March 28, 1996

Telecommunications in the 21st Century

David G. Messerschmitt

Department of Electrical Engineering and Computer Science

University of California

Berkeley, California 94720

Invited paper in the special section on “Dreams of Future Communications” of the Institute of Electronics, Information, and Communication Engineers English Transactions (Japan), January 1993

1.0 SUMMARY

The term telecommunications is derived from “tele”, meaning at a distance, and “communications”, meaning exchanging of information. The history of electronic communications has thus far been applied to the exchange of spoken, visual, and or textual information between pairs of people, pairs of machines, and people and machines. The role of telecommunications has been to provide a medium for the exchange of the information, with the burden placed on the communicating people or machines to initiate the communication and to interpret or process the information being exchanged.

In this paper we attempt to predict some future trends in telecommunications, reaching into the next century. Such predictions are inevitably incomplete, inaccurate, or both. Nevertheless, it is a useful exercise to try to anticipate these trends, and more importantly the issues and problems that will arise in the future, as a way of focusing near-term research efforts and suggesting opportunities. One of our hypotheses about the future is that telecommunications networks will become much more active in initiating, controlling, and participating in the exchange of information.

Our approach will be to first review some particularly important past developments, and then to try to predict the future in two ways: First, by extrapolating present trends and activities, and second, by criticizing current trends and anticipating problems looming on the horizon.

2.0 The Past

2.1 Digital Telephony

The single most significant advance in telecommunications technology in the past few decades has been the rapid evolution from analog to digital representations of signals. While this evolution has been driven in part by the integrated circuit technologies, also significant is the “regenerative effect” of digital representations; that is, the ability to store, copy, and retransmit the representation with an arbitrarily small degradation. As a digital representation, written language and later printing were much earlier applications of the regenerative effect, and had a tremendous impact on civilization. Digital communications has succeeded in expanding the scope of the regenerative effect beyond written representations of language and data to all modalities of communication, including speech and images, with the minor penalty of an insignificantly small quantization degradation.

Digital representations have also had the practical benefit of allowing compression by redundancy removal (this is a side benefit of the regenerative effect since degradation of the compressed version has much more impact than degradation of the original). Compression has accelerated the application of digital representations, since it enables more efficient use of scarce resources such as radio spectrum.

Another significant impact of digital representations has been the abstraction of information representation. All digitally represented forms of information, consisting of a bit stream, look very similar from the perspective of storage and transmission. “Integrated networking” and “multimedia computing” are both based on this common abstract representation of all forms of information.

2.2 Stored-Program Switching

The replacement of human operators and the development of intelligent switch controllers has been particularly significant in reducing costs and enabling advanced features. While the initial efforts were simply to automate the setting up of calls, and operations and maintenance functions as well, the incorporation of computer technology into network control is a very important development for the future. Any network control functions that can be conceptualized and implemented in software can be realized. Later, we expect this will lead to an active participation of the network in the communications.

3.0 The Future as an Extension of the Present

Some predictions about the future of telecommunications can be made by an extrapolation of current commercial trends and observation of what is currently happening in research laboratories.

3.1 More Bandwidth, Processing Power and Memory

The field of electronic communications has been extraordinarily successful at enabling point-to-point communications at high information rates, great distances, and low cost. Making reasonable extrapolations of current trends, the 21st century will be marked by an extraordinary growth in available broadband networking, making services such as high-resolution video conferencing routine and greatly increasing the possibilities for human interaction in social, political, and business relations.

While fiber optics has had a remarkable impact on the availability and cost of bandwidth, at least in principle if not yet in reality for the average user, the costs of processing and memory have also been advancing at similar rates. In most applications there is a direct trade-off between these three technologies: bandwidth, processing, and memory. For example, at the expense of processing we can conserve bandwidth to a remarkable degree by compression, at least in the case of speech/audio and images/video. Less widely recognized is the impact of memory technologies [2]. With a broadband network and sufficient memory at the destination, an audio or video presentation can be transported faster than real time. Or interactive network delay can be avoided in accessing a database by first transporting the entire database. If delay is not an issue, large files can be transported in the mail on a compact disk, perhaps at much lower cost. As yet another example, during a presentation the viewgraphs can be compressed as video and transported across the network, or they can be transported in a Postscript-like form at the beginning of the presentation and regenerated locally at the receiver site under the control of the presenter.

