

---

# **EXTROPY #10**

## **Winter/Spring 1993 (vol.4, no.2)**

---

<b>Editorial</b>	<b>Max More</b>	<b>04</b>
<b>Pigs in cyberspace</b>	<b>Hans Moravec</b>	<b>05</b>
<b>Protecting privacy with electronic cash</b>	<b>Hal Finney</b>	<b>08</b>
<b>Technological self-transformation: Expanding personal extropy</b>	<b>Max More</b>	<b>15</b>
<b>Mark Miller of Xanadu interviewed (Pt. 1) Creole physics and the credit theory of identity</b>	<b>Dave Krieger</b>	<b>25</b>
<b>Nanocomputers: 21st century hypercomputing</b>	<b>J. Storrs Hall</b>	<b>31</b>
<b>Extropy Institute news and information</b>		<b>38</b>
<b>Back issues listing</b>		<b>39</b>
 <b><i>The Transhuman Taste:</i></b>		
<b>Beyond the poor man's Extropianism: A review of two books about Ayn Rand and Objectivism Reviews by Mark Plus</b>		<b>40</b>
<b><i>Nanosystems: Molecular Machinery, Manufacturing, and Computation</i> Reviewed by J. Storrs Hall</b>		<b>44</b>
<b><i>Genius – The Life and Science of Richard Feynman</i> Reviewed by Harry Shapiro</b>		<b>45</b>
<b>Contributors</b>		<b>47</b>
<b>T-shirts and tapes</b>		<b>48</b>

# EDITORIAL

This issue continues our inexorable advance on several fronts: #10 is larger than ever, with 48 densely packed pages (and would have been 56 pages if not for cuts forced by escalating costs). Color enlivens the cover for the first time. Distribution, both in the USA and internationally, has grown greatly since last issue: 2,500 copies are being printed and distributed all over the planet. This external growth accompanies an influx of new writers with a talent for intelligent, radical, and clear expression.

Rapidly growing circulation and production quality immediately increased costs unmatched by short-term income gains, forcing us to trim the issue down from the original monster. Two items postponed until next issue are economist Julian Simon's analysis of how environmentalist Cassandras falsely alarm people regarding population and resources, and my review of *Rational Readings on Environmental Concerns*.

In this issue, David Krieger conducts the first in a series of interviews with leading transhumanist thinkers and researchers. David's first subject is Mark Miller, a software developer working on the fabled Xanadu hypertext project. Mature hypertext will vastly increase the interconnectedness of information, allowing us to uncover knowledge and opinion with far greater selectivity, productivity, and ease.

Master roboticist Hans Moravec returns with a sweeping look at our future expansion into cyberspace and physical space. Hal Finney's excellent introduction to electronic cash and public key cryptography reveals a powerful means of protecting our privacy while facilitating remote market transactions. A "cyberpunks" group has formed in Northern California to discuss and develop these techniques.

"Technological self-transformation" continues my development of the Extropian philosophy. J. Storrs Hall, moderator of the Internet newsgroup sci.nanotech, provides an illuminating introduction to the design and potential of nanocomputers—molecular-scale computers orders of magnitude smaller and faster than our current computing devices. He also reviews Drexler's long-awaited technical tome *Nanosystems*.

In "Beyond the poor man's Extropianism," Mark Plus reviews two recent books on Ayn Rand's philosophy of Objectivism, examining the extent to which Rand's intellectual descend-

dents have succeeded in building her ideas into a philosophy of life fit for today and tomorrow. Harry Shapiro reviews Gleick's book on the unconventional genius, Richard Feynman.

The spread of Extropian ideas is bursting beyond the confines of these pages: Apart from several e-mail lists spawned by Extropians, a growing number of reporters are calling ExI for our unique input on stories regarding advanced technologies and the future. On January 27 I will appear on *Breakthroughs: A TransCentury Update* – a public access cable-TV show in Los Angeles – explaining the Extropian worldview for thirty minutes to a potential audience of 120,000. We may also be mentioned soon in *Time*, and in a London newspaper. This exposure should attract new readers to *Extropy*, and bring new members into Extropy Institute, helping to ensure our sustained and expanded activity. If all goes well, *Extropy* may move to three or four issues per year in 1994.

Max More  
Editor

**NEXT ISSUE:** *Extropy* #11, is due to be mailed on July 15 1993. Likely features include:

- A memetic analysis of the spread of Extropian ideas.
- The construction and politics of ocean habitats for experimental living.
- Artificial languages and increased rationality.
- How to upload consciousness to a computer.
- Part II of the Mark Miller interview.
- Space colonization – future space launch and habitation systems.
- An article on "fuzzy logic" by its foremost proponent.
- Economist Julian Simon's "Bunkrupt: The Abstractions that Lead to Scars About Resources and Population Growth."
- Plus other possible articles, and the usual detailed reviews in *The Transhuman Taste*.

In the months before *Extropy* #11 is published, more Extropian reading will appear in issues 4, 5, and 6 of the Extropy Institute newsletter, *Exponent*. See p.38 for more information.

**MISSING CREDITS:** In a grievous oversight, resulting from last-minute haste, I omitted to credit Extropy Institute director Tom Morrow for producing the new five-arrow spiral *Extropy* emblem gracing the cover of *Extropy* #9. Tom also designed the ExI logo. My thanks and apologies to Tom.

**EXTROPY** — a measure of intelligence, information, energy, life, experience, diversity, opportunity and growth. Extropianism is the philosophy that seeks to increase extropy. The Extropian Principles are: (1) Boundless Expansion; (2) Self-Transformation; (3) Intelligent Technology; (4) Spontaneous Order; (5) Dynamic Optimism. [See *Extropy* #9]

**TRANSHUMANISM** — Philosophies of life (such as Extropianism) that seek the continuation and acceleration of the evolution of intelligent life beyond its currently human form and human limitations by means of science and technology, guided by life-promoting principles and values, while rejecting dogma and religion. [See *Extropy* #6]

# Pigs in cyberspace

Hans Moravec

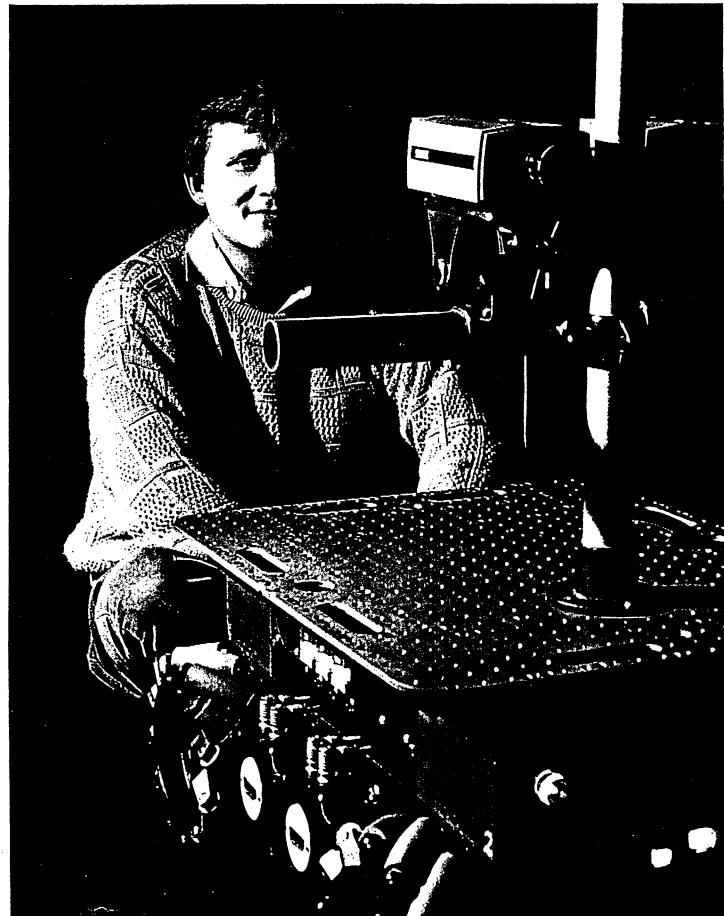
Robotics Institute  
Carnegie Mellon University  
Pittsburgh, PA 15213

© May 1992

Exploration and colonization of the universe awaits, but earth-adapted biological humans are ill-equipped to respond to the challenge. Machines have gone farther and seen more, limited though they presently are by insect-like behavioral inflexibility. As they become smarter over the coming decades, space will be theirs. Organizations of robots of ever increasing intelligence and sensory and motor ability will expand and transform what they occupy, working with matter, space and time. As they grow, a smaller and smaller fraction of their territory will be undeveloped frontier. Competitive success will depend more and more on using already available matter and space in ever more refined and useful forms. The process, analogous to the miniaturization that makes today's computers a trillion times more powerful than the mechanical calculators of the past, will gradually transform all activity from grossly physical homesteading of raw nature, to minimum-energy quantum transactions of computation. The final frontier will be urbanized, ultimately into an arena where every bit of activity is a meaningful computation: the inhabited portion of the universe will transformed into a cyberspace.

Because it will use resources more efficiently, a mature cyberspace of the distant future will be effectively *much* bigger than the present physical universe. While only an infinitesimal fraction of existing matter and space is doing interesting work, in a well developed cyberspace every bit will be part of a relevant computation or storing a useful datum. Over time, more compact and faster ways of using space and matter will be invented, and used to restructure the cyberspace, effectively increasing the amount of computational spacetime per unit of physical spacetime.

Computational speedups will affect the subjective experience of entities in the cyberspace in a paradoxical way. At first glimpse, there is no subjective effect, because everything, inside and outside the individual, speeds up equally. But, more subtly, speedup produces an expansion of the cyber universe, because, as thought accelerates, more *subjective* time passes during the fixed (probably lightspeed) *physical* transit time of a message between a given pair of locations—so those fixed locations seem to grow farther apart. Also, as information storage is made continually more efficient through both denser utilization of matter and more efficient encodings,



there will be increasingly more cyber-stuff between any two points. The effect may somewhat resemble the continuous-creation process in the old steady-state theory of the physical universe of Hoyle, Bondi and Gold, where hydrogen atoms appear just fast enough throughout the expanding cosmos to maintain a constant density.

A quantum-mechanical entropy calculation by Bekenstein suggests that the ultimate amount of information that can be stored given the mass and volume of a hydrogen atom is about a megabyte. But let's be conservative, and imagine that at some point in the future only "conventional" physics is in play, but every few atoms stores a useful bit. There are about  $10^{56}$  atoms in the solar system. I estimate that a human brain-equivalent can be encoded in less than  $10^{15}$  bits. If a body and surrounding environment takes a thousand times more storage in addition, a human, with immediate environment, might consume  $10^{18}$  bits. An AI with equivalent intelligence could probably get by with less, since it does without the body-simulation "life support" needed to keep a body-oriented human mind sane. So a city of a million human-scale inhabitants might be efficiently stored in  $10^{24}$  bits.

If the atoms of the solar system were cleverly rearranged so every 100 could represent a bit, then a single solar system could hold  $10^{30}$  cities—far more than the number ( $10^{22}$ ) of stars in the *visible universe*! Multiply that by  $10^{11}$  stars in a galaxy, and one gets  $10^{41}$  cities per galaxy. The visible universe, with  $10^{11}$  galaxies, would then have room for  $10^{51}$  cities—except that by the time intelligence has expanded that far, more efficient ways of using spacetime and encoding data would surely have been discovered, increasing the number much further.

## Mind without body?

Start with the concepts of telepresence and virtual reality. You wear a harness that, with optical, acoustical, mechanical and chemical devices controls all that you sense, and measures all of your actions. Its machinery presents pictures to your eyes, sounds to your ears, pressures and temperatures to your skin, forces to your muscles and even smells and tastes for the remaining senses. Telepresence results when the inputs and outputs of this harness connect to a distant machine that looks like a humanoid robot. The images from the robot's two camera eyes appear on your "eyeglass" viewscreens, and you hear through its ears, feel through its skin and smell through its chemical sensors. When you move your head or body, the robot moves in exact synchrony. When you reach for an object seen in the viewscreens, the robot reaches for the object, and when it makes contact, your muscles and skin feel the resulting weight, shape, texture and temperature. For most practical purposes you inhabit the robot's body—your sense of consciousness has migrated to the robot's location, in a true "out of body" experience.

Virtual reality retains the harness, but replaces the remote robot with a computer simulation of a body and its surroundings. When connected to a virtual reality, the location you seem to inhabit does not exist in the usual physical sense, rather you are in a kind of computer-generated dream. If the computer has access to data from the outside world, the simulation may contain some "real" items, for instance representations of other people connected via their own harnesses, or even views of the outside world, perhaps through simulated windows.

One might imagine a hybrid system where a virtual "central station" is surrounded by portals that open on to views of multiple real locations. While in the station one inhabits a simulated body, but when one steps through a portal, the harness link is seamlessly switched from the

simulation to a telepresence robot waiting at that location.

The technical challenges limit the availability, "fidelity" and affordability of telepresence and virtual reality systems today—in fact, they exist only in a few highly experimental demonstrations. But progress is being made, and it's possible to anticipate a time, a few decades hence, when people spend more time in remote and virtual realities than in their immediate surroundings, just as today most of us spend more time in artificial indoor surroundings than in the great outdoors. The remote bodies we will inhabit can be stronger, faster and have better senses than our "home" body. In fact, as our home body ages and weakens, we might compensate by turning up some kind of "volume control." Eventually, we might wish to bypass our atrophied muscles and dimmed senses altogether, if neurobiology learns enough to connect our sensory and motor

---

**The visible universe, with  $10^{11}$  galaxies, would then have room for  $10^{51}$  cities—except that by the time intelligence has expanded that far, more efficient ways of using spacetime and encoding data would surely have been discovered, increasing the number much further.**

---

nerves directly to electronic interfaces. Then all the harness hardware could be discarded as obsolete, along with our sense organs and muscles, and indeed most of our body. There would be no "home" experiences to return to, but our remote and virtual existences would be better than ever.

