech, gesture, and facial expression.

2. Understanding the capabilities, limitations, and status of teammates of varying natures.

3. Dynamically forming, adapting, coordinating, and monitoring teams. This includes discovering and allocating participants, roles, tasks, and resources, as well as training teams to develop mutual understanding and skill and creating and maintaining appropriate organizational structures.

4. Providing cognitive services such as assisting and guiding team members, providing the initial training required to bring new team members up to speed, and operating legacy systems.

5. Collaborating effectively across potentially large teams through appropriate use of such concepts as responsibility, authority, and autonomy.

Conclusion

Enhancing the productivity and capabilities of human beings is a most worthy challenge. The leverage of even modest success in augmenting the cognitive capabilities of human beings is immense because such an augmentation could potentially affect almost every domain of human endeavor.

There are, however, social as well as technological challenges that must be met in order to create cognitive partnerships. The greatest social challenge centers on the question of trust, and is closely related to Grand Research Challenge 4, Build Systems You Can Count On. To participate in the envisioned cognitive teams, individuals must trust the agents and robots with which they work to do what is expected of them, to always do something that is reasonable, to safeguard private information, to support their teammates in critical situations, and, most critically, to make life better rather than worse. Otherwise, human beings will not be inclined to participate in close collaborative relationships with robots and agents, or even to accept the idea of such relationships.

The other major social challenge focuses on such legal questions as who owns the agents and robots and who is liable for the actions of such teams. Neither of these social challenges is insurmountable, but addressing them requires a combination of progress in the technical, policy, and legal spheres.

GRAND CHALLENGE 3.

Provide a Teacher for Every Learner

Scalable, Learner-Centered Networks

Introduction

In today's high-tech, global economy, what you earn depends on what you learn. Making the American dream a reality for all Americans, not just a privileged few, requires that educational excellence become the norm rather than the exception. Although information technology is not a panacea for all of the shortfalls associated with our educational system, it offers the potential not only for significantly enhancing learning for all learners, but also for transforming the way we learn. The Grand Research Challenge of providing a teacher for every learner involves the development of a learning environment to make that transformation. Advances in information and communication technologies have the potential to enhance lifelong learning along the continuum from "K through gray." It is now within our power to create a future in which all Americans will be able to:

Advances in information and communication technologies have the potential to enhance lifelong learning along the continuum from "K through gray."

- Participate in networked and face-to-face communities of learners composed of peers, teachers, mentors, domain experts, avatars, and "cognitive" tutors that collectively approach the effectiveness of a one-on-one human tutor.
- More intuitively understand challenging concepts in mathematics and science using interactive simulations and information visualization.
- Create their own physical and/or digital artifacts that are personally meaningful to them and that challenge them to "learn by doing."
- Make the most of their talents, irrespective of their physical and mental disabilities.
- Tap into rich, universally accessible digital libraries with books, articles, music, paintings, primary source material, data sets, and 3-D representations of cultural and natural landmarks.
- Learn at their own pace and in their own style.
- Receive continuous, customized, and meaningful feedback and assessment.

- Acquire new skills in a way that is compelling and engaging.
- Learn anytime, anywhere—an advantage that is particularly important for adults struggling to balance the competing demands of work and family.

The Grand Research Challenge is to provide learning environments that approach the effectiveness of one teacher for every learner. Such systems, properly used, can produce a significantly better-educated populace by combining advances in learning sciences with advances in information technology.

Research Required by the Challenge

The task of developing the proposed system involves a number of technical challenges in five distinct, but interrelated, areas:

1. Cognitive tutors
2. Simulation-based (clip) models
3. Massive multiplayer online games
4. Collaborative authoring
5. Learning in context/just-in-time learning

For each area we discuss several technical challenges.