The point here is that there are alternatives available to the user, alternatives that don't necessarily entail broadband transport. From the user perspective, the relevant parameters are the total information to be transported, how soon it is needed, whether access is interactive, and the total cost (including bandwidth, processing, and memory). Note that we do not include subjective quality here, as we assume coding impairments will be negligible. The user will choose whatever means give the best trade-off among these parameters.

The focus of telecommunications in the past has been on real-time interactions, such as a voice call or video conference. This has led to a design culture in which throughput is the relevant performance parameter, and achieving high overall utilization of facilities subject to moderate latencies is the design goal. In contrast, in the local-area network (LAN), the bandwidth is typically allocated not in terms of high utilization, but in terms of the latency for transfer of large files or executables. The LAN has proven that customers are willing to pay for broadband transport because of its ability to provide low latency, even where their long-term throughput requirements would not justify that bandwidth. In conceptualizing the market for broadband networks, our telecommunications culture has driven us in the traditional direction of trying to achieve high utilization by continuing to emphasize real-time interactive services, trying to minimize the peak-to-average-bandwidth ratios, etc. I believe this to be a mistake, because it runs directly counter to the current commercial usage of broadband networking. My prediction is that in the 21st century broadband networks will be judged by the customers more in terms of latency for large transfers, and those networks that are the most cost-effective in providing that transfer will be the most successful commercially.

Unteathered access via microcellular radio and infrared systems will imply transparent access to telecommunications resources without regard to location. As has been pointed out [2], we are in a process of reversal in the transport of different services. For good historical reasons, services like voice telephone and data access that are naturally mobile have been delivered by wired media, while services like television that are inherently fixed-location have been transported by radio. Given a modicum of good sense on the part of the regulators, this situation will reverse itself in the 21st century. From our current state of knowledge, wireless access will have inherently limited aggregate bandwidth, giving a few users in each location a lot of bandwidth or a lot of users a little bandwidth. This appears to be for fundamental reasons, so that we can expect bandwidth to be a limitation for wireless access in the 21st century. Fortunately, as mentioned above, for most purposes added processing and memory are effective in overcoming this limitation. Thus, expanding the range of feasible processing, memory, and display technologies subject to the limitation of the battery technologies will be an increasingly important topic in the 21st century.

3.2 Telecomputing: An Industry for the 21st Century

It has been widely recognized for some time that there is a strong symbiotic relationship between computing and telecommunications. To date, however, that relationship has been manifested largely by the importance of computing in implementing networks, and the importance of networks in connecting computers to their peripherals. Sometime in the 21st century, telecommunications and computing will not be considered separate industries as they are today, and thus it is impossible to speculate on telecommunications in the 21st century without also speculating on 21st century computing. The industrial organization will more likely be split between terminal, application, and transport functions, where each is an essential component of both computing and telecommunications.

There has to date been an underlying intellectual distinction between computing and telecommunications. Telecommunications has emphasized real-time synchronous interaction, such as conversational voice and video conferencing, where a hard upper bound on delay was necessary. In contrast, computing has focused on non-real-time asynchronous communications, such as electronic mail or shared database access, where a lower communications latency is desirable but there is no hard upper bound. This has had a profound influence on the design of the respective systems, manifested for example in circuit vs. packet switching. Now this distinction is disappearing. Computing calls it “multimedia”, and telecommunications calls it “integrated networking”, but in both cases real-time and non-real-time services are combined. In the case of computing, the motivation is to mix different forms of representation (voice, image/video, text) within a single application. In the case of telecommunications, the primary impetus is to counter the proliferation of networks, but integration of different representations in the same service is an important side benefit. The net effect is that it is becoming increasingly difficult to draw intellectual distinctions between the design activities in the two industries.