The picture is that we are now a "brain in a vat," sustained by life-support machinery, and connected by wonderful electronic links, at will, to a series of "rented" artificial bodies at remote locations, or to simulated bodies in artificial realities. But the brain is a biological machine not designed to function forever, even in an optimal physical environment. As it begins to malfunction, might we not choose to use the same advanced neuro-

logical electronics that make possible our links to the external world, to replace the gray matter as it begins to fail? Bit by bit our brain is replaced by electronic equivalents, which work at least as well, leaving our personality and thoughts clearer than ever. Eventually everything has been replaced by manufactured parts. No vestige of our original body remains, but our thoughts and awareness continue. We will call this process, and other approaches with the same end result, the *downloading* of a human mind into a machine. After downloading, our personality is a pattern impressed on electronic hardware, and we may then find ways to move our minds to other similar hardware, just as a computer program and its data can be copied from processor to processor. So not only can our sense of awareness shift from place to place at the speed of communication, but the very components of our minds may ride on the same data channels. We might find ourselves distributed over many locations, one piece of our mind here, another piece there, and our sense of awareness at yet another place. Time becomes more flexible—when our mind resides in very fast hardware, one second of real time may provide a subjective year of thinking time, while a thousand years of real time spent on a passive storage medium may seem like no time at all. Can we then consider ourselves to be a mind without a body? Not quite.

A human totally deprived of bodily senses does not do well. After twelve hours in a sensory deprivation tank (where one floats in a body-temperature saline solution that produces almost no skin sensation, in total darkness and silence, with taste and smell and the sensations of breathing minimized) a subject will begin to hallucinate, as the mind, somewhat like a television tuned to a nonexistent channel, turns up the amplification, desperately looking for a signal, becoming ever less discriminating in the theories it offers to make sense of the random sensory noise it receives. Even the most extreme telepresence and virtual reality scenarios we have presented avoid complete bodylessness by always providing the mind with a consistent sensory (and motor) image, obtained from an actual remote robot body, or from a computer simulation. In those scenarios, a person may sometimes exist without a physical body, but never without the illusion of having one.

But in our computers there are already many entities that resemble truly bodiless minds. A typical computer chess program knows nothing about physical chess pieces or chessboards, or about the staring eyes of its opponent or the bright

lights of a tournament. Nor does it work with an internal simulation of those physical attributes. It reasons instead with a very efficient and compact mathematical representation of chess positions and moves. For the benefit of human players this internal representation is sometimes translated to a recognizable graphic on a computer screen, but such images mean nothing to the program that actually chooses the chess moves. For all practical purposes, the chess program's thoughts and sensations—its consciousness—is pure chess, with no taint of the physical, or any other, world. Much more than a human mind with a simulated body stored in a computer, a chess program is a mind without a body.

So now, imagine a future world where programs that do chess, mathematics, physics, engineering, art, business or whatever, have grown up to become at least as clever as the human mind. Imagine also the most of the inhabited universe has been converted to a computer network—a cyberspace—where such programs live, side by side with downloaded human minds and accompanying simulated human bodies. Suppose that all these entities make their living in something of a free market way, trading the products of their labor for the essentials of life—in this world memory space and computing cycles. Some entities do the equivalent of manual work, converting undeveloped parts of the universe into cyberspace, or improving the performance of existing patches, thus creating new wealth. Others work on physics or engineering problems whose solutions give the developers new and better ways to construct computing capacity. Some create programs that can become part of one's mental capacity. They trade their discoveries and inventions for more working space and time. There are entities that specialize as agents, collecting commissions in return for locating opportunities and negotiating deals for their clients. Others act as banks, storing and redistributing resources, buying and selling computing space, time and information. Some we might class as artists, creating structures that don't obviously result in physical resources, but which, for idiosyncratic reasons, are deemed valuable by some customers, and are traded at prices that fluctuate for subjective reasons. Some entities in the cyberworld will fail to produce enough value to support their requirements for existence—these eventually shrink and disappear, or merge with other ventures. Others will succeed and grow. The closest present day parallel is the growth, evolution, fragmentation and consolidation of corporations, whose op-

tions are shaped primarily by their economic performance.

A human would likely fare poorly in such a cyberspace. Unlike the streamlined artificial intelligences that zip about, making discoveries and deals, reconfiguring themselves to efficiently handle the data that constitutes their interactions, a human mind would lumber about in a massively inappropriate body simulation, analogous to someone in a deep diving suit plodding along among a troupe of acrobatic dolphins. Every interaction with the data world would first have to be analogized as some recognizable quasi-physical entity: other programs might be presented as animals, plants or demons, data items as books or treasure chests, accounting entries as coins or gold. Maintaining such fictions increases the

---

**We might then be tempted to replace some of our innermost mental processes with more cyberspace-appropriate programs purchased from the AIs, and so, bit by bit, transform ourselves into something much like them. Ultimately our thinking procedures could be totally liberated from any traces of our original body, indeed of any body.**

---

cost of doing business, as does operating the mind machinery that reduces the physical simulations into mental abstractions in the downloaded human mind. Though a few humans may find a niche exploiting their baroque construction to produce human-flavored art, more may feel a great economic incentive to streamline their interface to the cyberspace.

The streamlining could begin with the elimination of the body-simulation along with the portions of the downloaded mind dedicated to interpreting sense-data. These would be and replaced with simpler integrated simpler programs that produced approximately the same net effect in one's consciousness. One would

still view the cyber world in terms of location, color, smell, faces, and so on, but only those details we actually notice would be represented. We would still be at a disadvantage compared with the true artificial intelligences, who interact with the cyberspace in ways optimized for their tasks. We might then be tempted to replace some of our innermost mental processes with more cyberspace-appropriate programs purchased from the AIs, and so, bit by bit, transform ourselves into something much like them. Ultimately our thinking procedures could be totally liberated from any traces of our original body, indeed of any body. But the bodiless mind that results, wonderful though it may be in its clarity of thought and breadth of understanding, could in no sense be considered any longer human.

So, one way or another, the immensities of cyberspace will be teeming with very unhuman disembodied superminds, engaged in affairs of the future that are to human concerns as ours are to those of bacteria. But, once in a long while, humans do think of bacteria, even particular individual bacteria seen in particular microscopes. Similarly, a cyberbeing may occasionally bring to mind a human event of the distant past. If a sufficiently powerful mind makes a sufficiently large effort, such recall could occur with great detail—call it high fidelity. With enough fidelity, the situation of a remembered person, along with all the minutiae of her body, her thoughts, and feelings would be perfectly recreated in a kind of mental simulation: a cyberspace within a cyberspace where the person would be as alive there as anywhere. Sometimes the recall might be historically accurate, in other circumstances it could be artistically enhanced: it depends on the purposes of the cybermind. An evolving cyberspace becomes effectively ever more capacious and long lasting, and so can support ever more minds of ever greater power. If these minds spend only an infinitesimal fraction of their energy contemplating the human past, their sheer power should ensure that eventually our entire history is replayed many times in many places, and in many variations. The very moment we are now experiencing may actually be (almost certainly is) such a distributed mental event, and most likely is a complete fabrication that never happened physically. Alas, there is no way to sort it out from our perspective: we can only wallow in the scenery.

# Protecting privacy with electronic cash

by Hal Finney

How can we defend our privacy in an era of increased computerization? Today, our lives are subject to monitoring in a host of different ways. Every credit card transaction goes into a database. Our phone calls are logged by the phone company and used for its own marketing purposes. Our checks are photocopied and archived by the banks. And new "matching" techniques combine information from multiple databases, revealing even more detail about our lives. As computer databases grow, as more transactions take place electronically, over phone systems and computer networks, the possible forms of monitoring will grow with them<sup>1</sup>.

Predictably, most proposed solutions to this problem involve more government. One suggestion is to pass a set of laws designed to restrict information usage: "No information shall be used for a purpose different from that for which it was originally collected." Thus, income data collected by a bank through monitoring checking account activity could not be made available to mailing list companies; phone records could not be sold to telemarketing agencies, etc.

But this is a bad solution, for many reasons. The government is notoriously inefficient at enforcing existing laws, and the ease of collecting and using information suggests that it would be almost impossible to successfully enforce a law like this. The government also has a tendency to exempt itself from its own laws. It's unlikely that the IRS, for example, will happily give up the use of database matching, which it uses to track down tax evaders. And, of course, the very notion of trying to restrict the uses of information requires strict restrictions on the private actions of individuals which Extropians will find unacceptable.

But there is another solution, one advocated forcefully by computer scientist David Chaum of the Center for Mathematics and Computer Science in the Netherlands. While most people concerned with this problem have looked to paternalistic government solutions, Chaum has been quietly putting together the technical basis for a new way of organizing our financial and personal information. Rather than relying on new laws

and more government, Chaum looks to technical solutions. And these solutions rely on the ancient science devoted to keeping information confidential: cryptography.

Cryptography, the art of secret writing, has undergone a revolution in the last two decades, a revolution sparked by the invention of "public-key" cryptography. Seizing on this new technology, computer scientists have branched out into dozens of directions, pushing the frontiers of secrecy and confidentiality into new territory. And it is these new applications for cryptography which offer such promise for avoiding the dangers described above.

Chaum's approach to the protection of privacy can be thought of as having three layers. The first layer is public-key cryptography, which protects the privacy of individual messages. The second layer is anonymous messaging, which allows people to communicate via electronic mail ("email") without revealing their true identities. And the third layer is electronic money, which allows people to not only communicate, but to transact business via a computer network, with the same kind of privacy you get when you use cash. If you go into a store today and make a purchase with cash, no records are left tying you personally to the transaction. With no records, there is nothing to go into a computer database. The goal of electronic cash is to allow these same kinds of private transactions to take place electronically.

(Be aware that there are other proposals for "electronic money" which are not nearly so protective of individuals' privacy. Chaum's proposals are intended to preserve the privacy attributes of cash, so the term "digital cash" is appropriate. But other electronic replacements for cash not only lack its privacy, but would actually facilitate computer monitoring by putting more detailed information into databases, and by discouraging the use of cash. If you see a proposal for an electronic money system, check to see whether it has the ability to preserve the privacy of financial transactions the way paper money does today. If not, realize that the proposal is designed to harm, not help, individual privacy.)

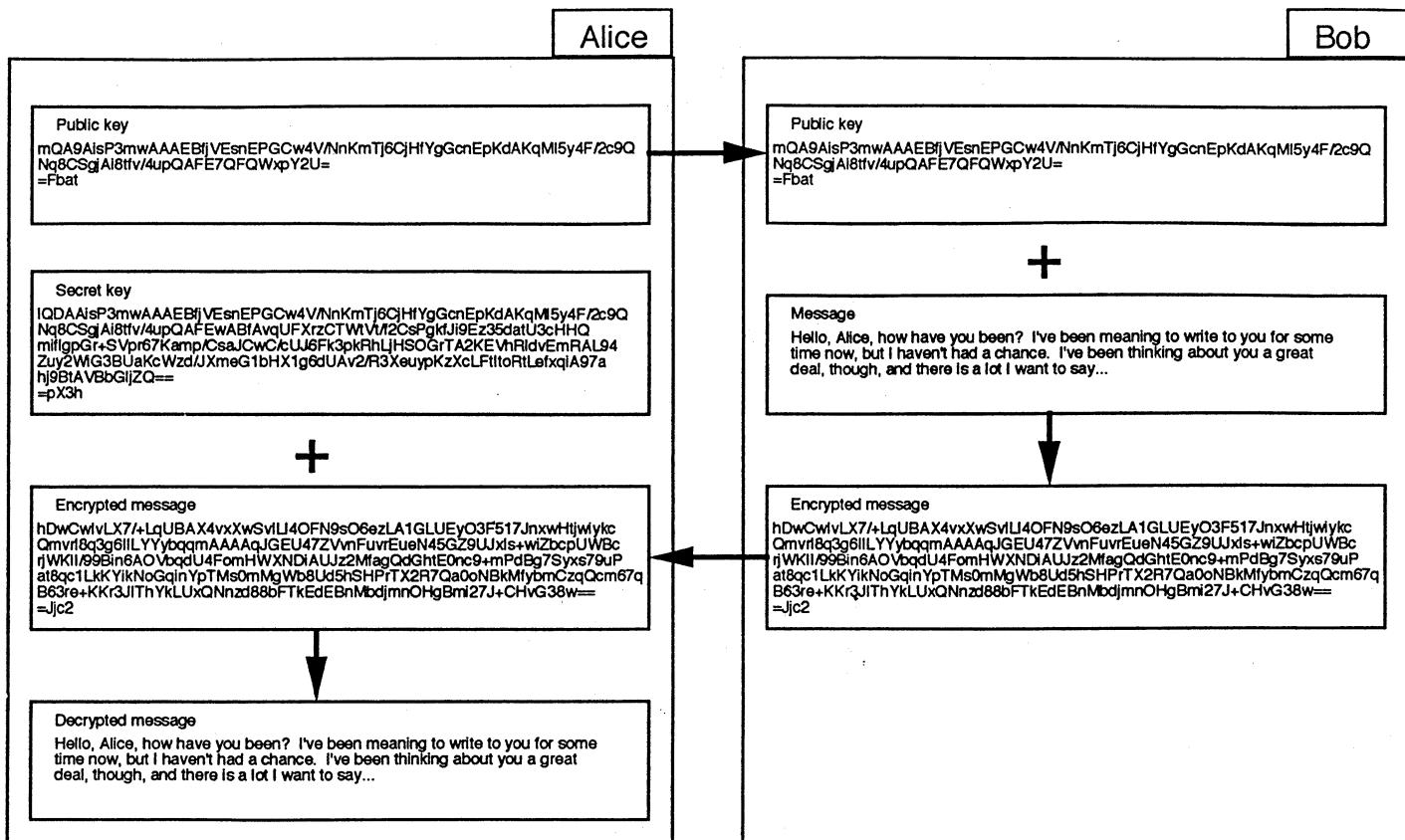


## Public-Key Cryptography

The first of the three layers in the privacy-protecting electronic money system is public-key cryptography. The basic concept of public-key cryptography, invented in 1976 by Diffie and Hellman<sup>2</sup>, is simple. Cryptographers have traditionally described an encryption system as being composed of two parts: an encryption method and a key. The encryption method is assumed to be publicly available, but the key is kept secret. If two people want to communicate, they agree on a secret key, and use that to encrypt and decrypt the message.