1. **Cognitive tutors**

   It has been demonstrated that it is possible for an automated tutor to improve student performance by roughly one standard deviation from the mean for some high school mathematics students. This is a dramatic result. One reason that such tutors are not widely available is because significant human effort is required to develop the specialized knowledge base for each different subject. In addition, we do not understand fully the conditions under which such a tutor will be effective. Significant progress must be made in crafting knowledge representations that are both interoperable and reusable. We need to develop models of the various styles in which a student learns, as well as appropriate pedagogies and assessment techniques.

   The knowledge representations that underlie such tutors should also be designed to incorporate new knowledge about a subject area, as well as advances in knowledge and techniques associated with both pedagogy and assessment.

2. **Simulation-based (clip) models**

   An important genre for next-generation educational software, particularly for training scientists, mathematicians, engineers, and technologists, is what might be called a clip model. By loose analogy to the well-known galleries of copy-and-paste 2-D clip art, a clip model is an interactive microworld, typically simulation- or rule-based, that expresses both geometry and behavior of the modeled entity or concept. It is a self-contained object ready to be embedded in a context such as a hypermedia learning module.

Clip models have two key additional properties that make them more powerful than today's largely stand-alone microworlds. First, they are designed to be combined to form larger models, for example, a heart model may be connected to a vascular system model and to a lung model to create a composite model that simulates respiration and oxygenation of the blood as it is distributed throughout the body. Second, no single model suffices for all learning purposes. Perhaps dozens, if not hundreds, of heart models are needed to meet the needs of learners at different levels of understanding and with different kinds of backgrounds and learning styles. Furthermore, some learning situations will require not just models of the "normal" heart, but also those that model various kinds of abnormalities and pathologies.

[T]here is an enormous . . . challenge in learning how to design models that can interoperate with other models that have different levels of complexity and modeling fidelity.

The fact that we need a family of models raises at least two profound technical problems. First, there is an issue of scale when the number of models in each family can grow almost without limits. In everyday life, the knowledge and assumptions we bring to any project are uneven; we may know more about one aspect of the subject than another. To what degree will it be possible to craft a general model that can be tailored to have the properties needed for different learning situations?

Second, there is an enormous technological, pedagogical, and "ontological engineering" challenge in learning how to design models that can interoperate with other models that have different levels of complexity and modeling fidelity. We lack understanding of how low-fidelity, simple models may accommodate the needs of higher-fidelity models with which they must exchange parameters to simulate the larger composite model. How can one heart model, or model family, serve a college freshman who is learning about heart anatomy and physiology, while that heart model interoperates with vascular and lung models at a much simpler level, say that of a high school biology student? The brute force approach of having all models simulate at the finest-grain level, while exposing only the attributes needed by peer models, is neither cost-effective for model construction nor computationally appealing.

It is likely that a multidisciplinary architecture team will need to design collections of a priori interoperable clip models in each domain of knowledge. This appears to be a Herculean labor, and raises the question of whether we can develop automation tools to expedite model development.

3. **Massive multiplayer online games**

   Anyone who watches someone play one of the popular multiplayer computer games will marvel at the display of concentration, tenacity, and ingenuity. One recent study concluded that the average EverQuest player was spending 4.7 hours a day in Norrath (one of the EverQuest virtual worlds), and that more than 30 percent of the adult players surveyed spent more time in Norrath than they did at work. One obvious question is whether virtual worlds could be used to teach children and adults something more than how to kill monsters and take their treasure.

   Currently, researchers are building experimental environments in which the learner identifies problems within the environment through observation, potentially with help from computer-generated avatars present in the environment. The learner posits solutions that are then played out in the simulated environment. The learner may be a middle-school student who is exploring a 3-D virtual 1880s town where illnesses plague the citizens. Alternatively, the learner may be a university administrator who must juggle the needs, even the demands, of students, faculty, athletes, alumni, unions, and the local city council.

   The learning environment may involve not one but multiple students who communicate with each other audibly or via instant messaging. It is a challenge to build an effective environment, especially one that can interoperate with other learning objects. It is another challenge to make that environment both cost-effective to construct and operate, but also adaptable to the needs of individuals as well as groups of learners.

