2. Future Visions

Speakers: Philip Anton, Joel Birnbaum, William Caelli, Colin Crook, David Farber, Robert Frederking, Erol Gelenbe, Eric Harslem, James Kearns, Douglas Lenat

Rapporteur: Steven Bankes

A set of distinguished speakers provided an initial plenary session of the conference with brief prepared remarks on likely future developments over the next 10-20 years in their respective fields of expertise. The purpose of this session was to provide participants a common set of visions, as a form of raw material for the later deliberations. We summarize key points from these presentations here. Many of these developments and their implications are revisited in more depth in the individual breakout sessions documented in sections 4-7, below.

Technologies and Artifacts

Information and communication technologies, and the products developed from them, are now a major economic force. The trade surplus for the U.S. in information products is 10% of the U.S. trade deficit and growing rapidly.

In deliberations about future developments, it is important to distinguish between underlying *technologies*, and the *artifacts* and *services* derived from them. Current jargon often confuses these concepts. Artifacts and services are the result of one or more technologies, and connect directly to end users. It is much easier to predict advances in technology than to predict the artifacts and services that will result from that technology (and its combination with others). For example, a forecast of IT made 20 years ago in IEEE Computer magazine got the technology performance about right, but missed all the major innovations in artifacts. Similarly, given an artifact, predicting the changes in social usage resulting from the deployment of that artifact can also be very difficult. Thus, in navigating our views of the future, it is important to maintain a distinction between enabling technology, artifacts and services that may be built using such technology, and the social changes that could result from using anticipated artifacts and services.

Communications Technology

It is very hard to predict the course of the next 20 years in communications technology, but given what we know, the following comments characterize our current expectations. (For a more extended, thoughtful essay on this important subject, see the white paper in Appendix A.)

Optical communication systems

The biggest technical development is likely to be in optical systems. We expect to see commercially available optical switches within the next 4-5 years, implying that end-to-end pure optical systems will be a reality within the timeframe being considered by this conference. Combined with other innovations such as multiwave technology (in which multiple wavelengths are used to carry multiple signals on a single optical fiber), this means that huge improvements in communications bandwidth will be available over the next decade or two.

This development will motivate a rethinking of various other technologies such as the design of computer processors and the Internet Protocol (IP). For example, the ability to route packets optically at very high speeds is questionable, so huge optical bandwidths could motivate a migration away from universal adoption of packet switched protocols for all applications. It may also create the opportunities for much more distributed processor architectures as communications bandwidths may come to rival those of computer backplanes. These developments should be considered *disruptive* to communications and computing industries in the U.S. and throughout the world, in that they will cause some large companies to become obsolete, provide new niches for others, and will force computer system architectures, operating systems, and applications programs to change radically.

The allocation of spectrum bandwidth has become quite antiquated and uneconomical, but is difficult to change. Perhaps we will see spectrum reallocation, but only over a time frame of ten years or longer.

Overall, ground-based optical systems will perhaps be the most revolutionary development during this period, and ultrawideband may be a very important technology exploiting this.

Wireless

The second most important communication technology development over the next 10-20 years will most likely be an explosion in wireless communication. It provides mobility and "everywhere" access. Cellular wireless systems will increasingly be accessed by all-purpose handheld devices combining the operation of a telephone with that of a World Wide Web browser and Personal Information Manager (such as Palm Pilot). These devices will be capable of handling several, or even all, of the different protocols and standards in use worldwide, so that one standard need not be agreed upon worldwide.

Another form of wireless quickly gaining in popularity is based on the "Bluetooth" technology. It allows short-range (e.g., 10 meter) wireless communication based on radio technology. Each device has a unique 48-bit address, and can transmit at one megabit/second, with a likely doubling of bandwidth in the next generation of devices. Such capabilities permit many appliances and devices within a home to "talk" with one another, and to sense the inhabitants (e.g., if they are wearing a Bluetooth-capable badge or carrying an enabled smart card). This will again create explosive new possibilities for networking technology on an intimate scale – within the home or office – with many novel applications, devices, and services to be developed. Many opportunities will be created, enabling entire new businesses to develop and compete worldwide.

