INFORMATION TECHNOLOGIES FOR ASTROPHYSICS CIRCA 2001

**A Position Paper for the
NASA Workshop on Astrophysical Information Systems
May 23-25, 1990**

by
Peter J. Denning

Director
Research Institute for Advanced Computer Science
NASA Ames Research Center

ABSTRACT: It is easy to extrapolate current trends to see where technologies relating to information systems in astrophysics and other disciplines will be by the end of the decade. These technologies include mineaturization, multiprocessing, software technology, networking, databases, graphics, pattern computation, and interdisciplinary studies. It is less easy to see what limits our current paradigms place on our thinking about technologies that will allow us to understand the laws governing very large systems about which we have large datasets. Three limiting paradigms are: saving all the bits collected by instruments or generated by supercomputers, obtaining technology for information compression, storage, and retrieval off the shelf, and the linear model of innovation. We must extend these paradigms to meet our goals for information technology at the end of the decade.

I have been asked to present some speculations on technologies that will be available to us in eleven years just after the turn of the century. I have even been asked to be "visionary"! I will indeed spend a few minutes telling you what I see.

Speculating is for me a pleasant and straightforward task. We can look for impressive developments in hardware, software, networking, databases, graphics, design aids, and interdisciplinary studies. A new style of computation -- pattern computing -- is emerging in the form of neural networks and associative memories that will be very helpful to us later in the decade.

What I can see is nonetheless of limited interest for me. I am far more interested in questions about what I cannot see. How do our traditional ways of thinking about our science limit the questions we ask and prevent us from seeing new approaches that will produce the innovations we require? What paradigms are we living in? What are the blind spots induced by those paradigms? What are we missing? What can we see new by stepping outside our paradigms? In short, what do we not see, and do not see that we do not see it?

It is easy for us to challenge someone else's paradigms -- and often unpleasant when someone challenges our own. The challenge often produces a startle reaction: we automatically find ourselves getting irritated, or saying "this cannot be right," or declaring "this person doesn't know what he's talking about."

I am sensitive to this. I want to challenge three of the paradigms you and I live in that affect our approach to information systems. At the same time, I want to offer some new possibilities that appear to those willing to step outside. Some of my challenges may irritate you. I ask that you say, "Oh! That's just my startle reaction," and listen on anyway.

What we can see now

By extrapolating today's trends, we can make educated guesses about eight major technologies by AD 2001.

MINIATURIZATION. We continue to refine our methods of building smaller, more power-frugal circuits. We routinely design circuits today with 100,000 transistors in the same amount of silicon as was in the first commercial transistors 25 years ago. The recent Sun SPARC RISC computer is faster and has more memory than the IBM 3033 ten years ago -- and costs under $5,000. DRAM memory chips have gone from 16K bits ten years ago to close to a million bits now and are likely to be 10 times that by the end of the decade. Look for chips of the year 2000 to offer speeds and memory comparable to today's Cray computers. Our design aids are so good that we can customize chips for special applications; look for "silicon subroutines" to be common after another ten years.

MULTIPROCESSING. Ten years ago, an advanced commercial multiprocessor was a machine with two to sixteen processing units. In one decade we have made considerable progress in mastering machines with thousands of processors. Such multicomputers are a necessity for our teraops processing goals of the mid to late 1990s. Today's Connection Machine has 65,536 ($=2^{16}$) processors; by the mid 1990s, look for one with just over 1,000,000 ($=2^{20}$) processors; by the late 1990s, look for machines of this type with over 8,000,000 processors. Look for the individual processors to have speeds beyond 100 mflops apiece. Look for considerable integration of processing, memory, and communication on each chip.

SOFTWARE TECHNOLOGY. For many years we have invested heavily in numerical software for new machines. This has paid off handsomely: since the 1940s, John Rice tells us, our PDE-solving systems have improved in speed by a

factor of 10^{12}; hardware improvements account for a factor of 10^6, algorithm improvements for the other factor of 10^6. Today's research efforts are showing us how to program the multiprocessors effectively. We are within reach of programming environments that will allow us to design highly parallel programs quickly and correctly by the mid to late 1990s.

NETWORKING. The globe is crisscrossed with communication links connecting computers, telephones, fax, radios, and televisions. I call this the phenomenon of worldnet. The distinction between a workstation and the worldwide network is blurring. In just ten years a workstation has shifted from being a personal toolkit to being a portal into the world; look for continued transformation so that by the end of the century we wear our computers, converse with them, and converse with others through them. Today's Research Internet backbone transfers data at the rate of 1.5 mbps, and NSFNET will install 45 mbps within the year. The gigabit optical fiber network should be with us by the mid 1990s. By the turn of the century our terrestrial networks will operate at 10 to 100 times that speed, depending mostly on advances in optical switch technologies and protocols. Look for the current satellite links, now running at 300 mbps, to be operating at speeds comparable with the terrestrial network. Look for networking infrastructure to reach into a sizable portion of businesses and homes in the US, Europe, And japan. Look for portable computers to be routinely connected by cellular links into the world network.

DATABASES. Mass storage systems and systems for archiving and retrieving information have been persistent problems -- our reach far exceeds our grasp. The largest direct access computational memory today is on the Cray YMP, 256 million 64-bit words. Look for this to increase significantly on multiprocessors where we can implement a uniform machine-wide virtual address space with little penalty for access between computers. Look for optical stores to become practical, replacing large disk storage "farms" with capacities of 10^{15} bits. The biggest problem will be finding information in these storage systems rather than transferring it in or out.

GRAPHICS. Look for continued improvements in resolution and function. What we today call HDTV will be the norm. Graphics libraries will permit a wide range of visualizations across many disciplines. Animations in real time will be routine. Head-mounted displays, data gloves, and other technologies associated with virtual reality will be common for scientific visualization and simulation.