But the integration of computing and telecommunications goes much beyond that. The computer originally arose out of the desire to automate large computations. However, the connection of computers by networks will fundamentally alter the role of the computer in the 21st century. The computational capability will be supplemented by the ability to access information on a vast scale, without regard to its physical location. While computation remains a major application of computing, increasingly the dominant role of the computer is to assist in the management and access to information. In actuality, this is largely a telecommunications function, not computing in the classical sense. We can identify a couple of reasons that the marriage of these two technologies is so important:

- *Spatial coherency of information*. In many human enterprises, there is the need for a logically-centralized database which gives the same answer when queried from different locations, and which can be updated from different locations. Take for example, an airline reservation system, that must accept updates from anyplace, and provide consistent information everywhere. The emergence of such databases has had an incredible impact on the productivity of transportation and financial industries, among others.
- *Temporal coherency of information*. As the rate of change in both human enterprise and human knowledge has increased, traditional means of dissemination of information such as printing technologies become inadequate because of the delay they introduce. As the delay in the dissemination of information is reduced, the rate at which knowledge advances is correspondingly increased, since knowledge generation depends on previous knowledge.

Both the spatial and temporal coherency of information are communications issues. One of the key roles of electronic and photonic telecommunications is to disseminate large amounts of information quickly and over large geographic areas.

The merging of the telecommunications and computing industry will be so complete that, for purposes of this paper, we might as well coin a name for this new field. I propose *telecomputing*, which is a concatenation of *telecom*munications and com*puting*.

3.3 From Efficiency to Complexity Management

The past development of telecommunications technology has been largely driven by efficiency considerations. The dominant goal of transmission has been to multiplex larger and larger numbers of telephone conversations into a given bandwidth, be it cable or radio media. This has led to a large effort in signal processing associated with compression of signal sources and higher spectral efficiency on bandwidth-limited channels.

Efficiency will continue to be important in some applications. The principle example is the finite bandwidth of the radio spectrum, which can be mitigated to some extent through frequency reuse. However, it is also clear that an increasing portion of technological effort will be devoted to complexity management, rather than efficiency [1]. This is driven by two complementary factors: First, the declining cost of electronics and photonics hardware leads to increased structural complexity at a modest cost, and second, the applications become less cost-sensitive as they penetrates almost all aspects of our commercial endeavors. The first technical community to be preoccupied with the difficulties of complexity management was computer science and its subfield software development, because software is a conceptual task (largely divorced from physical constraints) that rapidly exhausts the organizational capacities of enterprises that attempt it. A second familiar example was integrated circuit development, which was forced to abandon efficiency as the dominant consideration in favor of structured and automated design approaches as the complexity of the designs rapidly increased.

Telecommunications is encountering the challenges of complexity management today. This occurred first in call processing, because it is largely a software development task. But today we see complexity management entering telecommunications in much more fundamental ways, as we begin to integrate a large number of different services and applications. That complexity management became a dominant issue in call processing associated with only a single very simple service, the 64 kb/s circuit, is indicative of how difficult it will be to manage networks that provide thousands of different services. We will be forced to largely abandon efficiency considerations in favor of inherently inefficient structured approaches to service provisioning. Fortunately, this is not inconsistent with localized efficiency, such as in the radio portion of a connection. Further, the declining costs of hardware technologies and the broadband fiber optic medium allow us to compromise efficiency without abandoning cost effectiveness.

The object-oriented software methodologies have proven very effective for structured implementation of large software systems. The essential idea is to group data and functions modifying or accessing the data together into objects, forcing all external interfaces through the functions. Object-orientation has yet to be applied to distributed systems, in the sense of transporting objects rather that just the data contained in the objects. Since executables tend to be very large, broadband networks are required to transport objects in a reasonable time. Object-orientation is thus an example of a complexity-management technique that will prove to be a major impetus for broadband networks.

Perhaps a telling example from another venue is the economic system. The market system of commerce has proven the most successful, at least among the systems that have been tried, at making overall economic progress. In the market system we can clearly see major inefficiencies: parallel development of similar projects in different organizations, large bureaucracies to regulate and prevent abuses, unemployed workers. And yet a more centralized form of management of economic activity, which can easily eliminate these inefficiencies, has proven itself unable to cope with the inherent organizational complexities of the process. We can expect that telecomputing will have to face similar challenges, and will arrive at similar distributed organization consisting of autonomous entities negotiating among themselves. Such an organization necessarily introduces inefficiencies, since such negotiation necessarily lacks global knowledge.