"Virtual mailbox" and "cognitive packets" technology

Some common denominators among the cheap and robust communications systems that are likely to appear in this period are the virtual mailbox and networks using what are termed "cognitive packets."

The virtual mailbox is a terminal device or accessible service that allows text to voice and voice to text conversion. It will be possible to listen to voice or text messages, dictate text messages, view pictures, listen to music, and engage in a wide range of of media interactions. An artifact may take the form of a portable device used by the more well-to-do, with some type of portable multimedia interface. Or a virtual version may be accessed by everyone through public access points -- some future version of what are now phone booths, public access computer terminals, wireless phones, post offices, or information cafes. Whatever form it takes, the virtual mailbox will have the ability to “chase down” its user (or at least be chased down by him or her), allowing connectivity for sending or receiving messages regardless of location.

These virtual mailboxes will be enabled by cognitive (smart) packet networks that are able to home packets to recipients. It will allow the sender to be informed of delivery or connection. It will support diverse services: letters, voice messages, payments, video clips, instructional material. These cognitive packet networks will put the intelligence in the packets rather than the routers. Packets will route themselves, will learn to achieve goals by exchanging information with other packets, and will allow for high levels of functionality while requiring only very cheap router hardware.

It is likely that communications technology over the next few years will be characterized by the convergence of previously separate media, as the telecom industry, the computer industry, and various media industries combine to create a single infocom industry. This industry will provide seamless data, voice, and video service, with anywhere/anytime connectivity, technology integration, and application convergence.

Software

Machine-accessible common-sense knowledge

At least some of the long anticipated goals of artificial intelligence (AI) will eventually be achieved. Existing systems are not very smart because they have so little knowledge, especially common-sense knowledge that human correspondents take for granted. Handcrafting such knowledge is very labor intensive. Systems can learn from example, but in the absence of contextual knowledge, their rate of learning is similarly very inefficient. One major example of a system addressing this problem is the “Cyc Knowledge Server,” under development by Cycorp (www.cycorp.com). The knowledge base underlying this system is being hand constructed with the goal of getting to the cross-over point where it can begin to efficiently learn for itself from human sources such as newspapers.

In designing Cyc, it was necessary to give up on having consistency in its knowledge base. Instead it is designed to be locally consistent, with its global knowledge base consisting of a tapestry of such local contexts.

In general, we expect slow progress, but progress nevertheless, toward the important but elusive goal of having a considerable body of common-sense knowledge available to computational/informational systems.

Machine Translation

Another, somewhat related, technology vital to providing a more natural user interface to computer and information resources is machine translation (MT) between natural languages. Automatic translation between human languages would have important social implications as it becomes more capable. The conference participants were told that the current state of machine translation is captured by the truism that among the three attributes of MT -- high quality, general purpose, and fully automatic -- one can have any two, but not all three at once. This tradeoff comes from the design choice between doing "shallow" translation between sentences of actual languages, or "deep" translation where the source language is translated first to some semantically explicit interlingual representation and from there to the target language. Shallow MT systems -- which tend to use dictionary lookup of words and some simple cues -- tend to be general purpose, inexpensive, and easy to use, but of low quality. Deep systems can achieve higher quality, but are more expensive and require a constrained domain of discourse. An example of a successful deep system is that built for Caterpillar, which can translate such documents as training manuals well but only if they deal with bulldozers and other specific company products.

Expected developments in MT over the next ten years include:

- the ability to conduct Web searches across documents stored in multiple natural languages
- rapid deployment of shallow MT for many "minor" languages as the need arises, and for routine business communications in major languages
- "translating telephones" for minor and limited domains such as travel agencies or help desks.

Over the next 20 years more fundamental improvements may occur in MT, but they would require fundamental breakthroughs that cannot be predicted or expected at this time.

It is important to note that the limiting factor in progress towards MT is software design and knowledge engineering, not hardware. It is therefore not reliable to predict high levels of machine translator capability based solely on projected hardware improvements.