Public key cryptography introduced the idea that there could be two keys rather than one. One key, the public key, is known to everyone, and is used to encrypt messages. The other key, the secret key, is known only to you, and is used to decrypt messages. Public and secret keys are created in pairs, with each public key corresponding to one secret key, and vice versa.

So, to use a public-key system, you first create a public/secret key pair. You tell all your friends your public key, while keeping your secret key secret. When they want to send to you, they encrypt the message using your public key. The resulting encrypted message is readable only by using your secret key. This means that even the person who encrypted the message can't decrypt it! If he forgets what his original message said, and he deleted it, he has no chance of reconstructing the original. Only you can do that. This is the



**Figure 1**

paradox of public key cryptography - that a person can transform a message in such a way that they can't un-transform it, even though they know the exact formula used to make the transformation.

Figure 1 illustrates the steps involved in using a public-key system. (The keys and messages are based on actual output from Phillip Zimmermann's free public-key program, PGP.) Alice, on the left, first creates a public and secret key pair, the top two boxes on that side. The top box, the public key, she sends to her friend Bob, on the right. The second box is her secret key, which she keeps private. Bob, on the right, receives and saves Alice's public key. Then, when he wants to write to her, he composes a message, shown in the second box on that side. With a public-key encryption program like PGP, he encrypts the message using Alice's public key, producing output such as the third box on the right. This encrypted message is what he sends to Alice, as shown in the arrow leading back to the left side. Alice uses her saved secret key to decrypt the message from Bob, allowing her to reconstruct Bob's original message, shown as the last box on the left side.

There is no longer any need for public-key cryptography to be mysterious. There are now public-domain software packages which will let you experiment with public-key cryptography on your own computer, including Zimmermann's PGP and others. See the "Access" box for

information on how to get them.

### Anonymous Messages

Public-key cryptography allows people to communicate electronically with privacy and security. You can send messages safe from prying eyes using these techniques. But this is just a step towards the solution to the privacy problems we face. The next step provides the second layer of privacy: anonymous messages - messages whose source and destination can't be traced.

This is necessary because of the goal of providing in an electronic network the privacy of an ordinary cash transaction. Just as a merchant will accept cash from a customer without demanding proof of identity, we also want our electronic money system to allow similar transactions to take place, without the identity of the people involved being revealed to each other, or even to someone who is monitoring the network.

There are problems with providing anonymous messaging in current email systems. The national email networks are composed of thousands of machines, interconnected through a variety of gateways and message-passing systems. The fundamental necessity for a message to be delivered in such a system is that it be addressed appropriately. Typically an email address consists of a user's name, and the name of the computer system

which is his electronic "home". As the message works its way through the network, routing information is added to it, to keep a record of where the message came from and what machines it passed through en route to its destination. In this system, all messages are prominently stamped with their source and destination. Providing anonymous messages in such a system at first appears impossible.

Chaum has proposed two separate systems for overcoming this problem<sup>3</sup>. I will focus here on what he calls a "Mix" as it is simpler and more appropriate for the application of anonymous electronic mail. The notion of a Mix is simple. It is basically a message forwarding service. An analogy with ordinary paper mail may be helpful. Imagine that you want to send a letter to a friend, but in such a way that even someone monitoring your outgoing mail would not know that you were doing this. One solution would be to put your letter into an envelope addressed to your friend, then to place this envelope inside a larger envelope which you would send to someone else, along with a note asking them to forward the letter to your friend. This would hide the true destination of your mail from someone who was watching your outgoing envelopes.

Chaum's Mixes use this basic idea, but applied to email and improved with public-key cryptography. A Mix is a computer program capable of receiving email. It receives messages which contain

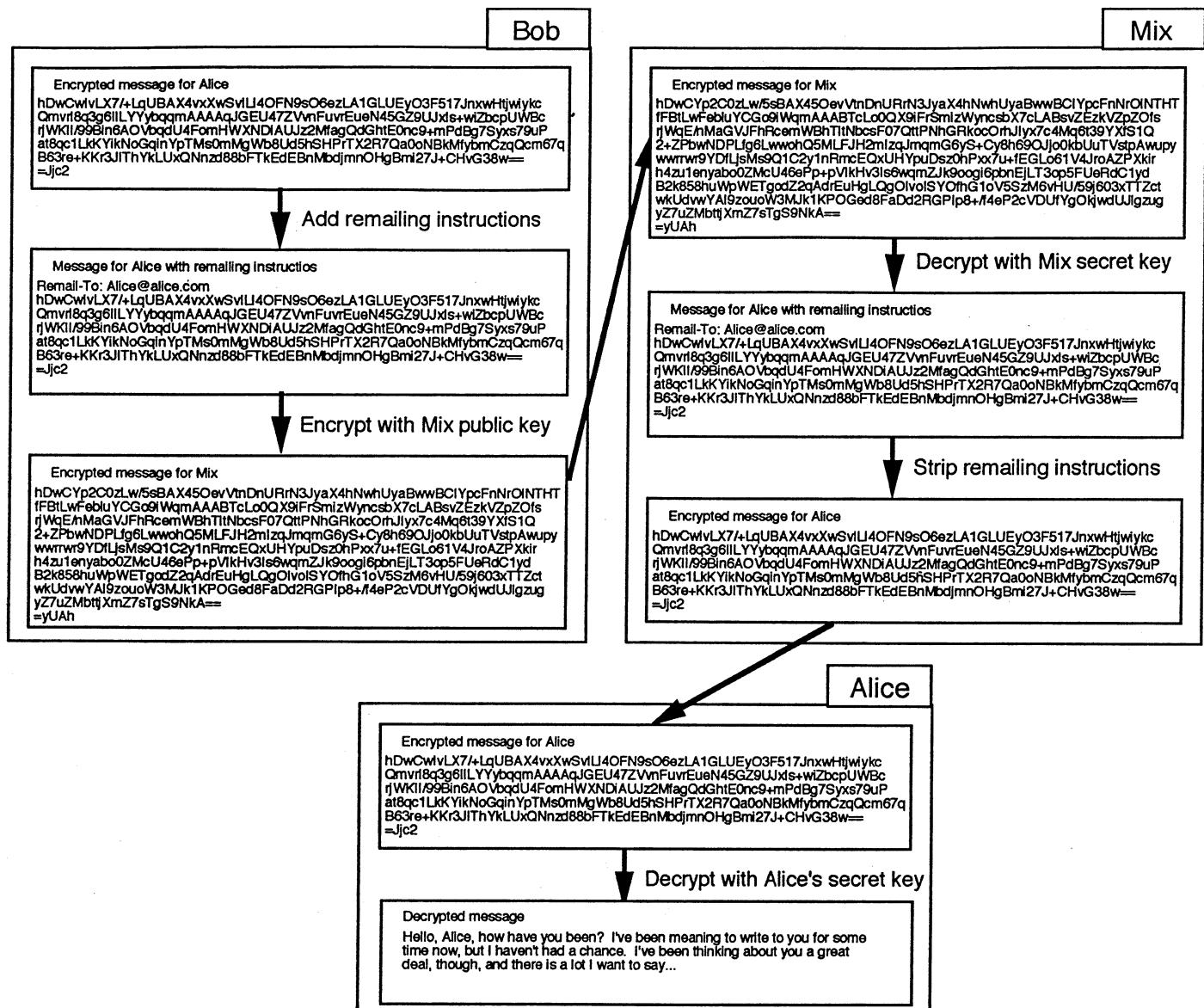


Figure 2

requests for remailing to another address, and basically just strips off these remailing instructions and forwards the messages as requested. Chaum adds security by having a different public key for each Mix. Now, instead of just sending the message with its forwarding request, the message plus forwarding info is encrypted with the Mix's public key before being sent to the Mix. The Mix simply decrypts the incoming message with its secret key, revealing the forwarding information, and sends the message on.

To protect the privacy of the sender, the Mix removes information about the original sender of the message before sending it. For even greater security, it's possible for the original sender to specify a "Cascade" of Mixes, a whole chain of Mixes that the message should go through before finally being sent to its destination. That way even if one of the Mixes is corrupt, it still can't determine who is

sending to whom.

Using Mixes, then, the basic requirement for anonymous mail is met. A message en route in the network does not have to reveal its source and destination. It may be coming from a Mix, going to a Mix, or some combination of these.

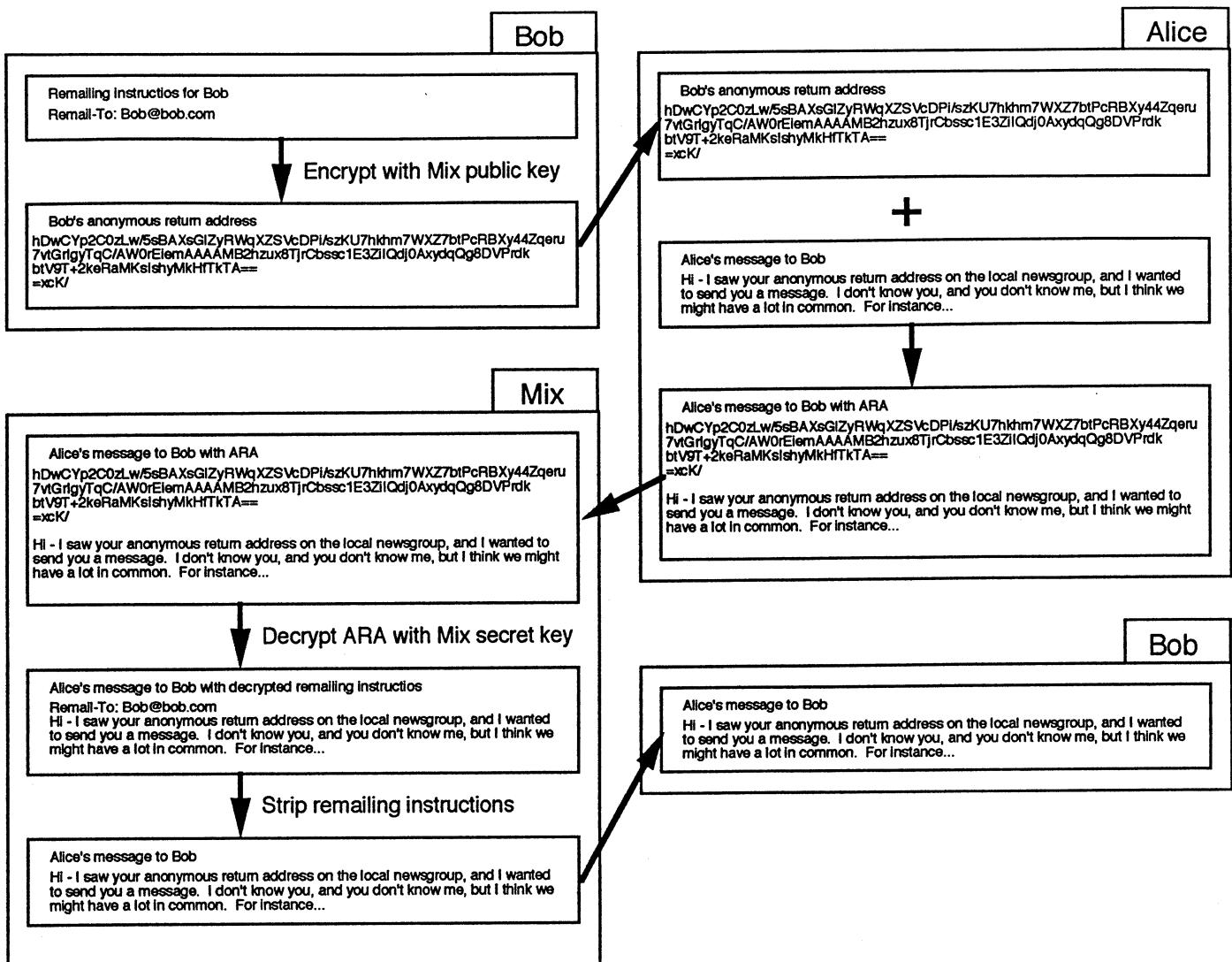
Figure 2 shows an example of an anonymous message as it is forwarded through a Mix, using public-key cryptography to protect its privacy. As in Figure 1, Bob wants to send his encrypted message to Alice, but this time he wants to use a Mix to provide more confidentiality. Starting with the encrypted message from Figure 1, Bob (on the left, this time) first adds remailing instructions which will be interpreted by the Mix. These will include Alice's email address in some format specified by the Mix. (This example uses a simplified form of commands currently being used in experimental remailers.) Then he encrypts the whole message with

the Mix's public key and sends it to the Mix.