3.4 Changing the Role of Standardization

The past focus of the telecommunications industry has been largely the packaging of universal services with widespread interest, and the standardization of those services. Let us critically examine each of these aspects.

The view that the telecommunications network provides a direct service to the end user dates from the very invention of the telephone: Rather than providing a voiceband channel, the industry enabled users to talk to one another. Today the interest in new services continues, including such things as enhanced call-management functions and video conferencing. I would submit that there are very few “universal” services. Rather, most services that might be conceptualized are of interest to smaller groups of people, and yet the aggregate of such services may represent a very significant commercial opportunity.

If we look to the computer industry for inspiration, we see that the concept of the packaged application has been discredited; for example, the stand-alone word processor. In addition, there are very few “universal” applications, such as word processing. Rather, the set of applications is highly fragmented, and most of these applications were not conceptualized or implemented by computer hardware or service companies (would a hardware company or timeshare vendor have invented the spreadsheet?). These fragmented applications are developed by small vendors, or by the users themselves. The result is a very dynamic and effective system for innovation.

The essential observation is that computer companies generally focus on the platform for application deployment, such as the CPU, peripherals, and operating system. They encourage other companies to specialize in developing applications. This process has proven so effective in meeting the needs of specialized user groups, and in speeding the rate of innovation in the computer industry, that it will inevitably become the dominant model for telecommunications services in the 21st century. The telecommunications service providers and their vendors will focus on providing the platform for services, such as transport, call-switching functions, and an open-system signalling interface (such as ISDN already provides). Other companies will focus on developing a wealth of specialized services based on telecommunications and computer platforms. Telecommunications companies will benefit because of a much higher velocity of new services, with an attendant increase in volume, and users will benefit from a staggering variety of available services, and even the ability to develop their own.

All this implies that telecommunications service providers will continue to lose control over the applications for which their networks are used, and will derive only a portion of the total service revenue. Nevertheless, they will benefit greatly from the transport revenues generated by rapidly increasing traffic, just as they derive substantial revenue today from calls completed to answering machines (that would otherwise not answer) and facsimile machines (which often require new telephone lines).

Related to this issue is standardization. The historical philosophy of the telecommunications industry is that since two or more users must participate in a service, it must be standardized in its entirety. Standardization is in some ways a tremendous impediment to progress. Aside from the delay in deployment of new technology it introduces, it often results in technical solutions designed by committee without the benefit of implementation or user input, and results in solutions that sometimes don’t satisfy user needs. Fortunately, in the 21st century most aspects of services, even signal-processing-based services like video, will be software-defined, and this offers the opportunity to bypass much standardization. Software-defined aspects of a service can be downloaded from a central repository (or from one of the participating users) over the ubiquitous 21st century broadband networks, avoiding standardization of the associated functionality. Of course standardization is still required at the basic level of the underlying transport mechanisms and the language used to describe the service, and this standardization will prove to be challenging. The goal will be to enable new services to be implemented, tested, and deployed without modifications to the platform (transport or centralized control entities) and without the delay of an intervening standardization activity.

Telecommunications has also been plagued in the past by the “community of interest” problem. Who among your friends will be the first to purchase their video conferencing set, if they have no one with whom to conference? Here again, the advance of hardware technology, allowing software-defined services, will be of tremendous benefit. If a service definition can be downloaded from a central point to platforms owned by each user participating in the service, a large community can immediately participate in a new service. The incentives for vendors to invest in such services will be dramatically increased.

Transporting objects over a network, as mentioned earlier, can provide a conceptual framework for distributing service definition over broadband networks, as well as for implementing software-defined distributed telecommunications applications. One current role of transport standardization at the application layers is to define the content of the bit stream passing over the network. One conceptual framework for eliminating this standardization is the distribution of objects over the network. Just as one of the motivations of object-oriented programming is to eliminate the standardization of data structures, standardizing instead the functional interfaces, it can also eliminate the detailed standardization of the bit stream being transported by a network.