PATTERN COMPUTATION. Three styles of computation are widely used today: signal processing, numeric processing, and symbolic processing. (Symbolic processing is the basis of machines that do logical inference within ai systems and languages like Prolog.) A fourth style is emerging, variously called pattern processing, associative processing, and neural processing. Its computational model -- a network of many-input threshold circuits -- is inspired by biological systems. These neural networks can store and retrieve large bit vectors that represent encoded sensory patterns. Although such systems have been the subject of speculation since the beginning of the era of electronic computing (1940s), circuit technology did not permit their construction until recently. Many new approaches to vision and speech recognition are now being tested in neural networks. Look for

this type of computing to attain maturity by the end of the century. It will not replace the other three types, but will certainly augment them. It will provide learning capabilities that are not attainable within rule-based expert systems.

INTERDISCIPLINARY STUDIES. Look for more interactions between experts in different disciplines. For example, many parallel algorithms now being developed for numerical computing will be transferred into earth sciences simulations and data analyses.

What we cannot see

Most of us here are scientists and engineers. Most of us here have worked in one discipline most of our lives. We are mostly men and mostly white. Most of us come from Judeo-Christian traditions.

These statements are facts about our common cultural background. They are neither "good" nor "bad"; they inform us about the body of shared assumptions that constitute our common wisdom about how science works, what science is important for public policy, what is innovation, what questions are worth investigating, what is true, what is good research, which data are valuable, and many similar questions. We seldom reflect on the common presuppositions given to us by our traditions. Most of the time, we are not even aware of our presuppositions. We are blind to them.

Let me give you an example. We often use the word paradigm to refer to the framework of preunderstandings in which we interpret the world. We have been taught, and we teach our students, that the great discoveries of science have happened when the discoverer challenged the current paradigm and stepped outside of it. At the same time, as recognized masters of our scientific domains, we resist changes that might leave us in less esteemed positions. Thus we have a love-hate relationship with paradigms: we like challenging the paradigms of others and we dislike others challenging our own. We especially dislike anyone suggesting that we are blind in some domain of importance to us.

Let me give you another example. As scientists we say that the scientific method consists of formulating hypotheses about the world, using them to make predictions, performing experiments to collect data, and analyzing the data for support or contradiction of the hypotheses. This method is based on a presupposition that the world is a fixed reality to be discovered. Our job is to probe the world with experiments and pass on our findings as validated models. In this preunderstanding, it is natural to say that someone discovered a new particle, discovered a new theorem, or discovered a new fact about the world; it sounds strange to say that someone invented a new particle, invented a new theorem, or invented a new fact about the world. And yet some scientists, notably chemists and molecular biologists, are engaged in a process of invention rather than discovery. The terminology of invention is natural in the paradigm of engineering. Have you ever noticed that physicists and mathematicians like to talk about the great discoveries of science while chemists and engineers like to talk about the great inventions? Because their paradigms are different, scientists and engineers often

disagree on what is "fundamental".

In his book, *Science in Action* [Harvard University Press, 1987], Bruno Latour painstakingly analyzes literature before, during, and after great discoveries and great inventions. He distinguishes between the simplified story we tell about science when looking back after the fact, and the complex web of conversations, debates, and controversies that exist before the "discovery" is accepted by the community. By tracing the literature, he demonstrates that statements are elevated to the status of "facts" only after no one has been able to mount a convincing dissent. Thus, he says, science is a process of constructing facts. Not just any statement can be accepted as fact -- a large community of people must accept the statement and must be incapable with resources and methods available to them of adducing new evidence that casts doubt on the statement.

Latour calls on the two-faced god Janus to contrast the retrospective view (an old man looking leftward, seeing "ready made science") with the in-action present view (young man looking rightward, seeing "science in the making"). See Figure 1. Examples of statements made by Latour's Janus are:

OLD: "Just get the facts straight."
YOUNG: "Get rid of the useless facts."

OLD: "Just get the most efficient machine."
YOUNG: "Decide on what efficiency should be."

OLD: "Once the machine works, people will be convinced."
YOUNG: "The machine will work when all the relevant people are convinced."

OLD: "When things are true, they hold."
YOUNG: "When things hold, they start becoming true."

OLD: "Science is not bent by the multitude of opinions."
YOUNG: "How to be stronger than the multitude of opinions?"

OLD: "Nature is the cause that allowed the controversies to be settled."
YOUNG: "Nature will be the consequence of the settlement."