Upon receipt, the Mix reverses the steps which Bob applied. It decrypts the message using its own secret key, then strips off the remailing instructions which Bob added. The resulting message (which the Mix can't read, being encrypted using Alice's secret key) is then forwarded to Alice as specified in the remailing instructions. As before, Alice receives and decrypts the message using her secret key. But this time, the message path has been protected by the Mix, and the fact that Alice and Bob are communicating is kept confidential.

## Anonymous Return Addresses

We need something more advanced than message anonymity for truly private mes-



**Figure 3**

saging, though. These anonymous messages are basically “one-way”. I can send you a message, with the source and destination hidden, and when you receive the message you won’t have any way of knowing who sent it. This means that you can’t reply to me. We need the ability to have such replies.

Here, we have a seemingly paradoxical requirement: being able to reply to someone without knowing either who they are or what their email address is. Chaum shows how this can be solved using public-key cryptography and Mixes. The basic idea is what Chaum calls an Anonymous Return Address (ARA). In its simplest form, I create an ARA by taking my regular email address and encrypting it with the public key of a particular Mix—call it MixA for this example. I send this resulting block of encrypted text along with my message to you, through a Cascade of Mixes.

Now, when you receive the message, you see no return address, but you do see the block of text that is the ARA. You can

reply to me without knowing who I am by sending your reply back to MixA, along with the ARA itself. MixA decrypts the ARA using its secret key, getting back my original email address that I encrypted. Using this email address, it is able to forward the mail to me. I was able to receive this message from you, although you have no knowledge of my true identity.

Figure 3 shows this process graphically. Bob, in the upper left, creates his ARA by encrypting a remailing command, similar to what was used in Figure 2, with the Mix’s public key. He then includes this ARA in messages which he anonymously sends or publicly posts. In the example, Alice sees Bob’s ARA and wishes to respond to him, even though she doesn’t know his email address. She composes her message, in the second box on the right, then combines her message with Bob’s ARA. The combined message is sent to the Mix. The Mix now uses its secret key to decrypt the ARA portion of the message, revealing the remailing in-

structions which Bob encrypted to create the ARA. The remainder of the process is just as in Figure 3. The Mix strips off the remailing request and forwards the message to Bob’s address, as shown.

These tools open many possibilities. With Mixes, Cascades, and ARAs, people can communicate without knowing other people’s true identities. You can make an anonymous posting to a public message board, include your ARA, and receive replies from scores of people who don’t know who you are. Some of them may reply anonymously and include their own ARAs. People can end up communicating with each other with none of them knowing the true identity of any of the others.

(Some “Chat” or “CB Simulator” systems today offer the illusion of such anonymous communication, but in most cases the system operators can easily break through the cover of handles and pseudonyms and discover true identities. With a Cascade of Mixes, no single Mix can establish this relationship. As long as

even one Mix of the Cascade remains uncorrupted, your identity is safe.)

## Digital Pseudonyms

Anonymous messages bring forth the usefulness of "digital pseudonyms," another concept from Chaum. With no identification of the source of messages, there would seem to be no way of verifying that two messages came from the same person. There could be a problem with imposters pretending to be other people, resulting in utter confusion. To solve this problem, we need another concept from public-key cryptography: the "digital signature."

As described above, public-key cryptography allows messages encrypted with my public key to be decrypted with my secret key. However, it works the other way around as well. Messages can be encrypted with my *secret* key and then decrypted only with my *public* key. This property is what is used to implement the digital signature. If I encrypt a document with my secret key, anyone can decrypt it with my public key. And since my secret key is secret, only I can do this type of encryption. That means that if a document can be decrypted with my public key, then I, and only I, must have encrypted it with my secret key. This is considered a digital signature, in the sense that it is a proof that I was the one that "signed" (that is, encrypted) the document.

The digital signature concept can be used to solve the imposter problem by allowing for "digital pseudonyms." My digital pseudonym is simply a public/secret key pair, where, as usual, I let the public part be known. Typically, I'd publicize it along with my ARA. Now, to prove that a given message is from me and no one else, I sign the message using the secret key of my digital pseudonym. Any set of messages signed by that same digital pseudonym is therefore known to come from me, because only I know the secret key. People may not know who I am, but I can still maintain a stable public persona on the computer nets via my digital pseudonym. And there is no danger of anyone else successfully masquerading as me.

With public-key cryptography, Mixes, and digital pseudonyms, we have all we need for a network of people communicating privately and anonymously. Now, we need a way for them to transact business while maintaining these conditions.

## Electronic Money

The next step, the third layer in our description, is digital cash-electronic money. Cash, ordinary folding paper money, is one of the last bastions of privacy in our financial lives. And many of the problems described above - the losses of privacy, the increase in computerized infor-

mation - could be avoided if cash could be used more easily. But cash has many disadvantages. It can be lost, or stolen, and it's not safe to carry in large quantities. Also, it is useless for purchases made electronically, over the phone or (in the future) over computer networks. Digital cash is designed to combine the advantages of electronic payment systems - the safety and convenience - with the advantages of paper money - the privacy and anonymity.

Once again, we are faced with paradoxes in the notion of digital cash. Since digital cash may be sent by email and other electronic methods, it must basically be an information pattern - in concrete terms, some pattern of letters and numbers. How could such a string of characters have value, in the same sense that the dollar bill in your wallet does? What about counterfeiting? Couldn't another copy of the character string be created trivially? What prevents a person from "spending" the same money twice?

To answer these questions we turn again to public-key cryptography. Realize, though, that electronic money is an active area of research in cryptography. Many people have proposed different systems for electronic cash, each of which has its own advantages and disadvantages. I will present here a simplified concept to give a feel for the problems and solutions which exist<sup>4</sup>.

One way to think of digital cash is by analogy to the early days of paper money. At one time, paper money was not the monopoly of governments that it is today. Instead, paper money was "bank notes", often given as receipts for the deposit of gold or similar "real money" in bank vaults. These notes would carry a description of what they were worth, such as, "Redeemable for one ounce of gold." A particular bank note could be redeemed at the issuing bank for its face value. People used these bank notes as we use paper money today. They were valuable because they were backed by materials of value in the bank vaults.

In a sense, then, a bank note can be viewed as a signed document, a promise to perform a redemption for the bearer who presents it at the bank. This suggests a way of thinking of digital money. Instead of a paper note with an engraved signature, we instead would use an electronic mail message with a digital signature.

An electronic bank could, like the banks of old, have valuable materials in its vaults. Today, these would likely be dollars or other government currency, but they could be gold or other commodities. Using these as backing, it would issue bank notes. These would be electronic messages, digitally signed by the bank's secret key, promising to transfer a specified sum to the account of whom ever presented the note to the bank (or, if desired, to redeem the note in dollars or

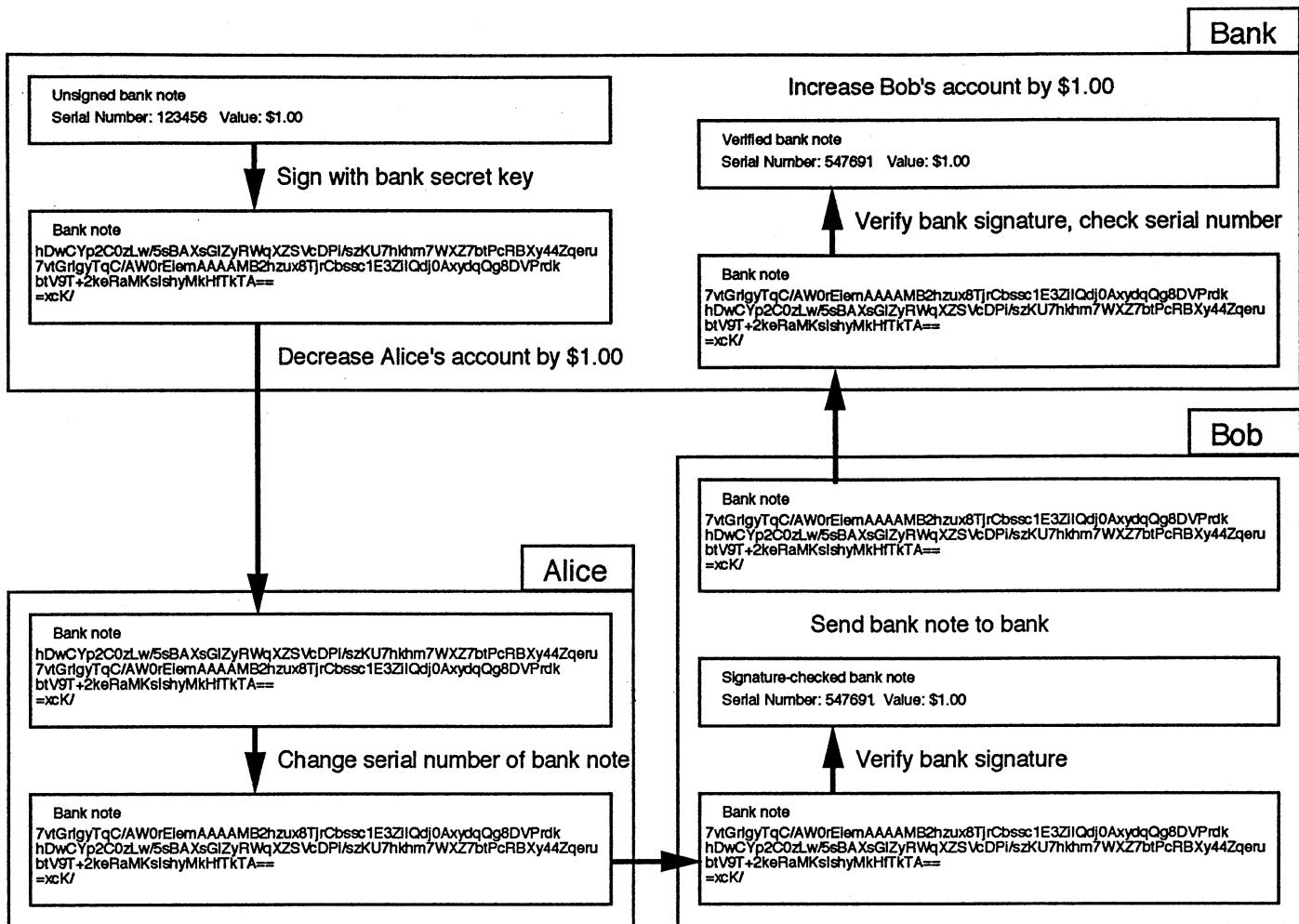
other valuables.)

Here is how it might work. You open an account with an electronic bank, depositing some money as in any bank. The bank then credits your account with your initial balance. Now, suppose you are going to want to make an electronic payment to me. Prior to any transactions, you would send a message to the bank, requesting one or more bank notes in specified denominations. (This is exactly analogous to withdrawing cash from your regular bank account.) The bank debits your account, creates new bank note messages, and sends them to you. They are sent to you as signed messages, encrypted with the bank's secret key. When decrypted with the bank's public key, which everyone knows, a one-dollar digital bank note would say, in effect, "This note is worth \$1.00, payable on demand." It would also include a unique serial number, like the serial number on a dollar bill.

The serial number is important; as we will see below, it is used by the bank to make sure that a particular note is accepted for deposit only once. But putting serial numbers on the bank notes hurts anonymity; the bank can remember which account a bank note was withdrawn from, and then when it is deposited the bank will know that the depositer is doing business with the withdrawer. To avoid this, Chaum introduces a clever mathematical trick (too complex to describe here) which allows the serial number to be randomly changed as the note is withdrawn from the bank. The bank note still retains its proper form and value, but the serial number is different from the one the bank saw. This allows the bank to check that the same note isn't deposited more than once, while making it impossible for the bank to determine who withdrew any note that is deposited.

When you are ready to purchase something from me, you simply email me the appropriate bank note messages. I can check that they are legitimate bank notes by using the bank's public key to verify its signature. I then email the notes to the bank, which checks that the account numbers on the notes have not been deposited before this. If they are valid bank notes, the bank credits my account for the face value of the notes. Your account was decreased when you withdrew the bank notes, which you held like cash, and mine was increased when I sent them to the bank. The result is similar to how it would work if you withdrew (paper) cash from the bank, mailed it to me, and I deposited the cash in my own account.

Figure 4 shows a similar transaction between Alice and Bob. The bank, in the upper left corner, creates a digital bank note by signing a message which specifies the serial number and value of the note, and sends it to Alice. Alice, as she withdraws it, uses Chaum's technique to alter the serial number so that the bank will not recognize the note as being from this with-



**Figure 4**

drawal. She then pays Bob electronically by sending the bank note to him. Bob checks the note's validity by decrypting using the bank's public key to check its signature. He then sends the note to the bank, which checks the serial number to confirm that this bank note hasn't been spent used before. The serial number is different from that in Alice's withdrawal, preventing the bank from linking the two transactions.

With this simple picture in mind, we can begin to answer some of the objections listed above. Bank notes cannot be forged because only the bank knows the secret key that is used to issue them. Other people will therefore not be able to create bank notes of their own. Also, anyone can check that a bank note is not a forgery by verifying the bank's digital signature on the note. As for the copying issue, preventing a person from spending the same bank note more than once, this is handled by checking with the bank to see if the serial number on the note had been used before before accepting a bank note as payment. If it had been, the note would not be accepted. Any attempt to re-use a bank note will be detected because the serial number will be a duplicate of one

used before. This means, too, that once you "spend" your digital cash by emailing it to someone, you should delete it from your computer, as it will be of no further value to you.