4. **Collaborative authoring**

   Some progress has been made over the past several years in the fields of computer-supported cooperative work and computer-supported cooperative learning. Shared whiteboards and file editors are just beginning to address the collaborative authoring problem for textual objects, such as documents and spreadsheets. Streaming audio and video and graphical animations can now be easily embedded within documents, hypermedia presentations, and some cooperative environments.

   Authoring in the broadest sense may serve a purpose beyond building an authored artifact. The main purpose may be collaborative problem-solving or cooperative learning, and the artifact may simply be documentation of the result of the intellectual effort. The supporting system must effectively support the interactive, cooperative process, not just artifact construction.

   Technical issues involving synchronization, latency hiding, distributed object replication, copy convergence, caching, and interface support for complex tasks are just some of the key areas in which significant advances must be made. In addition, fundamental research must be conducted to understand the spectrum of activities, from programming to scripting to authoring, that are carried out by both humans and by software.

5. **Learning in context/just-in-time learning**
   It is critical that each individual learner be able to learn in the physical, social, cultural, and virtual context most appropriate to that learner. In addition, the learner may have a very focused objective for a particular learning session. Just-in-time learning over such a broad range of contexts presents a number of technical challenges, including the support of reliable and ubiquitous computing, access and control of remote instruments, flexible digital object sharing, and user interfaces for small-format mobile devices. The automated teacher must tailor delivered information to fit that context and the goals of the just-in-time learner.

Conclusion

Our vision of teachers and learners immersed in a network of rich learning objects that are continually enriched and enhanced by the participants is achievable through anticipated advances in information technology and learning sciences. As a result of these efforts, we will be closer to achieving such goals as having every child reading effectively by the end of the third grade.

The most daunting challenges, however, are not technological, but institutional. Change of any kind in our school system is difficult, given its decentralization. People continue to disagree about both the ends and the means of K-12 education, making sustained movement in any direction difficult. As a result, most schools still lack the necessary computing, networking, and technical resources to capitalize on advances in pedagogical technology, and teachers often are not given the time and training they need to integrate these new technologies into the classroom.

GRAND CHALLENGE 4.

Build Systems You Can Count On

Introduction

A defining characteristic of our age is our reliance on vast, complex, and intertwined information networks. These networks enable the exchange, analysis, and control of information on a scale and of a quality that has never before been realized. Information networks link elements in our electrical grid, communications systems, financial markets, and national defense system. They support the critical infrastructure that is responsible for much of the productivity behind our economic growth in recent years, and they provide the foundation for the quality of life we enjoy and for advances that we contemplate in areas such as medicine, materials, and communications.

In many applications today, the reliability and security of these [complex] systems is unacceptable, creating problems that range from lagging productivity to dangerous vulnerability.

As we gain experience with these networks, the attributes of the information utilities required to realize these advances become increasingly well-defined. They include:

- Global reach: Uniform service should be available and affordable worldwide.
- Scalability: We should be able to increase the capacity of any service by a factor of 10^6 merely by devoting additional resources.
- Persistence: Data that we commit to the utility should be accessible in 100 years.
- Efficient administration: There should be a very high ratio of users to support staff, perhaps as high as a million to one.

We have also learned that as our reliance on these networks grows, so does our vulnerability. This was dramatically illustrated by the events of September 11, 2001, when regional communications and information utilities were disrupted with the loss of the twin towers. On a less catastrophic scale, the continued expansion of information technology is threatened by a relentless succession of bugs, crashes, worms, viruses, and security breaches, and by the increasing difficulty of administering systems. In many applications today, the reliability and security of these systems is unacceptable, creating problems that range from lagging productivity to dangerous vulnerability.

Accordingly, there are two other qualities that we require of all computing and communication systems: security and reliability/availability. Society requires systems you can count on. The objective is to design a new generation of systems that will make today's applications reliable and secure, that can be used with confidence in all elements of our global critical infrastructure, and that will enable entirely new generations of applications that can continue to transform our lives.