It is interesting that although the young man's statements are typical of the ones we make while "doing science", we quickly adopt the old man's views as soon as the "science is done." Our research papers, for example, describe orderly, systematic investigations proceeding from problem descriptions, to experiments, to data collections and analyses, to conclusions. The description tells a story that never happened: it fits neatly inside the scientific-method paradigm while the discovery itself is made inside a network of ongoing conversations. We do this also with the history of science. We trace an idea back to its roots, giving the first articulator the full credit. (If the idea is great enough, we give its original articulator a Nobel Prize.) The complex, dynamic web of conversations and controversies disappears. I will argue shortly that this paradigm of science is
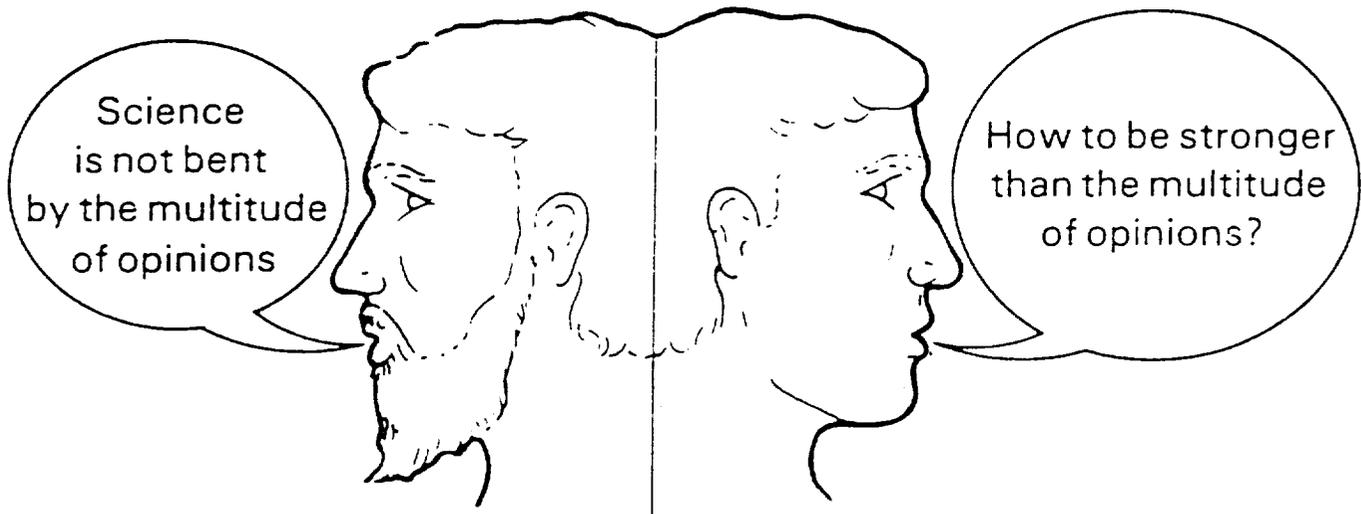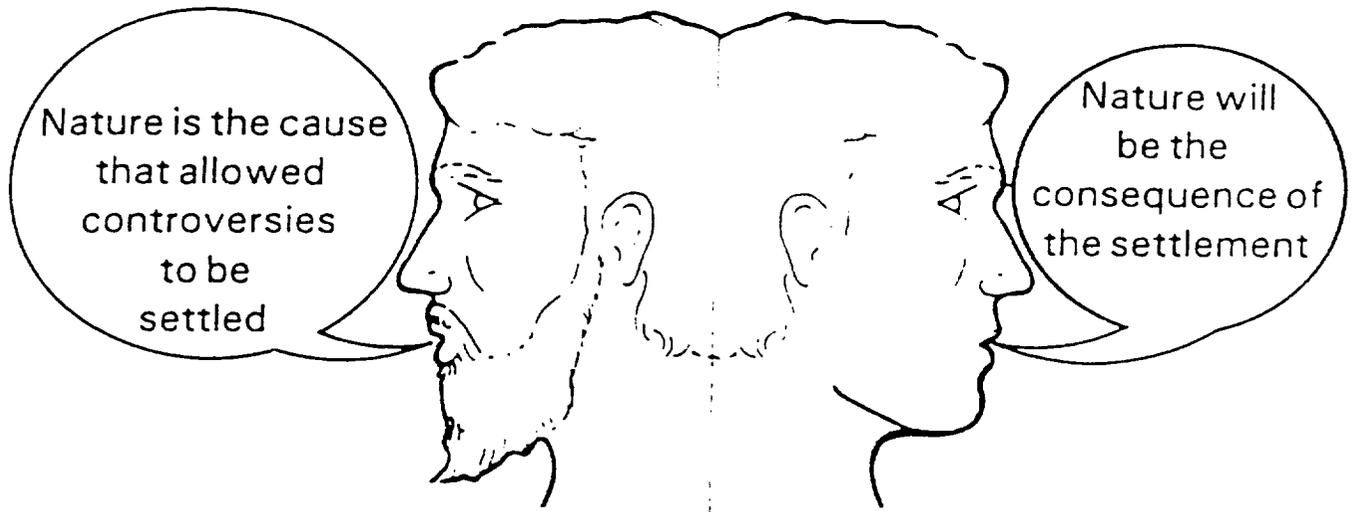
FIGURE 1. In his book, *Science in Action*, Bruno Latour illustrates the contrasts between the view of science after a statement has been accepted as fact (leftward looking face of Janus) and the view while statements are being defined and debated (rightward looking face).

linked to our nation's difficulties to compete effectively in world markets.

I see three major paradigms that shape our thinking about information systems. The first I call *saving all the bits*. Those in this paradigm argue that all bits from instruments and massive computations must be saved, either because the cost of recovering them is too high or because some important discovery might be lost forever. I will show two examples of new technologies that offer the possibility of increasing our power to make new discoveries without having to save all the bits.

The second of the three paradigms I call *obtaining technology off the shelf*. Those in this paradigm argue that NASA ought not sponsor its own research in information system technologies because research money ought to be spent on science and because the needed technology can be acquired from the commercial sector. I argue that this paradigm equates networking with connectivity and ignores networking as a way of collaborating. I argue that NASA has unique mission requirements that do not now appear in the market, and will not over the coming decade; thus I see that the commercial sector will be incapable of delivering the innovations NASA requires.

The third paradigm I call *the linear model of innovation*. Those in this paradigm argue that every innovation begins with a discovery or invention and passes successively through the stages of development, production, and marketing on the way to the customer. They see research as the noble beginning of all innovation. I argue that in reality a cyclical model is at work. Most innovation is accomplished by refinements over successive generations of a science or technology. I argue that NASA must design research programs to create and sustain cycles of innovation that involve NASA, university researchers, and commercial partners. The Numerical Aerodynamic Simulation (NAS) project is a successful instance of a cyclical model of innovation in NASA.

I will now discuss each of these paradigms in more detail.

Saving all the bits

I often hear from colleagues in earth sciences, astronomy, physics, and other disciplines that after we start up an expensive instrument or complete a massive computation, we must save all the bits generated by that instrument or computation. The arguments for this are first, the cost of the instrument or computation is so great that we cannot afford the loss of the information produced, and second, some rare event may be recorded in those bits and their loss would be a great loss for science. I have heard debates in which these points are made with such vehemence that I am left with the impression that saving the bits is not merely a question of cost, it is a moral imperative.