This simple scheme gives some of the flavor of electronic cash, but it still has awkward features. The need to check with the bank for each transaction may be inconvenient in many environments. And the fixed denominations of the bank notes described here, the inability to split them into smaller pieces, will also limit their usefulness. Chaum and others have proposed more complex systems which solve these problems in different ways.<sup>5</sup> With these more advanced systems, the anonymity, privacy, and convenience of cash transactions can be achieved even in a purely electronic environment.

### Electronic Money in Practice

Having described the three layers of privacy protection, we can now see how electronic transactions can maintain individual privacy. Public-key cryptography protects the confidentiality of messages, as well as playing a key role in the other layers. Anonymous messaging further

allows people to communicate without revealing more about themselves than they choose. And electronic money combines the anonymity of cash with the convenience of electronic payments. David Chaum has described variations of these techniques that can extend privacy protection to many other areas of our lives as well.<sup>6</sup>

Although my description of digital cash has been in terms of computer networks with email message transactions, it can be applied on a more local scale as well. With credit card sized computers, digital cash could just as easily be used to pay for groceries at the local supermarket as to order software from an anonymous supplier on the computer networks. "Smartcard" computers using digital cash could replace credit or debit cards for many purposes. The same types of messages would be used, with the interaction being between your smart card and the merchant's card reader.

On the nets themselves, any goods or services which are primarily information-based would be natural candidates for digital cash purchases. Today this might include such things as software, electronic magazines, even electronic books. In the

future, with higher-bandwidth networks, it may be possible to purchase music and video recordings across the nets.

As another example, digital cash and anonymous remailers (such as Chaum's Mixes) have a synergistic relationship; that is, each directly benefits the other. Without anonymous remailers, digital cash would be pointless, as the desired confidentiality would be lost with each transaction, with message source and destination blatantly displayed in the electronic mail messages. And in the other direction, digital cash can be used to support anonymous remailing services. There could be a wide range of Mix services available on the nets; some would be free, and presumably offer relatively simple services, but others would charge, and would offer more service or more expensive security precautions. Such for-profit remailers could be paid for by digital cash.

What are the prospects for the eventual implementation of digital cash systems and the other technologies described here? Some experiments are already beginning. David Chaum has started a company, DigiCash, based in Amsterdam, which is attempting to set up an electronic money system on a small scale. As with any new business concept, though, especially in the conservative financial community, it will take time before a new system like this is widely used.

The many laws and regulations covering the banking and financial services industries in most Western nations will undoubtedly slow the acceptance of digital cash. Some have predicted that the initial success of electronic money may be in the form of a technically illegal "black market" where crypto-hackers buy and sell information, using cryptography to protect against government crackdowns.

In the nearer term, the tools are in place now for people to begin experimenting with the other concepts discussed here. Public-key cryptography is becoming a reality on the computer networks. And experimental remailers with integrated public-key cryptosystems are already in use on a small scale. Digital-pseudonym-based anonymous message posting should begin happening within the next year. The field is moving rapidly, as privacy advocates around the world hurry to bring these systems into existence before governments and other large institutions can react. See the "Access" box for information on how you can play a part in this quiet revolution.

We are on a path today which, if nothing changes, will lead to a world with the potential for greater government power, intrusion, and control. We can change this; these technologies can revolutionize the relationship between individuals and organizations, putting them both on an equal footing for the first time. Cryptography can make possible a world in which people have control over information about themselves, not because government has granted them that control, but because only they possess the cryptographic keys to reveal that information. This is the world we are working to create.

## Notes

<sup>1</sup>For a review of the status of current monitoring technology, see [Clarke 88].

<sup>2</sup>See [Diffie 76].

<sup>3</sup>The "Mix" is described in [Chaum 81]. Chaum's other solution, the "DC-Net", is described in [Chaum 88A].

<sup>4</sup>The electronic money scheme I describe is a simplification of Chaum's first proposal in [Chaum 88B].

<sup>5</sup>For more proposals about electronic cash, see: [Even 83], [Chaum 85], [Okamoto 89], [Okamoto 90], [Hayes 90], and [Chaum 90].

<sup>6</sup>See [Chaum 85] and [Chaum 92].

## Access

### Public-Key Cryptography

Philip Zimmermann's free program PGP ("Pretty Good Privacy") is a widely available implementation of public key cryptography. It features high speed and has excellent key management, and operates on many systems, including PC compatibles, Macintoshes, and most Unix-based workstations. At publication time, version 2.1 was current. Readers with Internet access should be able to find PGP on such hosts as princeton.edu (/pub/pgp20) and pencil.cs.missouri.edu (/pub/crypt). Many of the larger bulletin-board systems carry PGP as well. Send email to Hugh Miller at <info-pgp-request@lucpul.it.luc.edu> for current information, or check the Usenet newsgroup alt.security.pgp.

Mark Riordan's free program RIPEM was in beta test at press time, with a release expected soon. Contact the author at <mrr@scss3.cl.msu.edu> for information about availability.

The Internet PEM ("Privacy Enhanced Mail") standard was due to be completed soon at press time. PEM uses a key-management hierarchy in which users register their public keys with a centralized organization. Free software implementing the basic public-key algorithms was expected to be available soon after the standard is finalized. Mail to <pem-dev@tis.com> for more information on availability.

### Anonymous Remailers

An email discussion group exists which is devoted to the topics of encryption, remailers, digital cash, and other topics related to these. Experimental anonymous remailers are under development at press time and should be widely available soon. Contact <cypherpunks-request@toad.com> for information.

## References

- [Chaum 81] Chaum, D., *Untraceable Electronic Mail, Return Addresses and Digital Pseudonyms*. Communications of the ACM, vol. 24, n. 2, p. 84-88, February, 1981.
- [Chaum 85] Chaum, D., *Security without Identification: Transaction Systems to make Big Brother Obsolete*. Communications of the ACM, vol. 28, n. 10, p. 1030-1044, October, 1985.
- [Chaum 88A] Chaum, D., *The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability*. Journal of Cryptology, vol. 1, p. 65-75, 1988.
- [Chaum 88B] Chaum, D., Fiat, A., Naor, M., *Untraceable electronic cash*. In: Advances in Cryptology - CRYPTO '88, p. 319-27, 1988.
- [Chaum 90] Chaum, D., *Showing Credentials without Identification: Transferring Signatures between Unconditionally Unlinkable Pseudonyms*. In: Advances in Cryptology - AUSCRYPT '90, p. 246-64, 1990.
- [Chaum 92] Chaum, D., *Achieving Electronic Privacy*. Scientific American, vol 267, n. 2, p. 96-101, August, 1992.
- [Clarke 88] Clarke, R., *Information Technology and Dataveillance*. Communications of the ACM, vol. 31, n. 5, p. 498-512, May, 1988.
- [Diffie 76] Diffie, W., Hellman, M., *New Directions in Cryptography*. IEEE Transactions on Information Theory, November, 1976, p. 644.
- [Even 83] Even, S., Goldreich, O., *Electronic Wallet*. In Advances in Cryptology - CRYPTO '83, p. 383-386, 1983.
- [Hayes 90] Hayes, B., *Anonymous One-Time Signatures and Flexible Untraceable Electronic Cash*. In: AusCrypt '90, p. 294-305, 1990.
- [Okamoto 89] Okamoto, T., Ohta, K., *Disposable Zero-Knowledge Authentications and Their Application to Untraceable Electronic Cash*. In: Advances In Cryptology - CRYPTO '89, p. 481-496, 1989.
- [Okamoto 90] Okamoto, T., Ohta, K., *Universal Electronic Cash*. In: Advances in Cryptology - CRYPTO '90, p. 324-337, 1990.

# Technological self-transformation

## Expanding personal extropy

Max More  
Extropy Institute



We, however, *want to be those who we are* – the new, the unique, the incomparable, those who give themselves their own law, those who create themselves!

[Friedrich Nietzsche, *The Gay Science*, p335]

‘Will a self.’ – Active, successful natures act, not according to the dictum ‘know thyself’, but as if there hovered before them the commandment: will a self and thou shalt become a self.

[Friedrich Nietzsche, *Assorted Opinions and Maxims*.]

### Self-transformation and personal extropy

Self-transformation is a process that increases personal extropy. Extropy is a measure of a system’s intelligence, information content, available energy, longevity, vitality, diversity, complexity, and capacity for growth. Clearly, I intend “self-transformation” to necessarily imply “*positive* self-transformation.” The changes that I will discuss, rather than being value-neutral, all in some way *amplify* the extropy in your life and person: They make you more intelligent and wiser, physically healthier and more vigorous, increasingly psychologically effective, more creative, rational, and productive, and more effective at gathering and filtering information.

The Extropian philosophy of life sees self-transformation as a primary virtue. A virtue is a psychological characteristic, a moral excellence that propels us to live superbly in a particular

manner. The virtue of *independence*, for example, is the characteristic exhibited by a person who makes their own decisions, using their own judgment, rather than allowing their decisions to be made by others by default. The virtue of *self-transformation* is a characteristic that reflects and empowers a person’s drive for physical, intellectual, moral, and psychological excellence. A commitment to self-transformation means a refusal to acquiesce in mediocrity, a questioning of limits to one’s potential, and a drive to perpetually overcome psychological, social, physiological, genetic, and neurological constraints.

Self-transformation is more than a mere preference. Ayn Rand identified the fundamental and ultimate choice perpetually confronting every organism: To live or to die. Without necessarily endorsing Rand’s view that *all* values and virtues can strictly be derived from the ultimate choice to live, I do affirm the claim that a commitment to successful and fulfilling living – Aristotle’s vision of human flourishing and Nietzsche’s will to power – requires a conscious effort to overcome ourselves, to push past all limits to our growth. Self-transformation is a virtue because it promotes our survival, our efficacy, and our well-being. As a dynamic process of self-overcoming, an internally generated drive to grow and thrive, it is the very essence and highest expression of life:

The nature of self-transformation provides part of the reason why the Extropian philosophy has no place for notions of gods and the supernatural. As Feuerbach and Nietzsche realized, positing a divine realm of perfection undermines life since ‘God’ and ‘spirit’ are reactive concepts, being defined in terms of what physical life is not. In the major religions, the World is base and corrupt, our physical existence transitory and inferior. Allegiance to life consists in resolving to seek vitality within, not from an external deity or a supernatural realm. A conception of a perfect, infinite, eternal, and wholly good supernatural realm, contrasted with the world of experience, degrades and strips our world and our physical selves of worth and significance. Life is fundamentally a ceaseless process, whose quintessence is a self-overcoming, a progression, a self-transformation and self-augmentation. Life’s purpose is not mere survival, for its energies remain once that basic goal is secured; nor is its purpose a drive to serve or glorify anything external, for then it would be self-alienating. The essence of life is what Nietzsche called the will to power – life’s perpetual drive toward its own increase and excellence. Extropic life can thus never manifest self-sacrifice or worship of superior beings.

The practice of self-transformation interacts synergistically with the other Extropian principles, as they are embodied in an individual’s actions.<sup>1</sup> *Dynamic optimism* serves to motivate continuous personal transformation, and to sustain transformative efforts, conquering barriers and overwhelming discouragement.

ment during times of difficulty or weariness. *Boundless expansion* as a society provides the context required for us to sustain truly long-term personal progress, to provide energy, space, and the framework for the diversity implicit in individual self-transformation. *Intelligent technology* in this context means directing science and technology toward the transcendence of our hitherto inherently limited abilities. *Spontaneous order* is an enabling condition, allowing each individual to pursue his or her *self*-transformation with minimal interference from others and to maximum mutual benefit from the resulting diversity.

## ST and extreme longevity

The extropian commitment to self-transformation coheres naturally with the extropian desire for extreme longevity and the quest for physical immortality. Practically immortal transhumans will need both to modify themselves continually to keep up with the world, and to ensure fulfillment over the long term. Completion of the ancient alchemists' quest for the key to abolishing aging is in sight. After crawling along for years with only minor successes (such as dietary restriction and Co-enzyme Q10 experiments), over the last five years interventionist gerontology has begun to employ the tools of molecular biology. Progress in both theory and practice is accelerating, and will make another leap forward once the analytical and interventionist tools of nanomedicine are brought to bear, probably by the end of the first quarter of the 21st century. It is no longer so radical to claim that we can expect enormous extension of the human lifespan by the middle of the 21st Century (and *perhaps* much sooner).<sup>2</sup>

Many of those who think superficially about the possibilities and consequences of indefinite lifespans paint a picture of stagnant persons grown bored with life. They picture those advanced in age to be psychologically decrepit. They equate deep maturity with boredom and ennui. This is a false image, except for those who (at *any* age) choose a passive life of stagnation despite the pressures of change. We need not look into the next century to see how this projection obscures rather than illuminates. Even now, with estimated lifespans approaching 80 years (ignoring future gerontological advances), the conditions of life relevant to personal transformation have undergone drastic changes as compared to the historically recent past. When most people died in their thirties or forties, or younger, and technology, culture, and social orga-

nization changed at a glacial rate, personal transformation was neither necessary nor lauded. A man working on the land in 16th century Europe felt no need to challenge his limits in favor of an innovative life. His job would likely remain unchanged throughout his entire life. He would probably be married for life, and the organization of his family would be set – no troubling alternative lifestyle to upset his equanimity. He would rarely, if ever, have to rise to the challenge of new technologies, important new ideas, or significant political change.