Telecommunications has also been plagued by very long depreciation intervals, with the undesirable side effect that all new equipment has to be compatible with very old equipment. A higher velocity of new services will force a faster obsolescence of hardware and software components. Both service providers and users will see clear benefits to replacement of components, and suppliers will find larger markets in replacement products. If for no other reason than this, suppliers and service providers should be very supportive of the platform approaches to service provision.

4.0 Some Future Challenges for Telecomputing

What might happen in the future that is not simply an extrapolation of the present? A useful starting point is to view the present critically, identifying shortcomings and problems with the current directions. There are five serious challenges that we can see looming, and the 21st century will have to face these challenges squarely. They are so serious, and inherently challenging, that we will call them *metachallenges*.

4.1 Metachallenge One: The Ownership of Information

The regenerative effect of digital representations made an important impact on society through the printing press, and this impact is only magnified by the temporal and geographical coherency of information enabled by telecomputing. It will enable us to capture not only text and data, but also virtuoso musical performances and the like in a form that can be transferred to future generations virtually without degradation.

In spite of its benefits, the regenerative effect creates a serious societal problem. The market system of commerce that has evolved is based on the concept of ownership of property. The essence of ownership is the ability of the owner to control the use of property. Information (where we include such things as software and audio and video performances) will be an increasingly important commodity in world commerce, and yet the privileges of information ownership become difficult to exercise because of the inherent ease with which it can be copied.

When ownership becomes ineffective, the market incentives to generate and enhance information are negated.

This problem arose first with the photocopier machine, but becomes more critical when information is represented in electronic form and networks and computers become more prevalent. Attempts to address this problem thus far have been *ad hoc* and largely ineffective [1]. I hope that the 21st century will see solutions to this problem, and I believe that telecommunications will play a key role. Ownership implies that all access or use of information can occur only with the permission of the owner. There are encryption techniques to insure that this permission is required, and ubiquitous worldwide networking will provide the means for users to request and owners to grant this permission.

Ubiquitous networking will also enhance the economic efficiencies of information production and consumption in other ways. The most effective model for the pricing of goods has been based on the utility to the consumer. Most goods are therefore priced not only on the basis of their functionality or application, but also based on usage. Easy examples are commodities like electricity and long-distance telephone service, but this applies to capital goods (like automobiles and appliances) because of their finite and fairly consistent useful life in relation to their replacement cost. Information violates this basic economic model because of its infinite lifetime (the regenerative effect) and very easy means of copying. For example, packaged software or books are normally assigned a fixed price independent of usage, and even that pricing is difficult to enforce due to the ease of copying. The result is a disincentive to the occasional user, while the major user pays too little (a disincentive to the software developer or author). Fortunately, ubiquitous networking, through its ability to enable the owner of information to control or monitor its access, will enable owners of information to derive revenue from all users, and to do usage-based pricing. We see this today in centralized information services, but in the 21st century this will extend broadly to much information, including that stored locally or freely copied by the user. Those who have grown accustomed to the “free” exchange of information will not be pleased, but it is necessary.

4.2 Metachallenge Two: The Filtering of Information

As mentioned previously, the temporal coherency of information afforded by global telecommunications, and other factors as well, result in a rapidly increasing velocity of information generation and accumulation [1]. We are all experiencing a rapid buildup of knowledge in our fields, perhaps increasing even geometrically. On the one hand the more rapid dissemination of information results in faster accumulation, as the time between an advance and its application or further development is shrunk, but on the other hand the increasing time we all spend keeping up with developments (rather than making new contributions) impedes progress. A geometrical increase in knowledge, coupled with the presumably limited capacity of each individual to absorb the knowledge, may quickly become the limiting factor to progress. More likely, there will be increasing fragmentation of fields through specialization, which itself impedes progress. This is not dissimilar to the multicomputing problem: Adding more processors does not necessarily increase computational throughput, and may actually decrease it!

The combination of the information and communication explosions will become an increasing factor in limiting individual as well as global productivity. Thus, in the 21st century, technological solutions that mitigate this problem will become increasingly important. Not to imply that this problem can be completely overcome, but rather the boundaries of feasible collective activity in light of personal limitations (and quality-of-life expectations) can be expanded considerably.