Those in this paradigm are perforce limited to questions about saving and moving bits. How shall we build a network with sufficient bandwidth to bring all the bits from instruments to us? How shall we build storage devices to hold them? How shall we build retrieval mechanisms that allow us to access them from around

the world? Data compression is of interest only if it is "lossless", i.e., it is a reversible mapping from the original data to the compressed data. "Smart instruments" that detect patterns in the data and inform us of those patterns are of little interest -- it is claimed, for example, that such "on-board processing" delayed the discovery of the ozone hole for several years.

As we speak, the Hubble Space Telescope is in limited operation, sending us on the order of 300 mbps via the TDRSS satellite link network to Goddard. This will be joined shortly with the ACT (advanced communications technology) satellite and, in a few years, the network of satellites making up the EOS (earth observing system). These are just a few of the growing number of advanced instruments we have put into space, any one of which can produce data streams at the rate of hundreds of mbps.

Let us do some simple arithmetic with the EOS alone. This system is expected to produce between 10^{12} and 10^{13} bits per day. (This is an enormous number. If we had one ant carrying each of those bits, a day's transmission would make a chain of ants stretching all the way from earth to sun.) It would take 2,500 CDs (compact optical disks) at about 4 gigabits capacity each to hold one day's data. Increases in optical storage density may allow this number to be reduced by a factor of 10 or 100 by the time EOS is on line. Where will all this storage be? Is Goddard going to be responsible for recording 2,500 disks daily? Even the national gigabit network will be inadequate to divert all those streams to other sites for recording elsewhere. And if we succeed in recording all the bits, how is anyone going to access them? How do I as a scientist ask for the records that might contain evidence of a particular event of interest? I am asking for a search of 2,500 disks representing one day's observations, 0.9 million disks for a year's, or 9 million disks if I want to examine trends over a ten-year period.

This scenario doesn't mention the data fusion problem that arises when an investigator requests to study several different data sources simultaneously for correlations. I have heard it said that advanced graphics will allow the investigator to visualize all the bits and see the correlations. But this statement is too glib: it hides the limitations on bandwidth of networks, speeds of graphics devices, methods of storing and retrieving the data, and algorithms for performing the correlations.

In short, the paradigm of saving all the bits forces us into an impossible situation: the rate and volume of the bits overwhelm our networks, storage devices, retrieval systems, and human capacities of comprehension.

Suppose we step outside the paradigm and say that there are important cases in which we do not need all the bits. What machines can we build that will monitor the data stream of an instrument, or sift through a database of recordings, and propose for us a statistical summary of what's there?

Let me give an example under test jointly by RIACS and the Artificial Intelligence Branch at NASA-Ames. Peter Cheeseman has developed a program called Autoclass that uses Bayesian inference to automatically discover the smallest set of statistically distinguishable classes of objects present in a database. In 1987 Autoclass was applied to the 5,425 records of spectra observed by the

Infrared Astronomical Satellite (IRAS) in 1983 and 1984. Each record contained two celestial coordinates and 94 intensities at preselected frequences in the range of wavelengths 7 to 23 microns. Autoclass reported most of the classes previously observed by astronomers, and most of the differences were acknowledged by astronomers as clearly representing unknown physical phenomena. NASA reissued the star catalog for the IRAS objects based on Autoclass's results.

One of these discoveries is shown in the accompanying Figure 2. Previous analyses had identified a set of 297 objects with strong silicate spectra. Autoclass partitioned this set into two parts. The class on the top left (171 objects) has a peak at 9.7 microns and the class on the top right (126 objects) has a peak at 10.0 microns. When the objects are plotted on a star map by their celestial coordinates (bottom), the right set shows a marked tendency to cluster around the galactic plane, confirming that the classification represents real differences between the classes of objects. Astronomers are studying this phenomenon to determine the cause.

There is nothing magic about Autoclass. It is a machine that can take a large set of records and group them into similarity classes using Bayesian inference. It is thus an instrument that permits finer resolution than is possible with the unaided human eye. It does not need to know anything about the discipline in which the data were collected; it does its work directly on the raw data.

The important point illustrated by Autoclass is that a machine can isolate a pattern that otherwise would have escaped notice by human observers. The machine enabled new discoveries, otherwise impossible.

Cheeseman suggests that an Autoclass analyzer could be attached to an instrument, where it would monitor the data stream and form its own assay of the distinguishable classes. It would transmit the class descriptions to human observers on the ground at significant reductions in bandwidth. If the human observer wanted to see all the details of specific objects, he could send a command instructing the analyzer to pipe all the bits straight through.

Let me give a second example. Also at RIACS we have a project studying an associative memory architecture called SDM (sparse distributed memory). See Figure 3. In the SDM each memory cell contains a name field (a vector of bits) and a data field (a vector of counters). When an address pattern (a bit vector) is presented, address decoders at all the cells simultaneously determine whether the given address and their own names are close by some measure such as Hamming distance; all the cells for which this is true participate in the read or write operation requested relative to the given address. Writing is accomplished by adding an image of the data vector to these counters, reading by statistically reconstructing a bit vector from these counters. We have a simulator running on the Connection Machine; it simulates a memory of 100,000 cells with bit vector lengths of 256, and it cycles 10 times a second.