Now, in 1993, few people can get through life without undergoing major change, and almost no one (the Amish may be exceptions) can avoid some changes to their lifestyle. Increasingly absurd in a young person is the belief that they can train for a career and stay in that career indefinitely. Workers in manufacturing have had to learn service skills; business executives have had to become proficient at using computers; and doctors have had to adapt to new diagnostic and treatment modalities. Family structure is no longer a given, to be entered into unconsciously, and by which to be bound for life. The male provider and head of the household is being challenged, as is the need for a large family, a partner of the other sex, and lifelong commitments. We have a growing choice of ideas, gadgets, cultures, sports, and games to choose from, and the choices grow at an accelerating rate due to population growth and economic pressures.

So, even now, self-transformation – learning new skills, modifying habits, selecting new interests and behaviors – is necessary if we are to stay involved in our protean world. Whereas the Old World smothered personal innovation in a sea of stasis, the contemporary world repeatedly electrifies us with the charge of change. The 21st Century – the era of the transition from the human to the transhuman – can only boost this current. While some conservatives will always seek to stagnate as far as possible, the pressures will all favor personal transformation. Biotechnology, nanotechnology, neural networks, synthetic intelligence, expansion into space, intelligence intensification, and neurochemical modification (plus innovations as yet unforeseen) will ensure the flow of change and the widening field of choice. We are used to associating 'advanced age' (as we now think of it) with lack of vigor, ill health, and senility. But the centenarians of next century will appear youthful and exude energy. Not only will physical illness become practically unknown, we will fully understand the basis of depression and lack of

enthusiasm, allowing us to choose to maintain ourselves in a perpetually high energy condition. With the termination of aging, chronic illness, and depression, advanced age will cease to imply weariness, retirement, or resignation.

Apart from transforming ourselves in order to keep up, we would-be immortals will find self-transformation necessary for a fulfilling, meaningful life. Whereas old views held that a meaningful life required a strong and stable bond to a particular community, a particular social role, and a particular god, the Extropian view sees meaning and fulfillment partly in the bonds one *chooses* to form, and in the *process* of growth, renewal, and the dissolving of old bonds and the forming of new ones.<sup>3</sup> A stimulating, challenging, and fulfilling life will require periodic, though not continual, metamorphosis. Rather than permanently retiring after six decades, the long-lived will take periodic temporary retirements in order to reflect on their current life, to slow the pace for a time, or to learn new skills. No matter how long we live, we can always find new interests, new fields of study, new friends, new cultures and subcultures, and new sports and games. By avoiding stasis, we can forever elude the existential boredom typified in tales such as Čapek's *The Makropulos Case*. Some may eventually choose to end their otherwise limitless lives, as does Elina Makropulos, but only those who do not challenge themselves to transform.

A long-lived and deeply mature person will be quite different from the humans of today (even ignoring the technological augmentations sketched below). In comparison, all of us today are callow, undeveloped infants. We make our decisions based on a narrow perspective arrived at after a small number of years, like the view of a dark auditorium illuminated by a solitary spotlight. Our senior selves will have come to understand their own and others' motivations, desires, and behavior far more deeply; they will have experimented with many more ideas, cultures, and relationships. These senior selves will look back on their first century of life, and see their early selves as immature and impulsive, ignorant and ignoble, making decisions largely in ignorance of the world and of their own selves. Up until the 20th Century, all the experience and wisdom accumulated by the oldest persons has been degraded by old age, and annihilated by death. The elders of the future will be able to build on their learning, evolving a level of sophistication and maturity that we mere neonates are incapable of fully comprehending.

Immortalism and self-transforma-

tion belong together in another way: Some of us now living will not remain alive until the abolition of aging. In light of this, longevists with sufficient foresight, independence, and determination, are making arrangements to have themselves placed into biostasis (currently in the unperfected form of cryonic suspension) in the event of cessation of life functions resulting from disease, accident, or old age. Opting for biostasis is a probable life-saver, but brings with it the possibility of true future shock. Patients remaining in biostasis for more than a few years will be greeted by a joltingly different world. The prospect of such a rude jolt shooting them from a familiar to a strange world has been enough to frighten people away from the idea of biostasis. The prospect unseats those people so much that they choose risking becoming food for worms over facing an alien future. The more committed we are to self-transformation, the less fear we will have of a sudden jump into the future. We will be familiar with the unfamiliar. We will be experienced in adapting, learning, and innovating. We may even reflect on this prospect more with excitement than trepidation.

## Self-definition

One thing is needed – ‘To give style’ to one’s character – a great and rare art! He exercises it who surveys all that his nature presents in strength and weakness and then molds it to an artistic plan until everything appears as art and reason, and even the weaknesses delight the eye.

[Friedrich Nietzsche, *The Gay Science*, p.290]

The history of humanity has been a history of the growth of our ability to define ourselves as individuals. Our future will see the continuation and deepening of this evolutionary process. The primitive lifeforms from which we evolved were completely defined by forces external to choice, such as genetic and environmental determinants. The continued importance of these factors manifests itself in the tired “nature vs. nurture,” or genetic heritage vs. environment, debate. Yet we, more than any previous organisms or earlier humans, have the power to define ourselves, to choose who and how to be. As Richard Dawkins argued in *The Selfish Gene*, as conscious beings we can understand and thus allow for the imperatives of our genes. We can rebel against the tyranny of the selfish replicators. This capacity for rebellion against our genetic programming is what allows us to control our reproduction, to redirect our sexual

# DEFINITIONS

**Biological fundamentalism:** A new conservatism that resists asexual reproduction, genetic engineering, altering human anatomy, overcoming death. A resistance to the evolution from the human to the posthuman.

**Ideal identity:** A internal model of our personality as we wish it to be; the person we seek to become.

**Immortalist:** A person who believes in the possibility of, and who seeks to attain, physical immortality.

**Longevist:** A person who seeks to extend their life beyond current norms (but who may not wish to live forever).

**Morphological freedom:** The ability to alter bodily form at will through technologies such as surgery, genetic engineering, nanotechnology, uploading.

**Nanomedicine:** The use of molecular-scale devices to repair damage and boost the immune system.

**The Net:** The interlinked collection of computer networks, including the Internet, allowing remote conversation, data processing, and information retrieval.

**Optimal Persona:** A personally constructed and sustained model of the person into whom you intend to develop.

**Smart drugs/nootropics:** Substances that, without negative side-effects, can enhance retention, recall, and concentration.

**Transhuman:** Someone in the transition stage from human to biologically, neurologically, and genetically posthuman. One who orients his/her thinking towards the future to prepare for coming changes and who seeks out and takes advantage of opportunities for self-advancement.

**Transbiomorphosis:** The transformation of the human body from a natural, biological organism into a superior, consciously designed vehicle of personality.

**Uploading:** The transference of personality patterns embodied in the brain to an appropriately configured supercomputer, allowing the same person to live in more powerful hardware.

energies, and to rethink and reorganize gender roles and family structures. As our environments have grown more diverse, so our range of choices has grown. Many humans may continue to be programmed by aspects of their environment (incoming information, family upbringing, geographical location, political ideology, predominant morality, religion, etc.), but now they must usually choose between competing programming forces. Here lies the budding of autonomy. The existence of diverse options facilitates – but does not guarantee – that any individual will make *conscious* choices.

The continuing increase in our behavioral, morphological, neurological, and genetic freedom can be seen in examples ranging from the superficial to the profound. Our choice of clothing can be used both to express something about ourselves (“I’m an efficient executive,” “I’m a Grateful Dead fan”) and to help us attain an appropriate mood. Adopting a more formal attire for certain times of the day may help us to focus on the task at hand. Dressing in sporting gear may generate a mood favoring exercise. Advertising provides crude role models and may help us to feel like a certain kind of

person by using the product.

The partial personas we find in books and in film – more encompassing and profound than those in advertising – act as templates guiding us in sculpting ourselves into the self we want to become. Many males try, at least partially, to emulate tough characters portrayed by Arnold Schwarzenegger, Clint Eastwood, Harrison Ford, and Jean-Claude van Damme. Lately women, too, have had the choice of personas portrayed in Sigourney Weaver's Ripley, and Linda Hamilton's Sarah Connor. Many readers of this journal have, to varying degrees, sought to emulate qualities found in the characters of writers Ayn Rand and Robert Heinlein. By focusing on the paradigm personalities in these didactic stories, we can home in our desired self without having to deduce the requisite behavior from abstract rules. An image of our intended result is more effective at promoting change than is an abstract set of prescriptions. In times of intellectual opposition and isolation, for instance, recalling an image of Rand's Howard Roark will stiffen our resolve and independence more than advising oneself to "be independent!"

Changes of name have long been a method for redefining and committing oneself to specific values. Traditionally, woman who marry have given up their surname for that of their husband. This action was a way of supporting the belief that men were primary and women were subservient. Adoption of a new name may reflect rejection of one's native culture, religion, or nationality, and identification with new communities. Assumed names are common on the computer networks, the new names frequently differing wildly from standard names. A number of Extropians have adopted new names, to project what *they value*, rather than retaining a label connecting them to an unchosen background.

The computer networks are now a major locus of self-definition, at least with respect to the ability to choose how we appear to others. Apart from use of assumed names on the Net, some like to present a virtual image different from their physical image. Some women use a normally male name and carefully maintain this pretense in order to see how differently they will be treated. Since what you can know about someone regardless of their wishes across the Net is severely limited by the medium, it is easy to present personal characteristics selectively, or even misleadingly. Those who want to interact with others free of sexist, racist or nationalist prejudice may withhold or conceal these facts about themselves. The precocious but acne-ridden

adolescent can contribute to discussions without fear that others will discount his or her contributions due to age or appearance.

What will happen as networks speeds increase, allowing affordable video and voice transmission? Will this force virtual images to conform to physical images? At first perhaps, but further computational advances will allow real-time modification or synthesis of our virtual appearance. An amusing and plausible example of this is found in Bruce Sterling's SF novel *Schismatrix*, where video images are processed to remove stubble and skin blemishes, and to enhance attractiveness in the absence of make-up. Morphing techniques and digital image processing are already moving us in this direction. Just as netters now use adopted names, many in the future will create synthetic faces and voices as vehicles of expression. Their bodies will largely cease to constrain their mode of expression; they will be able to choose a form mildly or drastically different from their actual form. Further in the future, if we upload ourselves and exist primarily in the computational world (downloading ourselves into a range of bodies as it suits us), the range of possible forms and their ease of adoption will become practically unlimited. The distinction between virtual image or identity and actual image will increasingly weaken over the coming years, and will dissolve entirely if we upload. Our synthetic images will have become our actual images.

Attention to posture and bodily motion is another path to achieving a desired psychological state. More physically drastic modifications, such as cosmetic surgery and implants, are becoming increasingly common. As costs fall and expertise climbs, more people will choose more radical surgical and physical modifications, especially once nanomedicine has superseded crude surgical alteration. We may yet see physiognomic choices as bizarre as those of the Urban Surgery youth group portrayed by SF writer Walter Jon Williams.

In the 21st Century, the depth and significance of self-transformation and augmentation will far exceed our current experience. Within a decade biologists will have decoded the human genetic program, and we will then accelerate our ability to understand and correct genetically-related physical (and psychological and intellectual) deficits, and to enhance normal abilities to transhuman levels. Today's gene therapy is a magnificent achievement, but will seem minor once nanomedicine is able to alter any of the DNA of a developed, adult human. Each

of us can then choose to alter mildly or massively our physical constitution. We can boost our immune systems, alter our facial features, become taller or shorter, stronger or more delicate, and sharpen our senses.

These capabilities will leave us still human, merely giving us a choice of the peaks of humanity. But genetic changes could be radical enough to make the appellation 'human' inaccurate. If some people's genetic coding is different enough from that of humans, they will be a distinct species. If the changes are *positive* changes, these new peaks in the evolutionary landscape will be *transhuman*. Genetic enhancement will be used alongside neurochemical modification, and other cognitive enhancements such as neural-computer integration (as described below). We who prize moving forward, thrusting past old limits, and seeking new abilities, will no longer be confined by our genetic, biological, and neurological heritage. We will ignore the biological fundamentalists who will invoke "God's plan," or "the natural order of things," in an effort to imprison us at the human level. We will move through the transhuman stage into posthumanity, where our physical and intellectual capacities will exceed a human's as a human's capacities exceed an ape's. To fully flower, self-transformation requires a rebellion against humanity. As Nietzsche put it:

I teach you the superman. Man is something that should be overcome. What have you done to overcome him?

[*Thus spake Zarathustra*, I prologue, p.3]

Emerging and future technologies of transformation will be resisted by some religious and humanist groups precisely because of their powerful abilities to change our human constitution. Some opposition will result from fears about purported dangers of these technologies. Granted, caution is warranted, but such fears typically are vastly overinflated – as the hysteria over genetic engineering research illustrated. Opposition will also be motivated by a vaguer notion that humans should not 'interfere' with 'God's plan,' or 'the natural order.' This notion should only amuse us who believe in no god and who understand that transformation of self and environment is perfectly natural for humans.

## Ideal self/Optimal Persona

In pursuing self-transformation we face a plethora of options of both ends and means. Questions about ends may include: Which skills shall I concentrate on

developing? Do I want to become more or less judgmental? More or less giving to others? What kind of occupation do I desire? Is physical cultivation important to me? With regard to means, we may ask: Should I stay in/return to formal education longer? Should I work for others or work for myself? Should I use the current generation of smart drugs? Should I sign up for cryonic suspension? Proceeding effectively in our self-sculpting will require a clearly defined goal. Knowing our destination will allow us to set priorities, and to choose among methods and means, picking those we find most desirable and believe most effective.