Research Required by the Challenge

The fundamental research problems that must be solved to create systems you can count on include:

1. System development tools that reduce the frequency and severity of bugs.
2. System administration tools that reduce the frequency and severity of configuration errors.
3. Understandable, deployable, and usable security.
4. New approaches to the composition of modular elements.
5. New approaches to federation.
6. Pervasive audit trails.
7. Self-adaptive systems.
8. Architectural enhancements to processors.

1. **System development tools that reduce the frequency and severity of bugs**
   We must develop new tools and techniques for reducing the likelihood of software bugs and for improving the trustworthiness of system upgrades and bug fixes. To increase the quality of code, programmers must have security-oriented tools for system development. These include improved programming languages and environments that reduce the likelihood of security-related errors, tools that audit programs for common security-related flaws, and additional run-time checking of security-related properties. Improvements in the system-upgrade and bug-fix processes also are critical. Among other tools, we need ways to test bug fixes reliably given the enormous number of configurations in the field, to design improved modularization techniques, and to develop techniques that help make installation easier and more reliable.

2. **System administration tools that reduce the frequency and severity of configuration errors**
   Human error is often cited as the cause for configuration errors. In reality, much of this so-called human error is not produced by carelessness or imperfect knowledge, but has its roots in the design of the system and its corresponding administrative interface. Certainly, systems and their components will be more robust if we engineer them to require far less human intervention, and to adapt and continue operating on their own under a wide range of conditions. Nonetheless, significant improvements will accrue to systems that are explicitly designed to work well with their constituencies of users, administrators, and maintainers.

Configuration management has yet to enjoy the kinds of static and dynamic checking, such as type checking, that are routinely used to catch errors in software.

3. **Understandable, deployable, and usable security**
   Today, cryptographic algorithms and standards for authentication, security, and privacy are far ahead of our ability to deploy, administer, and use security systems. All too often security policies are implemented mechanically, producing no real gains in security. We require new specification techniques for security policies that are meaningful to system administrators and end-users, so that security is deployed in a way that meets the expectations of users, administrators, and security engineers.

   Today, cryptographic algorithms and standards for authentication, security, and privacy are far ahead of our ability to deploy, administer, and use security systems.

   Improved techniques and tools for administering security should include accommodating common operational scenarios in a straightforward way. This might include the business failure of a security service provider, the merger or fission of security administrative domains, or the compromise of a critical resource. Because there is always a trade-off between security and convenience, there are opportunities for user-centered design to improve the acceptance and deployment of better security.

   Encryption is one specific instance that deserves special mention. Well-encrypted messages can move safely through dreadfully weak systems. Encryption is well understood, but not widely employed. There is a "usability gap" that translates directly into a "usage gap."

4. **New approaches to the composition of modular elements**
   The problem of creating a coherent system from modular components is well known: individual components can fail, the properties of many components may be incompletely specified, some components may actually be malignant, and properties of the whole system may not be easily derived from properties of the components. We must develop ways to compose systems from subsystems in ways that enable the systems to be more trustworthy than their components.

   These advances will require new ways of describing the properties that are actually needed for successful system operation and defining the properties that are actually assured by each subsystem. These descriptions will include many aspects: the modes and probabilities of failure; the fault-tolerance provided; some characteristics of performance such as worst case and average case; and abstract functionality. It will also be necessary to determine critical properties of compositions from partial descriptions, and to bound the system changes that may result from changes to (or addition or removal of) components. Finally, those who are responsible for systems and components must be able to write and understand these descriptions fluently and relate them to their own intentions. They must also be able to use automated tools that perform deductions about compositions.

5. **New approaches to federation**
   Today, federation is recognized as an architectural device that enables systems to grow and evolve. For example, federation can be used to allow directory servers created by different vendors, configured in different ways, and confined to different types of information, to appear as if they form a unified directory service. Individual servers and services can join or exit the federation while it operates. Although the federation technique appears promising for supporting system robustness (through redundancy) and evolution (as new services join a federation), we lack principles for designing new federated systems with guaranteed properties.