A fundamental problem here is the presumably limited personal resources of each individual in the face of an endeavor that necessarily spans many individuals. One way to mitigate this is by expanding the capabilities of each individual through the use of tools, thereby decreasing the number of individuals required. There are many examples of this already. The number of people required to design an integrated circuit has actually decreased (especially if efficiency is compromised) as its complexity has increased, due to the impact of computer-aided design. To date, relatively straightforward parts of the tasks, things that would scarcely intellectually challenge the most capable human, have been susceptible to automation. Extending this capability dramatically is largely in the domain of “artificial intelligence”. The accomplishments of AI have fallen far short of expectations to date, and I personally don’t expect machines to match the intellectual capabilities of humans in the 21st century. (In fact, I secretly hope that they don’t!) However, dramatic progress will be made, and the power of the individual will be greatly expanded by this technology.

Another aspect of this problem that is amenable to technological attack is information filtering. While the information that each individual must access expands rapidly, the universal knowledge base from which this must be extracted expands much more rapidly. The problem is accentuated by the increasing complexity of our systems (technological and economic), resulting in a more interdisciplinary approach to system design, and a corresponding increase in the relevant information and knowledge base. A clear inefficiency in the process is the manual filtering we must all do to cull the information we want from the mass of information we don’t absolutely need. In the 21st century we will see a telecomputing infrastructure that not only enables easy and almost instant access to information, but further tailors the information presented to each user to their own interests and needs. Such a system, to be fully effective, again reaches into the domain of artificial intelligence, because the capabilities needed are vastly more sophisticated than the “keyword search” capabilities now common. The essential challenge is that the system must understand the information, not simply recognize it, to be able to effectively filter it [3].

4.3 MetaChallenge Three: The Filtering of Communications

The increasing globalization of research and economic activity, largely enabled by modern telecommunications and transportation technologies, dramatically increases the number of people with whom we may have occasion to communicate. A classic problem in many large engineering systems with concurrent access to finite shared resources is “thrashing”, where the volume of activity associated with the arbitration of resource access grows to become so dominant, and the actual access to the resource actually shrinks. We see this phenomenon in computer systems, in communication networks, and in transportation systems, among others. Surely if we consider a worldwide collection of people communicating among themselves and with machines, the individual people represent finite resources that easily become swamped. Emerging technologies like personal communications are a double-edged sword: While they increase economic efficiency by enabling communications, they also impede personal effectiveness by generating constant interruptions, and in their extreme probably decrease the quality of life.

The most effective individuals are those who consciously prioritize tasks, prioritize them most appropriately, and perform only as many as they can do effectively. A similar principle applies to communications: As the volume of communications becomes burdensome, the most effective response is to prioritize them and participate only in the most important. The goal will be to enable desirable communications to occur with the least obtrusiveness, but to allow the individuals to control and prioritize them. The current philosophy in the design of telephone networks is to make it as easy as possible to connect to another individual. In the 21st century, this philosophy must change.

Biological systems offer some insight into more effective ways to organize man-made enterprises. Looking at a complex organism like man, we see a hierarchy of control mechanisms, from the conscious or perceptual at the top to the autonomous on the bottom. At the risk of over-simplifying this elegant picture, we can divide the control mechanisms into several layers:

- *Conscious* behavior (response to external environment), which includes all the things that we have to explicitly think about in our day-to-day life.
- *Subconscious* learned (not innate) behavior, such as forming words or reaching for an object, which is initiated consciously but which does not require detailed conscious control.
- *Autonomic* (not learned) basic functioning of the organism. This system has two parts: the *sympathetic* system and the *parasympathetic* system, where the former has the role of increasing resources (increasing blood supply or pupil diameter) and the latter reduces resources (decreasing blood supply or pupil diameter). This system also directly influences conscious behavior, for example through hunger.