In one experiment David Rogers sought to learn if a variant of SDM could learn the correlations between measurements and desired results. He fed SDM a stream of approximately 58,000 records of weather data from a station in Australia. Each record contained 12 measurements and a bit indicating whether rain fell in the
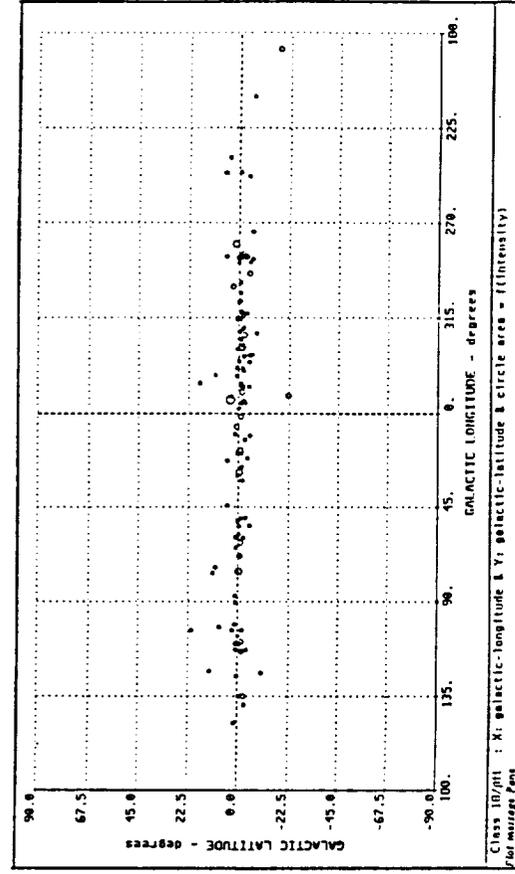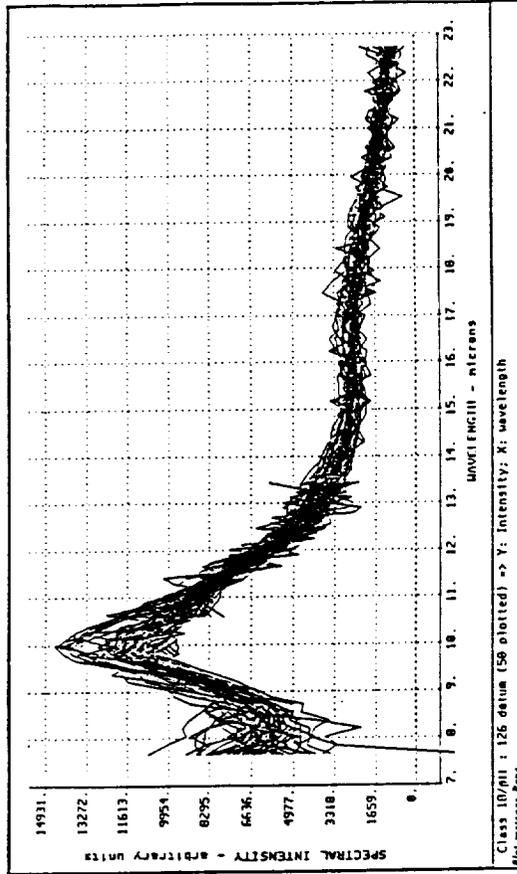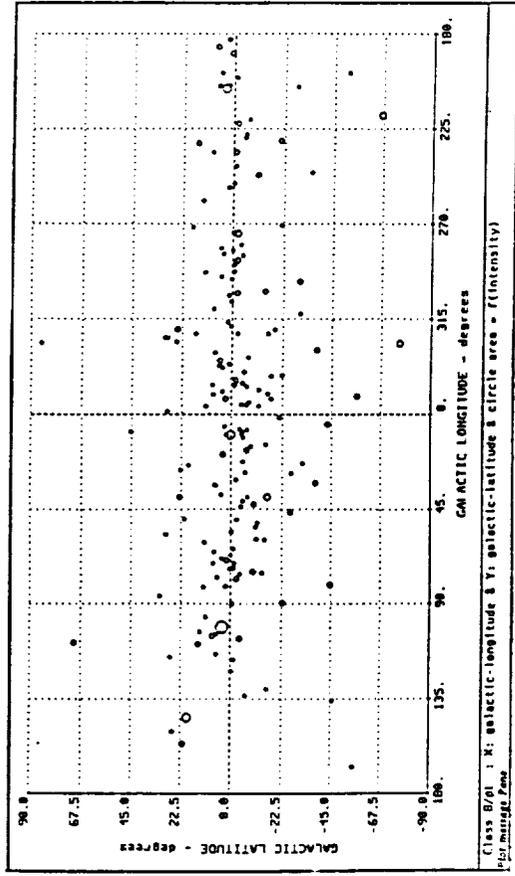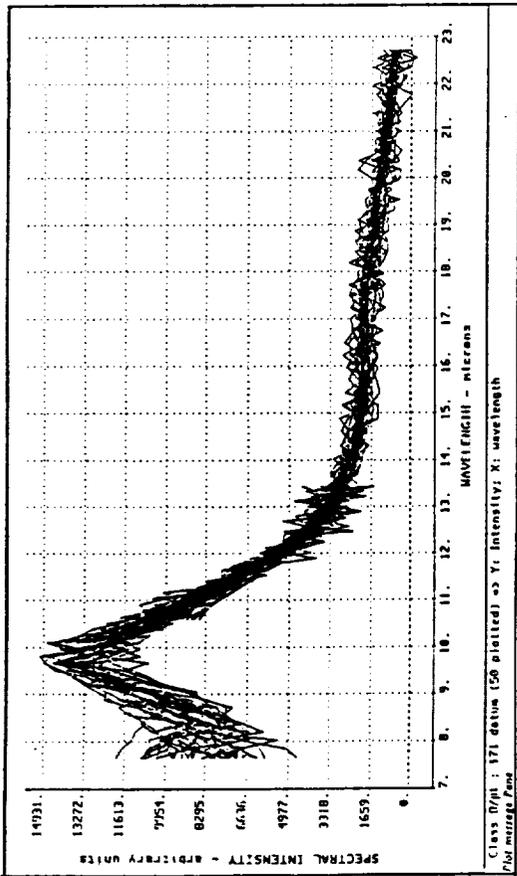
FIGURE 2. In 1983 and 1984, the Infrared Astronomical Satellite (IRAS) detected 5,425 stellar objects and measured their infrared spectra. A program called AUTOCLASS used Bayesian inference methods to discover the classes present in the data and determine the most probable class of each object. It discovered some classes that were significantly different from those previously known to astronomers. One such discovery is illustrated in the accompanying picture. Previous analysis had identified a set of 297 objects with strong silicate spectra. AUTOCLASS partitioned this set into two parts (top). The class on the left (171 objects) has a peak at 9.7 microns and the class on the right (126 objects) a peak at 10.0 microns. When the objects are plotted on a star map by their celestial coordinates (bottom), the right set shows a marked tendency to cluster around the galactic plane, confirming that the classification represents real differences between the classes of objects. AUTOCLASS did not use the celestial coordinates in its estimates of classes. Astronomers are studying the phenomenon further to determine the cause.