6. **Pervasive audit trails**
   Audit trails are used to detect, analyze, and repair errors in human and computer systems. They are used to track hackers who penetrate servers, as well as to unwind financial transactions that go awry. They can also be valuable in ensuring that privacy constraints are adhered to. Vastly improved auditing techniques are required to improve the robustness of computing and communication systems. Simple ways must be found to demonstrate convincingly to both a customer and a provider that transactions are executed and completed successfully.

7. **Self-adaptive systems**
   Network reliability and security will be increased by systems that observe the current condition of the global network and individual applications and adapt to deliver target levels of security and availability/reliability. They would perform this task without (or with minimal) human intervention. If a distributed denial-of-service attack is launched, for example, the system would replicate the service to alternate sites and coordinate the routing infrastructure to route requests to new places. Similarly, if particular network links fail, the system might heal itself by recruiting additional replicas and automatically retiring old ones. The goal is to achieve a mean time to repair that is accomplished in electronic time scales (seconds and microseconds), rather than human time scales (minutes and hours).

8. **Architectural enhancements to processors**
   To provide a base on which to build trusted systems, the architecture of the underlying circuitry and instruction set will have to be redesigned so that trust is a key design criterion, along with speed, power, area, and other aspects. For example, smart cards are vulnerable to a non-invasive attack that uses power and timing measurements to extract the cryptographic key. A technical challenge is to design trusted hardware on which the system depends, in conjunction with the algorithms such as encryption, so that such attacks are not effective.

Conclusion

The aspirations that we have for our computing and communication systems, the functionality that we desire, and thus the complexity of these systems are constantly increasing. At the same time, our expectations for ease of use and reasonable costs are similarly demanding and open-ended: a system can never be designed, implemented, installed, administered, and used too easily or too inexpensively. Thus, producing systems you can count on will always be viewed as a work in progress, rather than as a destination with a definite end point.

Technological progress alone, however, will not be enough to stimulate widespread adoption of systems you can count on. Currently, the economics of computing dictates that commodity (off-the-shelf) systems be widely used in almost all applications. Security flaws and system failures are not only tolerated, they are expected. If systems you can count on are to be widely adopted, we must create meaningful metrics for system security and stability so that these attributes can be more widely understood and their development and operational overhead can be justified. Ultimately, we must demonstrate that the development of systems you can count on is a fundamental requirement for the further expansion of information technology into our lives, our economy, and our critical infrastructure.

GRAND CHALLENGE 5.

Conquer System Complexity

Build Systems With Billions of Parts

Introduction

Today, the complexity of very-large-scale information systems surpasses that of individual engineered physical systems, such as aircraft carriers or telecommunications satellites. Their complexity approaches that of an urban community or even the economy, where technicians may understand components, but can neither predict nor control the whole system.

The complexity of very-large-scale information systems has outstripped our ability to develop and manage them. The principles of hierarchy and abstraction provide a discipline for breaking down a system into its components, each with a fixed interface. Global design and specification techniques help assure that every component of a system is specified in detail before it is built. But the continued efficacy of these techniques fades in the face of our ambitions for future systems. We will require unprecedented performance, flexibility, and ease of use as we tackle ever more demanding problems and requirements that will drive unimaginable and—using current methods—unmanageably complex systems. We are fast approaching a complexity barrier.

The limiting factor is software, not hardware. Every year, Moore's Law continues to pay its dividend, yielding ever higher raw performance per dollar. Software is still being written at the same manual pace and has progressed only incrementally since the 1950s. Software organization hits a "complexity barrier" somewhere above 10 million lines of code. Without better ways of producing and structuring software, the complexity barrier will constrain our ambitions for system behaviors.