Overall the organism operates in a hierarchical fashion: It survives because of the autonomous functioning, it is adaptive enough to learn a set of routine subconscious behaviors that are appropriate for its environment, and it deals with non-routine situations at the conscious level. The role of conscious control of behavior is seemingly minimized, and as much functioning as possible is relegated to the subconscious. It has been observed that one design goal for good computer human interfaces is to make them *unobtrusive* [4]. I would suggest that one aspect of this is to allow the user to subconsciously internalize as much of the interface as possible. Another aspect is to make it “intuitive”, which can be interpreted as conforming more closely to previously-learned subconscious behavior (like the desktop metaphor of windowing systems).

Almost all communication and interaction with other people, verbal or written, desired or not, is performed at the conscious level. Communications is inherently obtrusive, resulting in too much time spent or too many interruptions. A goal of telecommunications in the 21st century should be, as has been proposed for computer systems [4], to move toward becoming less obtrusive. In our view, the way to accomplish this is by making subconscious or even autonomous an increasing set of behaviors that are currently conscious, while reserving the most appropriate and desirable communications for conscious attention. More specifically, we can identify two specific needs: The prioritizing and filtering of communications, similar to the filtering of information, and the autonomous scheduling of synchronous interactions (covered in the next section).

The goal of communications filtering is to allow the individuals involved to prioritize their potential interactions, encouraging those that should occur and, without conscious effort, block those that are lower priority. Superficially this looks similar to information filtering, but in fact it appears to be much more difficult. In information filtering, the individual expresses a set of priorities, which are implemented by an autonomous agent by scanning the available information looking for those that meet the criteria. In communications filtering, there is still the context of the proposed interaction (similar to information), but actually two or more sets of priorities that matter, those of the originator(s) and those of the recipient(s). I might be able to structure priorities based on the identity of the originator, for example ruling out all salesmen, but for most originators my priority would be based on both the identity of the originator and the context of the proposed communication. In fact, the latter is likely to be more important than the former.

The filtering of communications must involve a complicated process of negotiation between recipient and originator, or their autonomous agents. It may be more accurate to call these agents subconscious, rather than autonomous, since they must embody adaptable learned behavior. The same is true of information filtering: The priorities and interests of any one individual are not static, but depend dynamically not only on their innate interests but also on the context of the larger society. If a new body of knowledge arises, a subconscious agent can ask the individual if they are interested, or can pass some information to the individual and watch their behavior.

Both information and communications filtering are natural functions for the network. Information filters will potentially gather information from a number of sources, which are naturally available to the network. Communication filtering involves a negotiation among users of the network or their autonomous agents, and again the network is the entity that naturally has available the requisite information. Of course, both information filtering and communications filtering could be performed by human personal assistants; in fact, this is a common approach for many busy individuals today. However, to expand this model will result in more and more people as personal assistants and fewer and fewer primary contributors. Just as the telephone network had to automate call setup (or else every man, woman and child would have to be a telephone operator), likewise the personal filtering agent function must be partially or (for some people) totally automated.

4.4 Metachallenge Four: Scheduling of Synchronous Interactions

The very fact that we make subconscious or autonomous many routine administrative functions should in itself reduce the volume of communications. An additional way that many individuals today reduce the tyranny of interruptions is to rely on asynchronous communications modes, like electronic mail or voice mail. This allows the processing of communications to be more efficient and less obtrusive, by concentrating it in particular parts of the day. However, this also leads to considerable inefficiency. A complex task or interaction, one that requires many back-and-forth exchanges, can occur much more expeditiously if the latency in the exchange can be reduced. Many of us have the experience of trying to conduct such a conversation by electronic mail or voice mail, but that same interaction could be performed much more effectively, consuming less time, and involving fewer interruptions, if the parties to the interaction could simply converse in real time. But the use of voice mail greatly discourages those synchronous interactions, because it encourages people to seldom answer their phone! Taken to its extreme, where neither party ever answers the phone but only initiates return calls, no synchronous interaction is feasible.