34

1 1 0 1 1 0 0 1 1 1 0 1 1 0 1 1 0 1... (17)

1 0 1 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1... (512) ◄—— Good selection

0 0 0 0 0 1 1 1 1 0 0 1 0 0 0 0 0 1... (222)

0 0 1 1 0 1 1 0 0 1 0 1 1 0 1 1 0 0... (32)

1 0 1 1 1 0 1 1 0 0 0 0 1 0 1 1 0 1... (87)

0 0 1 0 1 0 1 1 1 1 1 1 0 1 0 0 0... (156)

1 1 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1 0... (302) ◄—— Good selection

0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0... (3)

0 1 1 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0... (1) ◄—— Bad selection

1 0 1 1 0 1 0 1 1 0 0 1 1 1 0 0 1 1... (55)

1 1 0 0 0 1 0 1 1 1 0 1 0 1 0 1 0 1... (216)

...etc... (...)

Population of Binary strings Fitness scores

First parent

Second parent

New member
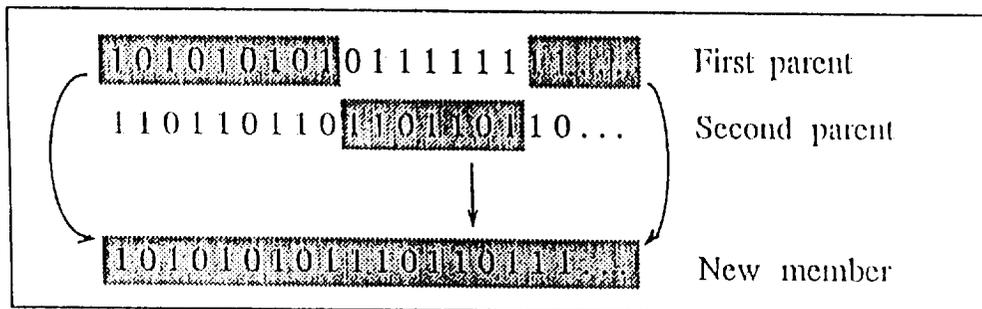
FIGURE 3. The genetic sparse distributed memory is an associative memory system whose addresses are dynamically modified during training so that they collectively evolve toward a set that is capable of best prediction of a future data element. The idea of address modification is based on Holland's genetic algorithm.

measurement period. The measurements were encoded into a 256-bit vector, and the rain bit of the next period was used as data. Just before the actual next-period rain bit was stored, the SDM was asked to retrieve its version of the bit. If the retrieved bit agreed with the bit about to be written, each selected cell had 1 added to its "success count". At intervals the two highest scoring cells were cross-bred by combining pieces of their names; the new name thus created replaced the name in the lowest- scoring cell. This is the principle used in genetic algorithms, and Rogers calls his variant the genetic memory.

At the end of the experiment, Rogers found that the memory gave accurate predictions of rain. By examining the name fields of all memory cells, he was able to determine which subset of the measurements were the most correlated with the occurrence of rain in the next measurement period.

The genetic memory is a machine that can be fed a stream of data. It organizes itself to become a consistent predictor of a specified pattern.

Both these examples show that it is possible to build machines that can recognize or predict patterns in data without knowing the "meaning" of the patterns. Such machines may eventually be fast enough to deal with large data streams in real time. By the end of the decade they may be well enough advanced that they can serve on space probes and space-borne instruments, where they can monitor streams that would be incomprehensible to us directly. With these machines, we can significantly reduce the number of bits that must be saved, and we can increase the likelihood that we will not lose latent discoveries by burying them forever in a large database. The same machines can also pore through databases looking for patterns and forming class descriptions for all the bits we've already saved.

I am not alone in this conclusion. In *Science*, 11 May 1990, journalist Mitchell Waldrop documents the rising concern in the science community about the volumes of data that will be generated by supercomputers and by instruments. He likens the coming situation with drinking from a fire hose: "Instant access to far-flung databases could soon be a reality, but how will we swallow a trillion bytes a day?" He is drawn to a proposal by Robert Kahn and Vinton Cerf to create surrogate processes that would roam the networks looking for data of a particular kind, returning home with their findings. Called knowbots (short for knowledge robots), these processes would resemble benign viruses in their operation. The article ends without saying how knowbots might work. What do you suppose would go inside? Machines that perform automatic discovery, pattern matching, and prediction.

Technology off the shelf

Over the past decade I've repeatedly heard representatives of scientific disciplines giving testimony to NSF, NASA, ONR, advising those agencies against engaging in research on networking. They have argued that the research dollars should be spent on science, that networking is technology, not science, and that the government can acquire the technology it needs "off the shelf" from the

commercial sector. This way of thinking has stopped NASA from engaging in research on its networking needs, and it nearly stopped the NSFnet from being formed. The high performance computing initiative plan departs only slightly from this way of thinking by specifying a technology project to produce a gigabit network by 1995 that will be taken over by the commercial sector. This paradigm does not distinguish networking as connectivity from networking as a way of collaborating.