The dampening presence of the complexity barrier is evident in the nearly universal reluctance of system administrators to modify existing functional systems, even when upgrades will better serve the needs of users. There is widespread recognition that our systems are quite brittle and respond to change in unpredictable ways. In addition, our systems have become very labor intensive and costly to maintain. Administrative expenses, made up almost entirely of people costs, represent from 60 to 80 percent of the total cost of information system ownership. We can ill afford even greater complexity.

Ironically, systems that we depend on to promote change are themselves fast becoming obstacles to change. We do not know how to design more complex systems, how to deploy them, how to administer them, or how to manage them through their life cycle. We need computing systems that are agents of change, rather than deterrents to it. There is an urgent need for new principles, tools, and techniques to build future computing systems.

The Grand Research Challenge is to conquer system complexity. Meeting this challenge requires a reformulation at all levels of computer architecture, software organization, and system design to break through the complexity barrier and create more robust systems. We must be able to design and implement systems that can autonomously adapt, maintain, repair, and heal themselves. Such innovations will substantially lower the total cost of ownership, reduce the need for intense manual supervision, and increase the future reliability and scalability of our systems.

Research Required by the Challenge

The complexity of an information system may derive from its composition. It may be a federation of complex subsystems, such as a data center that runs multiple large applications, or a federation of huge numbers of simple components. Dynamic change, which is sometimes measured in orders of magnitude, induces immense stress. Dynamic change may occur due to:

- Scaling of size to include more components, additional software functionality, or users.
- Changing requirements, changing technology, and/or changing operating conditions.
- Resource fluctuation including available power, bandwidth, latency, memory capacity, and processing capacity.

To break through the complexity barrier, we must make complex systems easier to design, easier to administer, and easier to use. Because large-scale systems are already too difficult for humans to configure, maintain, and tune, they must necessarily become self-sustaining along multiple dimensions:

- **Self-configuration.** Today, installing, configuring, and integrating large complex systems is challenging, time-consuming, and error-prone—even for experts. Large Web sites or corporate data centers are typically a haphazard accretion of servers, routers, databases, and other technologies on several different platforms from several different vendors, the full configuration and functionality of which cannot be grasped by any single human mind. Complex systems of the future will be able to configure themselves automatically in accordance with high-level policies (representing business-level objectives, for example) that specify what is desired, not how it is to be accomplished.

- **Self-optimizing.** The performance of a complex system is nonlinearly dependent on the settings of tunable parameters that relate to usage conditions. An untuned system with parameters set to their default values can perform very poorly, either individually or coupled with other complex systems. In the future, complex systems will achieve self-optimization by monitoring, experimenting with, and tuning their own parameters and by making appropriate choices about in-sourcing or out-sourcing functions to other components of the system.

- **Self-maintaining and robust.** To overcome the limitations of complexity, we must design systems capable of maintaining and adjusting their operation in the face of changing workloads, demands, and external conditions.

- **Self-healing and self-protecting.** Complex computing systems of the future will be self-healing—capable of detecting, diagnosing, and repairing localized problems that arise from bugs or failures in software or hardware. They will also be self-protecting in at least two different senses. First, they will defend the system as a whole against large-scale, correlated problems arising from malicious attacks or cascading failures that remain uncorrected by self-healing measures. Second, they will anticipate potential problems (perhaps based on early reports from sensors or components) and take steps to avoid them, or at least to mitigate their effects.

- **Self-differentiating.** Complex systems of the future will use fewer parts with pre-determined behavior and, instead, will create different behavior from similar parts. In other words, they will develop in much the same way as an organism. This will lead to vastly less expensive design and manufacturing of complex systems.

We do not know how to build self-sustaining software today. Among the research challenges are:

1. Understanding and controlling emergent behavior
2. Learning in multi-agent systems
3. Negotiation and optimization
4. Architectures and networks
5. Programming languages and computational models

1. Understanding and controlling emergent behavior

Self-sustaining behavior is characterized as emergent behavior, that is, system-wide behavior that derives from a collection of many modules of software that were designed and coded only with local functionality in mind. One crucial research challenge is to define appropriate models and techniques for understanding, designing, and controlling emergent behavior.