Voice understanding is very important for ease of input for Japanese and Chinese users of computers. This technology may consequently be viewed as, to some possibly major extent, pacing the rate at which those nations adopt IT.

The foregoing suggests that over the next 20 years, MT cannot be counted on to help many developing nations enter the information age. The reason is that MT and voice understanding technologies may be robust enough to enable high-quality translation between the most-frequently-used four or five languages in the world – and then in perhaps restricted domains. But the "shallow" translation that might be available for lesser-used (and perhaps then less profitable!) languages will provide only poor and unreliable translations of the informational riches of the Web (which will be available predominantly in English into the next decade or two, at least) into these other languages.

Bio-, nano-, and materials technologies

The synergistic contributions of biotechnology, materials technology, and nanotechnology hold significant promise to impact and expand the information revolution. From a conservative perspective, advances in nanotechnology will be needed to continue Moore's Law of exponential shrinking of semiconductor scale and resulting exponential increases in computational densities. Beyond semiconductors, however, on-chip integration of logic and other components will likely include chemical sensors and components, electro-optical devices, and biological components as well as microelectromechanical systems (MEMS). These systems will also leverage smart materials that combine sensing and actuating capability, integrating processing and function at the same scale. Sensors will be increasingly embedded within systems and devices to provide seamless and unobtrusive presence. Information interfaces will also become more integrated and could even provide direct neural stimulation and augmentation or replacement of sensory organs. Molecular manufacturing may begin to have impacts by 2020 although the pace of progress is uncertain.

The synergistic effects of these technologies can perhaps best be shown through example trends and paradigms. Health care is likely to be revolutionized by genetically informed prognosis, prevention, drug selection and treatment. On-person and laboratory microsystem diagnostics will likely improve diagnostics and early detection. Smart cards will likely improve health records management and access while providing both security challenges and solutions. The whole field of drug development may be revolutionized through: better tailoring of drugs through patient genotyping; higher drug approval rates in clinical trials through better understanding of genotypic bases of disease;

reduced dependency on clinical trials through simulation; on-chip drug testing; and molecular modeling. Combined, these advances may lengthen life expectancies (especially in developed countries).

In another example, manufacturing may be revolutionized to enable customized product development at low cost as well as global part manufacturing. Advances in rapid prototyping are already combining computer part modeling with 3D manufacturing processes to facilitate customization and part specification. Manufacturing robots with improved sensors, agility, reconfigurability, and intelligence will likely facilitate manufacturing agility and reconfiguration at lower cost. Agile manufacturing could revolutionize logistics and part reserves for some industries or classes of components. It could also enable new countries/regions to acquire technologies that will allow them to develop manufacturing niches if facilitated by local resources and transportation economics.

The information revolution, then, is taking place in a wider *technology revolution* where information technology is a key component but where other technologies such as biotechnology, nanotechnology, and materials technology synergistically enable new paradigms and effects.

Computing devices and artifacts

Silicon technology continues to improve at a pace that follows Moore's Law. The number of transistors per chip doubles every 18-24 months. This means that processor performance follows a similar exponential growth trajectory. The microprocessor of the year 2000 can operate at 1 Gigahertz, and provide something like 1000 MIPS. In the year 2011, we expect 10 Gigahertz and 100,000 MIPS. This rate of improvement is expected to continue for some time yet. The semiconductor industry now estimates it will be able to achieve minimum feature size of approximately 70 nanometers (nm) by 2009. Components per chip will improve by 50% due to gains from lithography, 25% for device and circuit innovation, and 25% for increased chip size (manufacturability). Chip performance is improved 10-20% using copper vs. aluminum-copper.

The supercomputer of the year 1988 was the Cray Y-MP8/4128, which had one Gbyte of RAM, peak Mflops of 1333, weighed more than 5000 lbs and cost approx. $14M. Today, that same performance is expected of a Dell Notebook running a Pentium III, which will weigh under 5 lbs and cost under $4000.