---

**We will ignore the biological fundamentalists who will invoke "God's plan," or "the natural order of things," in an effort to imprison us at the human level. We will move through the transhuman stage into posthumanity, where our physical and intellectual capacities will exceed a human's as a human's capacities exceed an ape's.**

---

Setting a goal for self-transformation is best implemented by creating for ourselves a paradigm, an idealized model of the person we want to become. Comparing our present condition to our paradigm will allow us to steer a course through distractions and temptations more effectively than trying to reason our way along solely by using abstract rules, principles, and guidelines. Cybernetic control systems work on this principle. They have some map or representation of their destination, and continually compare their present state or location to the map, then make adjustments to keep on track. Our paradigm – which I will call the *ideal self*, or the *Optimal Persona*<sup>4</sup> – differs from the map of many cybernetic systems in that it is dynamic, not static. We decide who we want to become, and there is nothing to prevent us from changing our minds (though we should realize that constancy and tenacity are generally more produc-

tive than frequent changes of direction). Our ideal self should evolve as we revise the ranking of our values, as we see come to recognize new goals as worthwhile, and as we learn new behaviors that contribute more effectively to our ideal. The Optimal Persona is Nietzsche's *Übermensch*, the higher being existing within us as potential waiting to be actualized.

Constructing and periodically revising an Optimal Persona requires a high degree of self-awareness, an understanding of what we are motivated to do and what are the causes of these motivations. We need to have an idea of what we can reasonably expect to change in any given timeframe. Before setting out to change some of our personality characteristics so as to cohere with and support other characteristics, we should think critically about which parts of our current selves we have freely chosen and which we adopted unconsciously, absorbing them from the familial and cultural environment. For instance, someone might have been raised to be unfailingly polite, never speaking out directly against something they regard as mistaken or despicable. This person may have adopted this pattern of behavior, feeling it to be a part of them, yet critical reflection may lead them to decide that this behavior frustrates their more considered values and goals – those that are more truly personal because they were formed consciously in the light of their broader view of themselves and of the world. Beliefs, values, or behaviors adopted largely unreflectively (i.e., most of those personal qualities acquired early in life) may nevertheless form part of one's ideal self, but only after examining and reaffirming their worth and their coherence with other desired qualities. [See sidebar for a practical exercise based on the ideal self.]

Danger lurks in any desire to be different from our current self. Holding a conception of an Optimal Persona, and the critical self-examination necessary to actualize it, introduces the possibility of denying the worth of one's current self. Self-disgust and self-denigration, whether in recurring but transient episodes or as an enduring characteristic, is common among persons of high standards. Such people often treat themselves far more harshly for their shortcomings than they would another person. On making an error, on discovering their ignorance of something, or on failing their own standards, these individuals curse themselves, insult themselves, and in extreme cases may physically punish themselves. To avoid this, we need to be aware of the difference between dissatisfaction on the

one hand, and disgust, anger, and hatred on the other. You can respect and esteem yourself for what you can do, for what you get right, and for what you achieve, while simultaneously being dissatisfied with yourself. Dissatisfaction means you believe you can (and perhaps should) do better. Self-disgust means that you believe you *must* do better, that you are worthless unless you are perfect. Even if perfection is measured by your own standards, this kind of perfectionism is both painful and self-defeating.

Hatred of self brings depression and paralysis. A better response to mistakes, slips, and backsliding is to praise yourself for your current and past achievements and successes, while acknowledging the faulty behavior and focusing on ways to prevent a recurrence and to minimize the negative consequences. Similar remarks apply to the temptation to repress feelings that you believe to be inconsistent with your ideal.<sup>5</sup> One might, for example, feel fear when confronted with some threat or uncertainty. Repressing the emotion, rather than experiencing it and acting appropriately in response, will bury important information about yourself and your situation. A better response is to acknowledge the unwelcome feeling of fear, anxiety, anger, or weakness, and to investigate the possible ways of changing yourself or your situation so that these emotional responses will not occur. In the case of fear, this might mean working out or arming yourself for more physical confidence, or learning the skills needed to cope with a difficult situation. Or the best strategy might be to learn to live with the unpleasant emotion, reducing its severity by critically challenging the basis of the feeling, such as by thinking of a frustration as an inconvenience rather than as a disaster.

## **Continuity and self-direction**

Before describing some possible cognitive, physical, and psychological transformations possible today or in the future, I wish to stress the importance of the *self* in "self-transformation." Discussions of actual and hypothetical instances of radical transformation usually provoke the question, "But is the person after the change really the same person as the person before the change?" This is too complex an issue to explore adequately here.<sup>6</sup> I will limit myself to claiming that an important consideration determining whether a person undergoing dramatic change remains logically the same is the extent to which that person selects and directs the changes.

The sense in which the pre- and

# TECHNOLOGIES OF TRANSFORMATION

- **Nootropics (smart drugs):** Drugs that enhance retention, recall, and concentration.
- **Electronic networks:** Online databases, virtual fora discussions on numerous topics.
- **Hypertext:** Massively interconnected database; references can be tracked in both directions.
- **Genetic engineering**
- **Neurochemical modification:** Personal control of emotional and cognitive states.
- **Nanocomputer implants:** Molecular computer integrated with the brain, providing additional memory, processing power, and running decision-making programs.
- **Internal nanotech:** Nanoengineered immune system, strengthened bones, reinforced skull, backup organs, muscular augmentation.

post-transformation individuals are the same is the logical, not the qualitative sense. Clearly, by hypothesis, they are significantly different qualitatively. They are the same – they are logically identical – if they can reasonably be considered as two temporal stages of one persisting entity. If a person suffers a massive brain injury, causing loss of all rational capacities, major changes in emotional response, an inability to recognize close friends, and erasure of memory, the psychological connections between the earlier and the later individual are too tenuous for them to count as the same person (though we can say the same *body* persists). Loss of continuity need not (at least in principle) require a *loss* to occur. If the entire psychology of someone changes instantaneously and these are changes for the better (even according to their pre-change standards), personal continuity will have been destroyed, leaving behind a new person. Obviously, a spectrum exists between cases of total discontinuity and total absence of change; personal continuity may be disrupted to varying degrees.

There are two reasons why self-direction of one's transformation is important in maintaining continuity. The more obvious reason is that another person is less likely to make the changes in you that you would choose, either because they don't know what those are, or because the modifications they choose to make will be influenced by their own interests. The second reason is that continuity requires that later stages of an individual *develop* out of earlier stages, rather than simply usurping their place. Thus replacing 90% of a person's brain with neural matter of a different configuration would not preserve continuity, because in no way is the resulting person a development out of the earlier. Changing ourselves is more likely to result in continuous development rather than disruption of self since the outcome will better reflect our values and goals. In choosing which changes to make and when to make them, we will be better able to integrate the new or modified characteristics into our overall character. Control over our own transformations will grow in importance as more powerful

technologies (some described below) are introduced, e.g., genetic engineering, neurochemical modulation of mood and cognition, neural-computer integration.

## Cognitive self-transformation

This category of self-transformation encompasses intellectual virtues that foster personal growth, non-technological methods of enhancing intelligence and rationality, and technologies capable of augmenting our intellectual powers to a superhuman level.

The intellectual virtues are those enduring qualities of character that reveal themselves in our methods and habits of thinking. Rationality – the unlimited application of critical thinking – should be regarded as the primary intellectual virtue. Rational thinking means not believing assertions casually; it requires a habit of asking questions such as: "What is your evidence for that?" "According to whom?" Rationality means questioning, examining, assessing your own beliefs for their coherence and grounding, and an avoidance of belief in a proposition simply because it is easy or comforting.

Rationality does not allow room for accepting ideas on faith. 'Faith,' in the sense used here, means believing in something in the absence of or contrary to the evidence; I do not use the term to include trust in what someone says where that trust is justified on the basis of past experience. Faith in a method of personal transformation leads to stagnation, since the continuing failure of the method to produce results will be ignored or rationalized. Critical rationality play an essential role in effectively assessing competing means to our goals. For instance, we should be open to evidence showing the ineffectiveness of current nootropics (smart drugs), so as not to waste our resources and to free us to pursue other methods.

Determining where to draw the line between persistence and faith can be difficult. Some methods of self-improvement may only take effect after considerable and repeated effort; abandoning them too soon will as surely lead to stultification as will clinging to failed methods. Knowing when to abandon one avenue of exploration in favor of a fresh one will partly depend on knowledge about our own propensities either to give up prematurely or to persist irrationally. Critical rationality therefore requires a balance; it is not simply a matter of being constantly and supremely critical. A propensity constantly to criticize all attempts at self-improvement may reflect, not a reasonable caution, but an evasion of the respon-

sibility to choose a method and implement it.

Achieving this balance between persistence and critical analysis requires an ability to tolerate uncertainty and ambiguity. Authoritarian personalities, like authoritarian governments, cannot bear disagreement, uncertainty, or alternatives. They demand allegiance to a single goal, a single method, a single agency. Living extropically calls on us to develop and sustain the contrary ability to welcome alternatives, to encourage diversity of opinion, and to thrive on uncertainty. Only the ability to remain open to new information and evidence will be effective in our pursuit of self-transformation. We will heed R.A. Wilson's dictum: "Convictions make convicts," and will think in terms of working hypotheses rather than certain beliefs, and use the probabilistic categories of "fuzzy logic" rather than the black and white knife of classical logic. We can learn to enjoy the progress represented by being corrected more than the comfort of feeling certain. Not only will this be more personally effective, thriving on uncertainty and correction will reduce interpersonal conflict, allowing us to accept the merits of another's argument without needing to reject their entire argument or to attack the argument's proponent.

Cognitive self-transformation requires us to search for and employ the most effective methods of increasing our intelligence. Effective reasoning is possible only if we study the process of reasoning itself, in order to allow for weaknesses in typical ways of thinking, and to seek ways of augmenting our analytical and creative capacities. Meta-reasoning (reasoning about reasoning) will obviously include studying the fundamentals of logic, statistics, and some areas of mathematics. It may include the newer field of fuzzy logic, as implemented in the new generation of electronic goods to handle continuously varying quantities. We can further sharpen our reasoning in regard to decision-making by studying the fields of game theory and strategy, applying iterated Prisoners' Dilemma reasoning (as illustrated in Axelrod's enlightening work), and more controversially, Hofstadter-style superrationality. These fields, added to an understanding of human psychology, will increase our effectiveness in personal interactions, enhancing our ability to achieve our goals while leaving others feeling satisfied rather than frustrated.

Cognitive psychologists have demonstrated biases in human reasoning about the probability of an uncertain event, or the value of an uncertain quantity. They

have shown that we "rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors."<sup>7</sup> An example is the representativeness heuristic, in which probabilities are estimated according to the degree to which A is representative of B (the degree to which A resembles B). For example, if given a description of an individual who is shy, helpful, uninterested in the real world, but tidy and possessing a passion for organization, most people will guess out of a list of jobs that he is a librarian. This estimation will usually be made in disregard of prior probabilities; even if there are far more farmers than librarians, for instance, the individual will be assumed to be a librarian because of the closeness of his description to a stereotype of librarians. Avoiding the inappropriate use of heuristics such as this is tremendously difficult. Measuring probabilities by closeness to a stereotype appears to be a natural result of the functioning of our brains. Our neural networks form paradigms or exemplars, to which we compare incoming information.<sup>8</sup> The more we familiarize ourselves with these biases, the more frequently we will notice and correct them.

Other typical heuristics and biases include the availability heuristic, in which we overestimate the probability of an event because we easily recall a similar event. So, I might forego buying a Dell computer because I remember the problems my friend had with a Dell, even if there is no reason to believe this to be statistically significant. The Gambler's Fallacy, involving fallacious assumptions about probability and causation, is well known. Other errors arise from mistaken beliefs regarding statistical regression, from anchoring our estimates to previous estimates (a type of cognitive conformism), and from conflation of correlation and causation.

Many other methods of cognitive augmentation are available, even without employing technology. General semantics<sup>9</sup> warns of intellectual traps such as wholly identifying one thing with another ("John is a Republican"), and even offers a revision to English—E-prime<sup>10</sup>—which attempts to avoid use of forms of the verb *to be*. Artificial languages such as Loglan/Lojban might repay our study by providing a linguistic medium specifically designed to prevent unintentional ambiguity. Philosophy of science can improve our understanding of experimental procedure and scientific warrant<sup>11</sup>.

Speedreading techniques enable us to boost the efficiency of our information gathering, while numerous memory techniques allow us to retain more of the information acquired. These are just some of the non-technological means available for our project of cognitive enhancement, each of which deserves an article, at least, in itself.

Present and future technologies will further expand our intellectual capacities, in conjunction with the foregoing means. The current generation of nootropics (smart drugs) appear to be mildly effective for relatively young, healthy persons, but improved understanding of neurochemistry, synthesis of more powerful compounds, and more precise delivery mechanisms, should allow us to push back our biological and neurological limitations. Our capacities for organizing and presenting information are vastly expanded by use of personal computers, and the Net provides a practically endless source of documents, discussions, and expertise. The appearance, in 1993, of the first generation of personal digital assistants (PDAs) heralds an era of increasingly portable personal computing power and communications flexibility. Soon you will be able to contact most people, and access remote databases, no matter where you happen to be. Software agents and 'knowbots' will help us to gather the information that interests us, relieving us of tedious work hunting down and managing information. Recent experiments linking a biological neuron to a field effect transistor point to the day when our computerized assistants will be inside our heads. Eventually our computers will be tightly integrated with our brains, becoming part of us, and abolishing barriers to the attainment of transhuman intelligence.<sup>12</sup> We may also genetically engineer our brains to expand their capacities, and even upload our consciousness to superior hardware, thereby endowing ourselves with the unlimited potential of posthuman intelligence.<sup>13</sup>

## Physical self-transformation

Superior cognitive performance will not persist for long if our bodies are deteriorating, aging, and dying. Elevating personal extropy will therefore include physical self-transformation. As with cognitive enhancement, many physical improvements can be made without employing current or future technologies. Living extropically will involve a concern for maximizing our health through diet and exercise, from widespread practices such as high-fiber, low-fat foods, weight-training and aerobic exercise, to well-es-

tablished but lesser known practices such as the very low calorie, very high nutrition Walford (High/Low) Diet<sup>14</sup>, which has consistently reduced the incidence of many diseases and extended both mean and maximum lifespan in widely varying species.