We need a mathematical model that relates global behavior of complex systems to the local behaviors and the interaction patterns of the individual elements. We need to understand how to exploit the relationship in order to accomplish two things: First, we need to be able to map from desired global behavior to a set of behavioral rules and interaction rules that, if embedded within the individual elements, will induce the desired global behavior. No longer can we handcraft every variant of every heterogeneous component that will fit together in the larger system.

Second, we need to be able to map from a collection of local behavior and interaction specifications to derive the emergent system-wide behavior with respect to properties such as robustness, stability, functional behavior dynamics, and performance. Thermodynamics is such a model for interacting physical particles. No such satisfactory macroscopic model exists for information systems. Whatever form such a model takes, it must permit us to define and control emergent behavior using declarative, not imperative, techniques.

2. **Learning in multi-agent systems**
   Self-sustaining systems must learn about themselves and re-optimize their behavior in the face of unavoidably large, dynamic change. This Grand Research Challenge requires that we address a research objective similar to one discussed in Grand Research Challenge 3, Provide a Teacher for Every Learner. Machine learning by a single agent in relatively static environments has been well studied and is well supported by theoretical results. However, in complex systems, individual agents will continually adapt to their environment, which consists largely of other agents. There are virtually no major theorems and only a small number of empirical results that arise from this situation. Initial forays of research into learning by multi-agent systems have tended to combine methods of game theory with machine learning, but this approach may not scale.

3. **Negotiation and optimization**
   One reason that agents must learn is to re-optimize their behavior in the face of change and new human-specified objectives. To be successful, agents must change and optimize in concert. Agents must be cognizant of each other and each other's changes to avoid unfortunate instabilities and oscillation.

   Negotiation involving two or more agents is likely to be a common form of coordination. A third research challenge is to establish a solid theoretical foundation for negotiation from two perspectives. First, individual elements require analysis algorithms and negotiation protocols that determine what bidding or negotiation approach is most effective. Second, from the system level it needs to be possible to establish what overall system behavior will result from a mixture of negotiation algorithms employed by the population of elements, and to define the conditions under which multilateral (as opposed to bilateral) negotiations among elements are necessary and/or desirable.

4. **Architectures and networks**
   Key building blocks of complex information systems include computer architecture, particularly processor architecture, and network protocols. Neither provides ways to control emergent behaviors that arise as system complexity grows. For example, processors are

unaware of any notion of faults and errors at a higher level than themselves. Network protocols have proved to be remarkably resilient in the face of network growth of orders of magnitude of components. However, today's network protocols provide essentially no support for addressing overall system security. Indeed, they stymie progress in addressing well-understood security problems. We hypothesize that both hardware architectures and communication protocols will have to provide some basic support for emergent behavior properties. Defining what that support is to address issues such as system fault-tolerance, efficient energy usage, security, and so on, is a difficult research problem.

5. **Programming languages and computational models**
   A common rule of thumb in production software is that 70 percent of the code deals with error conditions and exceptions, and only the remaining 30 percent provides the functionality expected by users. One reason for this is that fault-handling code is in each and every local code module. In complex systems with billions of components, the portion of code devoted to functionality will decrease further; most code will deal with adaptation to dynamic change, not just error handling, but overall self-sustainment.

Today's programming languages and computational models are not adequate to the task. We need new languages and models for:

- Articulating the dynamic plumbing of millions or billions of interacting components.
- Specifying fault detection and processing across the system, not individually for each code module in which faults may arise.
- Reasoning about software behavior in the presence of failures.
- Specifying the semantics of large-scale distributed synchronization and coordination primitives.
- Discovering and adapting to change, for example, harnessing newly discovered resources.
- Guaranteeing system-wide application performance.

As was said earlier, we need to develop ways to define and control emergent behavior at the system level using declarative, not imperative, techniques.