Existing handheld devices such as the Palm VII or the RIM BlackBerry provide wireless e-mail, two-way paging, and web access for under $500. These services will be enhanced in the medium term to provide increasingly rich data access through such technologies as the Bluetooth protocol. At the same time mobile phones and pagers will add features to become increasingly “smart”. Eventually all these products will converge to create a (possibly modular) product whose overall characteristics and features cannot be fully anticipated. And the possible implications for user behavior are presently boundless!

These trends suggest that the following assumptions about the coming 20 years seem reasonable:

- There will be robust global information utilities with widely available service
- A variety of pervasive devices will be available, including wireless interfaces
- Information “appliances” will have intuitive interfaces accessible by many
- A large fraction (but very likely not a majority) of world population will become connected in one form or another
- Computing power will continue to rise according to Moore’s Law, or perhaps faster as concurrent architectures accelerate the gains from chip speeds alone
- Similar exponential increases will be seen in communications bandwidth
- Bio-sensors will be employed to identify users and detect and monitor other aspects of the environment
- Smart storage technologies will allow the personal high-capacity non-volatile storage of information

Given these assumptions, a variety of conclusions seem to follow:

- Sentient artifacts will make access to information technology and services much simpler and require much less sophisticated users
- Inexpensive artifacts imply very large user populations
- Artifacts will disappear into the environment, and the focus will move from devices to services with the central issues being utility, security, and the supply chain

- This reverses where the action happens, as large communications bandwidth and smart storage-enabled devices create a wide range of new engineering tradeoffs, such as between communications bandwidth and computation.

But, in spite of the rapid pace of technical progress, it may be difficult to overcome gaps in infrastructure to bring the have-nots into the game.

E-commerce and beyond

Information technology will create huge changes in commerce in the coming decades:

- Customers (including both businesses and end-users) will have very significant technical capabilities
- There will be billions of such customers
- There will be intense price competition
- New "value propositions" will be created – ways of creating value by providing wholly new services and facilities
- There will be increasing customer control and choice in market transactions
- Business must increasingly appeal to customer aesthetics.

Billions of customers and millions of businesses all interacting suggests the information economy will be a very complex system. Already, the Internet shows the statistics associated with classical "complex systems:" a power law dependency on site hit rates, links, and languages. This suggests that this economy will be characterized by the non-linear effects of such complex systems including lock-ins (of customers to specific sites) and the support of a diversity of small niches.

Such a system can be a rich domain for innovation, with a constant creation of new rules and structures. Napster may be a good example of how such a new paradigm can emerge, "from the bottom up."

Increasingly, the physical world may be forced to align with the virtual world that coordinates it. The tension between real and virtual may vary sector by sector, but it is anticipated that the virtual may dominate, except in emergencies where physical constraints will still be central. (For example, logistics folks must increasingly perform to Star Trek-inspired standards.)

This will result in a swirling mix of old and new economy, with customers as adaptive agents whose behavior can be shaped, and businesses as virtual enterprises. A variety of evolutionary and revolutionary enterprises can be anticipated. A mix of individual-business-government interactions will create a web ecology, with accompanying new social tensions. New social “things” will emerge.

Current e-commerce reflects a tool-building stage in a historic progression. In the longer term the prospects for fundamental change to business is real. Establishing trust is crucial to enabling business to business e-commerce. It enables transactions among partners to trusted relationships. Consequently, the set of concerns that often are given the label “privacy” could prove an important constraint that could impede the realization of these technology visions.

We speak easily of technology and usage of infrastructure, but what are the crucial (virtual?) artifacts that must be created to get to there? Computer Science research strives for “perfection”, but much of our daily experience is getting work done in spite of how bad the software is. Is it technology or engineering that drives/constrains the future?

Commoditization of interactions could serve to minimize the need for trust. Most of the interactions occurring in cyberspace, especially related to e-commerce, are in fact instances of a rather limited set of transactions: making requests or offers; making promises; fulfilling promises; making assertions. It is possible that use of greater precision and clarity in establishing such transactions could help enable systems establishing appropriate levels of trust within each type of transaction.

The availability of multiple personas for an individual person when interacting in cyberspace may allow us to deal with privacy issues in innovative and unprecedented ways.