I'm not challenging the statement that we must build an infrastructure of networks and databases that will allow data to be stored, shared, and analyzed in the scientific community. Many of the components of such an infrastructure are (or will be) available in the commercial market. In those cases, it is appropriate for the government to acquire the needed technologies "off the shelf."

I am challenging the notion that all NASA's networking needs can (or will) be satisfiable commercially. I am specifically challenging the notion that NASA needs no research efforts of its own that treat problems arising in the context of large networks of computers, databases, instruments -- and scientists collaborating over large distances.

NASA is the only organization on earth with the data needs of the magnitudes outlined earlier. No commercial organization has such needs. No commerical customers demand products that would cope with such bandwidths or volumes of data. NASA has defined a unique set of requirements. We are simply not going to cope with all the data with our current ways of thinking: we need wholly new ways of thinking about and handling data. This is true for each major NASA scientific community. NASA earth scientists, I say, must organize their own research program to study data collection, recording, retrieval, fusion, analysis, and understanding in their disciplines. No one else is looking at these questions.

Linear model of innovation

Many innovations will be needed to achieve the goals for information systems by the turn of the century. Most of us think about how to bring those innovations about within the confines of a "linear model" of innovation. (See Figure 4.) This is the familiar model that says every innovation begins with a discovery or invention (usually by some individual or at some institution) and passes successively through the stages of development, production, and marketing on the way to the customer. We use the term research to refer to institutional activities that systematically seek to spawn new discoveries that feed the pipeline. We see research as the noble beginning of all innovation.

In my discussion of Latour, I noted that this model seems to fit what we see when we look back from the present to the past moment when the idea was first articulated. That retrospective history seems to contain the stages noted above.

But the retrospective model is limiting because it hides the intricate webs of conversation, false starts, controversies, and iterations that take place while we seek to make a technology usable by many people.
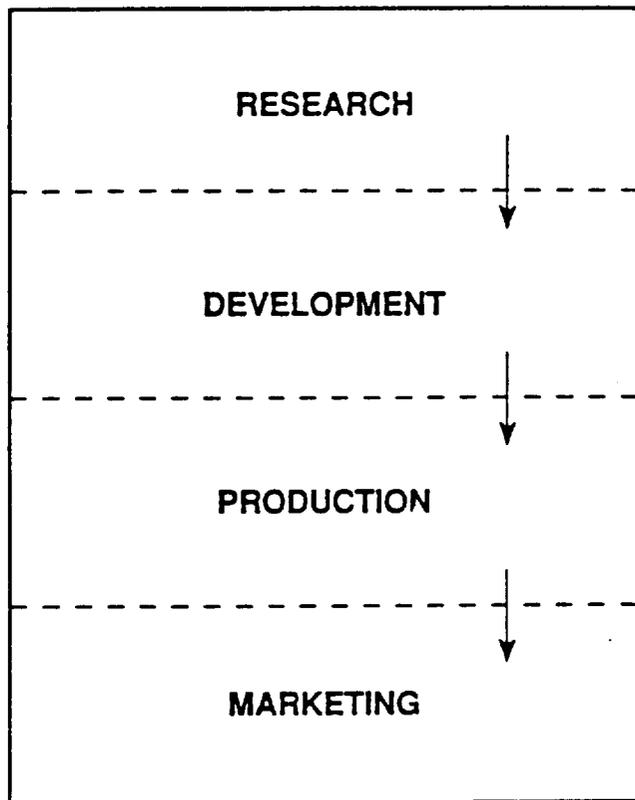
FIGURE 4. Steve Kline, among others, has challenged the linear model of innovation, which holds that ideas are generated during research and then flow through a pipeline of development, production, and marketing on the way to customers. He depicts the model as shown here.

Stephen Jay Kline published a report called "Innovation Styles in Japan and the United States [Stanford University, Department of Mechanical Engineering, Report INN-3, December 1989]. He analyzed in some detail how the actual process of innovation differs markedly from the linear model given to us by our cultural paradigm. Kline reprints a figure compiled by Christopher Hill of the Library of Congress in 1986 showing an inverse relation between Nobel Prizes and growth of GNP, just the opposite of what one would expect if innovation took place according to the linear model. (See Figure 5.) Kline shows that an accurate model consists of many feedback cycles among the various stages of development of a technology: research permeates and sustains all the stages.

Writing in *Scientific American* in June 1990, Ralph Gomory also criticizes the linear model and says that a cyclical model is actually at work in most cases of innovation. While some innovations have been introduced by a linear model, most occur by successive refinements over a series of generations of a product.

Why is this relevant to NASA? As we lay our plans for research in during the 1990s, we must not fall into the trap of thinking that NASA earth scientists will be the original source of many future discoveries that will benefit all of earth science and then eventually all of society. We must instead design our research programs to create and sustain cycles of innovation that involve NASA, university researchers, and commercial partners. We are much more likely to reach our goals of 2001 AD by engaging in cycles of innovation than by setting ourselves up to be either the source of new ideas or the recipient of new ideas generated by others.

The Numerical Aerodynamic Simulation (NAS) facility at Ames illustrates the approach. A major component of the work needed to achieve the national goal of complete simulation of an aircraft inside a computer is technological: namely the acquisition of supercomputers. The planners of the NAS, however, recognized that the architectures of supercomputers such as the Cray-1 and Cyber 205 could not be extended to deliver the needed teraflops computational rates. They argued that the requirement for such speeds was unique to NASA, and thus NASA would have to work closely with commercial partners to foster the development of supercomputers with thousands of processors. They argued that a research component was also needed to develop entirely new kinds of algorithms to exploit the machines and assist the aircraft companies to use the NAS. The NAS they designed has many cycles of activity in it including partnerships with industry, aircraft companies, other supercomputing centers, and universities; it also has a research group on site supporting all these activities. This facility embodies a cyclical model of innovation. It is of obvious value to the US aircraft industry and the nation. It is a smashing success.