Conclusion

The preceding four Grand Research Challenges are all closely linked with societal goals. The Grand Research Challenge of conquering system complexity is more technical, but it is our capstone Grand Challenge. Addressing each of the other challenges successfully requires mastering some aspects of complexity. In other words, complexity must be at least partially addressed in order to design and build any of the systems envisioned in the other four Grand Challenges, and it must be solved if we are to build the next generation of information systems capable of realizing the Grand Research Challenges put forth in this report.

APPENDIX A.

CRA Grand Research Challenges Conference Attendees

June 23-25, 2002

Kamal Abdali
Gregory Abowd
Tom Anderson
William Aspray
Ruzena Bajcsy
Bob Balzer
Gordon Bell
Joel Birnbaum
Gaetano Borriello
Ron Brachman
Eric Brewer
Andrei Broder
Rod Brooks
Agnes Chan
Dan Cooke
Tom DeFanti
Keith Edwards
Deborah Estrin
Jim Foley
Peter Freeman
Johannes Gehrke
David Goldberg
Seth Goldstein
Leana Golubchik
Ambuj Goyal
Marti Hearst
Joe Hellerstein
Bert Herzog
Tony Hoare
Jim Horning
Mary Jane Irwin
Ravi Jain
Anant Jhingran
Anita Jones
Tom Kalil
Alan Kay
Vinod Khosla
Len Kleinrock
Larry Landweber
Ed Lazowska
Bob Metcalfe
Jim Morris
Todd Mowry
Robin Murphy
Rich Newton
Jordan Pollack
Kishore Ramachandran
Raj Reddy
Steve Reiss
Paul Rosenbloom
Jean Scholtz
Mary Shaw
Dan Siewiorek
Elliot Soloway
Alfred Spector
Bob Sproull
Jack Stankovic
Bruce Sterling
Valerie Taylor
Jon Turner
Jeff Ullman
Amin Vahdat
Andy van Dam
Uzi Vishkin
Dick Waters
David Wetherall
Steve White
William Wulf
Bryant York
Taieb (Ty) Znati

Report from the Front Lines of the Conference

This conference was all discussion and "no papers," quite unlike a typical technical conference. One formal, technical survey presentation by Ambuj Goyal and Alfred Spector initiated the first plenary meeting. After that, participants determined topic, content, and schedule. Each evening at 10 p.m., the organizing committee met to integrate the results of the day and to organize a schedule for the next day, informed by the wishes of the participants.

Splinter groups were encouraged. They self-organized, changed the course of discussion, and influenced the resulting challenge choices. One participant described this conference as being "like the ARPA principal investigator conferences of the 1970s"—that is, unstructured with stimulating debate and serious information exchange from dawn to late evening. The carefully planned recreational excursions were mostly ignored by participants, who preferred continued discussion over relaxation.

This conference proved to be a welcome occasion for leading computer scientists and engineers to exchange ideas at the interface between the social and scientific dimensions of their work—and they embraced the opportunity with enthusiasm. Debate and formal votes determined the five Grand Research Challenges described in this report, despite the efforts of the organizing committee to narrow the number down to three challenges. Our participants were convinced that by considering a more faceted perspective on the ways technology can contribute to the social good, they would arrive at a clearer perception of the research challenges before us.

Organizing Committee Members

William Aspray, Computing Research Association
Ambuj Goyal,* IBM Watson Research Center
Mary Jane Irwin,* Penn State University
Anita Jones, University of Virginia (Chair)
Ed Lazowska,* University of Washington
Dave Patterson,* University of California, Berkeley
Jordan Pollack, Brandeis University
Bob Sproull, Sun Microsystems
William Wulf, National Academy of Engineering and University of Virginia

*Member, Computing Research Association Board of Directors 2002
Grand Research Challenges website: http://www.cra.org/Activities/grand.challenges/

COMPUTING RESEARCH ASSOCIATION

1100 17th Street, NW, Suite 507
Washington, DC 20036-4632
Tel: (202) 234-2111; Fax: (202) 667-1066
E-mail: info@cra.org; URL: http://www.cra.org