The role of communications filtering is to limit interactions to those that meet appropriate criteria (the biological parasympathetic function), but it does nothing to enable synchronous interaction, if that is appropriate (the biological sympathetic function). Globalization of activity accentuates these problems by reducing the range of candidate times for synchronous interactions across different time zones. Scheduling such a synchronous interaction becomes in itself very obtrusive, often involving conscious asynchronous interchanges between the parties to find an acceptable time for synchronous interaction. Fortunately, the scheduling of synchronous interactions is particularly amenable to automation. The goal is to autonomously initiate these interactions at mutually available times, without the burden of interactions solely for scheduling purposes. Like communications filtering, this is a natural function of the network, because the network is logically connected to all parties whose time must be scheduled and the network is the entity that will initiate the synchronous interaction at the mutually available and acceptable time.

4.5 Metachallenge Five: Integrating People and Machines

The history of technological innovation and industrialization has been to relegate an increasing set of tasks to machines, and simultaneously advance the skills and tasks performed by people. This occurred first in physical tasks, and today is happening in conceptual and administrative tasks. Technological history would suggest that the role of people is not diminished in this process, and in fact becomes more critical as people assume increasingly higher-level skills and functions. However, the role of people is diminished when viewed in the larger system context; that is, as an increasing fraction of the work (albeit the less sophisticated work) is performed by machines.

To date we have largely viewed telecommunications and computing technologies as tools for people to use in accomplishing tasks, like writing or designing. This was certainly an accurate perspective for physical tasks in the industrial revolution, because people provided all the brainpower for the system functioning. The most important design element of these tools is their human interface. But in the modern telephone network we see an example of a fundamentally different entity arising, a large system that operates largely autonomously from people, and performs an enormously important function for society. To be sure, people have to get involved in solving the really difficult problems, but the routine operation and even most of the maintenance of the system is performed autonomously. The system is even beginning to converse with its human users using speech.

When the system begins to take over increasingly less routine functions previously performed by people, such as the manipulation and disposition of the greatest portion of information flowing through society, then the view of the system as a tool will be increasingly less appropriate. Rather, in the 21st century we will come to a systems approach, in which machines and technology will come to be viewed much more as equal partners in the functioning of society. Of course, the humanists will resist this view, and perhaps never come to accept it. And of course they will be right from the perspective that all the systems, technological and human, were created by and for the benefit of people. But from a design point of view, systems engineering will need to take a different perspective from today. Both humans and machines will have to be viewed as complementary elements of a complex functioning machine. The metachallenge will be to identify those unique roles for humans that leverage their unique capabilities, in the context of the larger societal system function.

The last three metachallenges that were described are clear illustrations of this. They all describe a future in which, as their capabilities increase, our machines take over a larger and larger set of functions, while freeing people to consciously deal with that they do best. The challenge to the designers is to design machines that subsume an increasing number of sophisticated tasks, while integrating people into the system in their unique roles. The challenge to people is to continue to upgrade their skills as more routine functions are increasingly automated.

5.0 Conclusions

We have taken a look into the future of telecommunications as both an extrapolation of the present activities and as a response to some looming problems. We can summarize a few major conclusions of this exercise: The organization of the telecommunications industry will be largely merged with the computer industry in the 21st century. This has been anticipated and discussed for some time, but there have been some important distinctions between the two fields that are now finally disappearing. The declining cost of hardware and an increasingly dominant role for programmable solutions will result in a substantive shift in the role of standardization. New innovations in telecommunications services will occur dramatically more rapidly in the future, and there will be a proliferation of a great variety of specialized services. While the telecommunications system has been largely passive, responding to the requests of individual users, in the future it will become a much more active and autonomous entity. It will not only enable communications among individuals, but actively facilitate them by taking into account the interests and schedules of those individuals. Also, the network will participate in the determination of what communications actually takes place, actively discouraging that which is unwanted or unnecessary.

The references below discuss some of these issues from a different perspective, and they all make interesting reading.

6.0 References

1. Robert Lucky, *Silicon Dreams*, New York: St. Martin Press, 1989.
2. Nicholas P. Negroponte, “Products and Services for Computer Networks”, *Scientific American*, Sept. 1991.
3. Lawrence G. Tesler, “Networked Computing in the 1990s”, *Scientific American*, Sept. 1991.
4. Mark Weiser, “The Computer for the 21st Century”, *Scientific American*, Sept. 1991.