Conclusions

We live in three paradigms that can impose severe limitations on what NASA can accomplish in an information systems program during the 1990s. It is not necessary to give up these paradigms; they have been useful in the past. It is, however, necessary to avoid being limited by them.
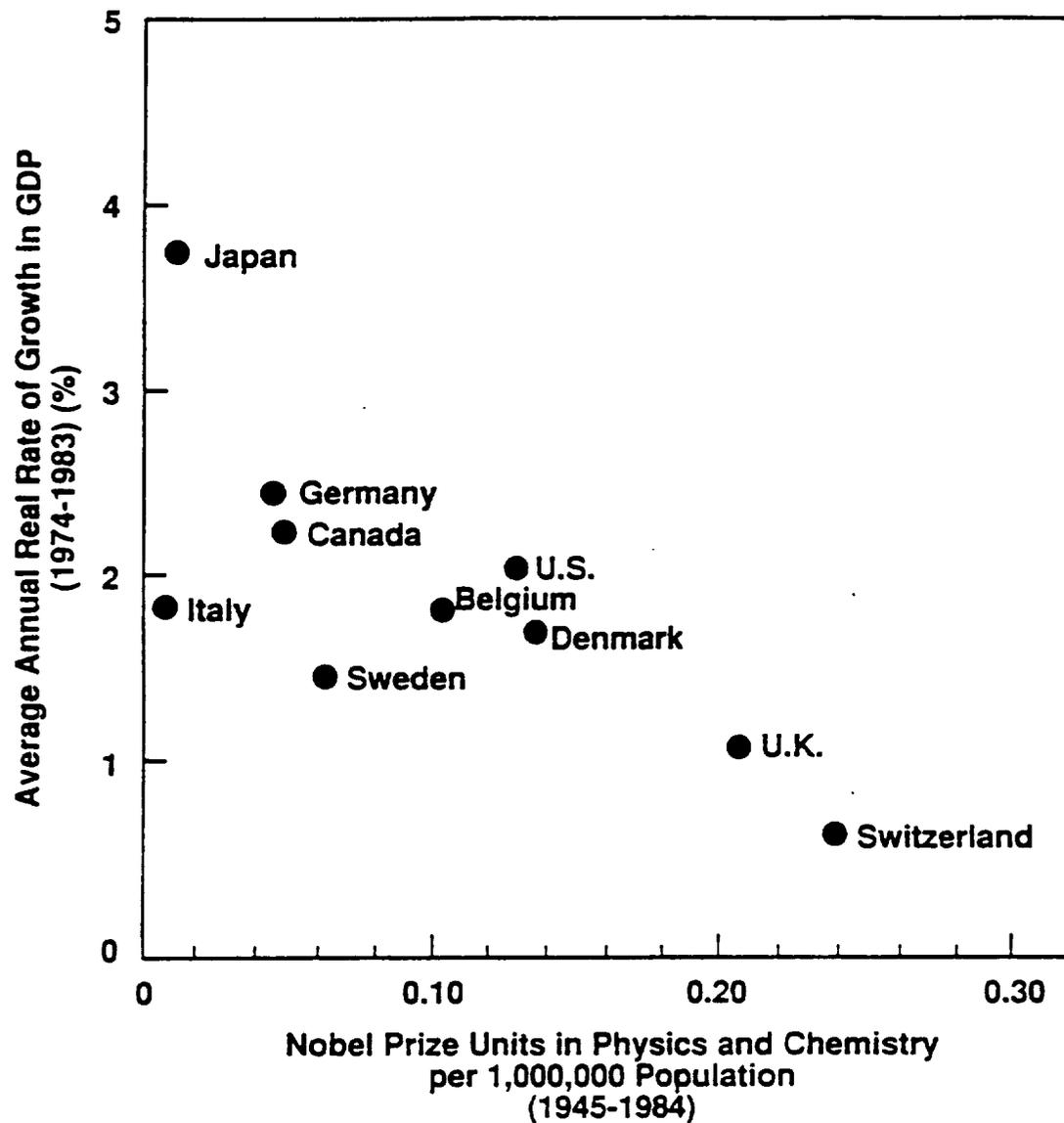
FIGURE 5. Kline cites a Congressional study by Hill in 1986 as striking evidence against the linear model. Hill found inverse correlation between the number of Nobel Prizes and the annual growth of a country's economy.

* C. Hill, Congressional Research Service Study, April 16, 1986.

To go beyond the save-all-the-bits way of thinking, I recommend that NASA include research on machines that can perform automatic discovery, pattern identification, prediction, correlation, and fusion. Such machines would allow us to make more discoveries without having to store all the bits generated by instruments. They could be part of the instrument itself, and could be shut off during intervals when all the bits are needed.

To go beyond the technology-off-the-shelf way of thinking, I recommend that NASA declare that most of its requirements in information management are unique to the agency because of the magnitude of the needed bandwidths and storage and the size of the participating scientific community. I recommend that NASA undertake research programs that will assure the presence of technology needed for the NASA missions.

To go beyond the linear-model-of-innovation way of thinking, I recommend that NASA position itself as a sustainer of the cycles of innovation that will be needed to produce the technologies required for NASA missions in earth sciences during the late 1990s.

Editor's Note:

Peter Denning was the Director of RIACS at the time this paper was written. RIACS is an institute of the Universities Space Research Association at the NASA Ames Research Center. RIACS works to assist Ames in fulfilling its missions by engaging computing researchers from the universities with Ames in parallel computation, learning systems, and networking. Denning has a Ph.D. in electrical engineering from MIT. He served 15 years in computer science departments at Princeton and then at Purdue. He has written extensively and, since coming to Ames in 1983, he has focused on interfaces between research in computing and in other disciplines.