Even given an extropian commitment to physical transformation, we face conflicting choices. Physical transformation refers to a collection of goals, including health, longevity, strength, resilience, speed, stamina, suppleness, and beauty. Some of these goals may be, to a degree, mutually inconsistent. For instance, the Walford Diet is tremendously effective at promoting health and longevity, but will preclude extensive muscle-building. If we choose primarily to pursue the peaks of performance, whether strength, stamina speed, or suppleness, we will likely have to sacrifice some health, longevity, or possible beauty (depending on your standards). Aerobic exercise exceeding about 30 minutes, three to four times weekly, will increase stamina but produce no further cardiovascular protection, while producing more free radical activity and injuries, and (temporarily) suppressing immune function. Injecting anabolic steroids will reliably increase muscle mass and strength, but sustained use brings several deleterious health effects. So, while all Extropians ought to challenge themselves with exercise and a careful diet, the particular mix of performance vs. health and longevity will be a personal choice.

The decades that lie ahead will bring technologies of transformation enabling us to modify, augment, and replace our human, biological bodies with superior vehicles worthy of our evolving intelligence. Increasingly, those of us desiring bodies beyond those evolved by natural processes, will engage in a process of what I call *transbiomorphosis* – the engineering of improved bodies by intervening in biological processes, and by incrementally replacing our biological forms with synthetic life-sustaining bodies. As Nietzsche realized in the late Nineteenth century, humanity is not the end of the story of evolution:

...Man is a rope, fastened between animal and superman - a rope over an abyss...what is great in man is that he is a bridge and not a goal.

[Friedrich Nietzsche, *Thus Spake Zarathustra*, Pt.1, p.3]

Already we can enhance our health and longevity with a multitude of nutritional supplements and drugs. The 1990s have seen the beginning of gene therapy; we can expect genetic engineering to progress from restoring defective systems (today's

## VISUALIZE YOUR OPTIMAL PERSONA

Reprogram your behaviors and habits by regularly practicing this exercise.

(1) Develop in detail an image of your ideal self. Write down your desired characteristics. Then select two or three behaviors to change or goals to achieve.

(2) Sit up straight. Breathe in deeply, hold the breath while tensing all your muscles. After a few seconds, blow out the breath while releasing the muscular tension. Repeat. Then start and maintain regular deep breathing (try breathing in for two seconds, holding it for eight seconds, and breathing out for four seconds).

(3) Bring into focus a picture of your Optimal Persona, seeing yourself behaving as you want, achieving your goals. Practice until you can use all sensory modalities: Make your internal image clear and moving, and hear what is being said, even imagining the smells and sensations of the situation.

Continue for 15-20 minutes. Use this exercise at least once a day, focusing on just a few characteristics in each session.

For more details on a variety of visualization techniques, see Tony Robbins, *Unlimited Power*, Adelaide Bry, *Visualization: Directing the Movies of Your Mind*, Jose Silva, *The Silva Mind Control Method*.

medical paradigm) to pushing back natural limits (tomorrow's medical paradigm). Much discussed artificial organs will be a temporary measure, merely a stand-in until the arrival of nanotechnological medicine which, without cutting or poisoning will cure disease, regenerate limbs, reverse aging, and will allow us to reinforce our bones, massively strengthen our immune systems, and re-engineer our bodily structure as we please.<sup>15</sup> Apart from structural enhancement, we can anticipate unprecedented control over our appearance, including the possibility of complete and reversible change of gender.

Transbiomorphosis will involve the merging of our machines and technolo-

gies with the human body. Earlier, I said that computers will continue to shrink, while their power grows, and the degree of interconnectivity with our brains increases until they become part of our brains. We can also expect our senses to be sharpened and new senses to be added through interfaces with mechanical sensors.

Most people feel alarmed or horrified by the prospect of human-machine integration or merging. They fear, understandably, a loss of humanity in becoming "mechanized." This fear is fed by popular images, whether it is the collectivist monstrosity of the Borg in *Star Trek: The Next Generation*, the programmed behavior of *Robocop*, or the rigidly robotic

imagery in the pioneering techno music of *Kraftwerk*. Mechanization of this kind is indeed to be shunned, for it extracts the vitality of life, simplifying thought and behavior, subjecting the agent to programming and external control. This is the very antithesis of the Extropian drive for self-actualization, personal growth, and individual freedom. These undesirable and misleading connotations of the term 'mechanization' explain why I prefer to talk of transbiomorphosis. Our future integration with the products of technology will not be a mechanizing, constraining, subtracting process, but the very opposite.

Recent years have revealed a clear trend toward making our artifacts organic (in the abstract sense), fluid, responsive, and living. While rigidly programmed computers will probably always be used for some purposes (due to their blinding speed at logical computations), we are witnessing the emergence of connectionist machines – neural networks that learn from experience, adapt, and solve problems without a human determining the algorithms. Such artificial neural networks are modeled (more or less abstractly) on brain function. The new field of artificial life (A-Life)<sup>16</sup> attempts to evolve computational and robot organisms that share the characteristics of our familiar carbon-based life. Software designers are developing knowbots and other software agents that respond to our needs and desires, learning from us and helping us. Fuzzy logic is being implemented in electronic devices, obviating the need for rigid on-off responses. These and other examples illustrate ways in which some technologies are diverging from traditional rigid machine behavior, and evolving towards an organic, flexible, complex functions suitable for supplementing our limited brains.

## Psychological transformation

In this section I will comment on self-transformation as applied to personality and behavior. This will include what would normally be called morality – questions of what behaviors and dispositions are *good* and *bad*. However, standard notions of good and bad are deeply stained by the religious metaphysics that I have already rejected. In place of moral concepts we might evaluate our psychological characteristics as healthy and unhealthy or sick. As Norton says, expressing Nietzsche's view: "Health denotes all that contributes to ascendent vitality, while sickness characterizes whatever contributes to life's degeneration and demise..." (Norton, p.82.)

If we are to actualize our ideal selves, we must first choose that self. The self we encounter when we first look within may not be a self we have chosen. Before we can realize ourselves, we need to discover and choose ourselves. This requires a thorough, unrelenting self-examination in order to uncover the sources of our current psychology. We will expose the contribution of our family, our teachers, and our culture to our development, and will see that we absorbed many of those influences largely unconsciously and uncritically. Choosing an ideal self asks of us that we "revalue all values," that we look at our person as a fascinating stranger, and determine whether we wish to affirm, modify, or relinquish each of our important beliefs, habits, associations, relationships, and dispositions. Since self-transformation is a dynamic process and the ideal self an ever-evolving paradigm, the initial period of revaluation must be followed by recurrent self-examination and course correction. A range of psychological techniques can facilitate self-understanding (such as the sentence-completion exercise frequently employed by Nathaniel Branden<sup>17</sup>), as can meditation techniques. Some internal explorers report breakthroughs facilitated by drugs such as the empathetic MDMA ("Ecstasy") and the psychedelic LSD.

Once a self has been affirmed and an ideal self or Optimal Persona created, we need to take responsibility for our own lives, and we will demand the right of self-determination. In our quest for self-realization and transformation we will want to experiment with alternative methods of growth, including those discussed earlier. This requires that others not interfere with our free and responsible choices. Unfortunately, the government of the U.S.A., in common with every other state in the world, arrogates to itself the power to circumscribe our experimentation. In our pursuit of health, longevity, and cognitive enhancement, our greatest enemy in the U.S.A. is the Food and Drug Administration, with its monopolistic approval process.<sup>18</sup> The barriers to self-development raised by agencies such as this show that political awareness and action should form part of our plans for transformation. Only in a truly free community can we fully realize our potential.

External obstacles should never be used as an excuse for failing to explore our potentials in the many ways left to us. We can try out new careers and projects, developing new skills and aspects of ourselves; we can mix with different types of people; experiment with new types of relationships; and visit new locations, learning about diverse cultures. A commit-

ment to experimentation, flexibility, and personal evolution will protect us against our own dogmatism, stagnation, and the thoughtless comfort of conformity.

A core feature of successful psychological transformation is self-discipline. Without the ability to control our impulses and to maintain our carefully planned course, we will fritter away our energies in every direction. Lack of self-control will leave us vulnerable to those who would use us as tools for their own ends: "He who cannot obey himself is commanded." Self-discipline and the conscious self-guidance of our lives will allow us to achieve ever higher goals, as we raise our sights with each triumph. Effective self-rule will free us of the desire to control others. Contrary to popular interpretation, the *Übermensch* are not the Blond Beast, the conqueror and plunderer. They are those who neither rule others nor tolerate others' attempts to rule them. The developed, self-chosen self will exude benevolence, emanating its excess of health and self-confidence. As a well-crafted and integrated individual, the self-transforming person will have the strength to be honest and sincere, to reveal and express him/herself. One who has long practiced self-transformation will present an appearance of depth, stability, discipline, and of being at ease.<sup>19</sup>

Changing aspects of our personality and determining our mental state can be tremendously difficult. Joining the non-technological tools already at our disposal soon will be powerful means of cognitive and emotional modulation. Modification of our DNA and resulting brain structure may be able to alter the ancient evolved drives over which we currently have minimal control. For instance, we may be able to reliably control our drives for sex, for territory, and for violence. If we come to understand the relation between our brain structure and endemic desires for intellectual comfort and certainty, we might be able to modify ourselves (with a cautious eye on the consequences) to reduce our need to be proven 'correct,' and raise our tolerance for seriously considering alternative interpretations of the world.

Since our cognition and our emotions are deeply interwoven, future abilities to edit genes, modify the hormonal output of the neuroendocrine system, and to affect the levels of the numerous neurotransmitters, should grant us far greater choice of how we typically think and feel. We may develop chemical-releasing implants, controlled by a computer interfaced with our brains, that allow us to rapidly alter our state of mind, for instance to dramatically increase alertness,

or to disconnect sexual impulses when they are distracting, or to gear us up for a major intellectual challenge.<sup>20</sup> These possibilities may alarm some people, but they are merely extensions of everyday, much cruder methods of neurological and emotional control, such as the use of exercise, sex, food, drugs, and television.

I have attempted to demonstrate the centrality of self-transformation to an extropic life, and to explore several aspects of and means toward self-transformation. We face an open-ended future looming large with potential for defining and transforming ourselves to an extent unthinkable in all past human history. As is to be expected in regard to this topic, I have focused on selves as individuals, since we each have to take charge of our own destiny, and accept responsibility for who we are and who we can become. This stress on the individual should not be taken to denigrate the extensive contribution to self-development afforded by suitable groups and cultures. We need not be isolated, totally self-sustaining achievers. Support and encouragement by fellow extropic-minded persons is enormously valuable. Extropic friendships, cultural groups, and activities provide a stimulus for us to move onward, upward, outward. Let us encourage each other by setting examples of what can be achieved, let us share our discoveries, and accelerate ourselves toward the attainment of individual and cultural excellence.

## NOTES

<sup>1</sup>I have developed other Extropic principles in previous issues of *Extropy*: Spontaneous Order in "Order Without Orderers," *Extropy* #7 (vol.3, no.1), and "Dynamic Optimism," *Extropy* #8 (vol.3, no.2), and summarized all five principles in "The Extropic Principles 2.0," *Extropy* #9 (vol.4, no.1).

<sup>2</sup>A survey of the increasingly optimistic views of professional gerontologists appeared in *Life*, October 1992.

<sup>3</sup>See "Transhumanism: A Futurist Philosophy." *Extropy* #6 (Summer 1990).

<sup>4</sup>I borrow the term "Optimal Persona" from Bruce Sterling's excellent near-future novel, *Islands in the Net*, though in *Islands* the Optimal Persona is a common hallucination rather than a consciously constructed model.

<sup>5</sup>A good discussion of this can be found in books by Nathaniel Branden, especially *The Disowned Self*, and *Honoring the Self*.

<sup>6</sup>I analyze this issue in detail in Chapter 6 of my Ph.D. dissertation (in progress), *The Diachronic Self: Identity, Continuity, Transformation*. Contact

me at the Extropy Institute address if you would like a copy.

<sup>7</sup>Amos Tversky and Daniel Kahneman, p.3 of Kahneman, Slovic, Tversky, 1982. See also, Nisbett and Ross, 1980.

<sup>8</sup>See Paul M. Churchland, 1989.

<sup>9</sup>The classic reference is Alfred Korzybski, *Science and Sanity*, 1933.)

<sup>10</sup>On E-Prime, see D. David Bourland, Jr. and Paul Dennisthorpe Johnstone, eds., 1991, and the Summer 1992 issue of *ETC. A Review of General Semantics*.

<sup>11</sup>A good place to start would be Klemke, Hollinger, Kline, eds. 1980, and Lambert and Brittan, Jr., 1970, 1979.

<sup>12</sup>See E.A. Wan, et al, 190, and J.D. Foley, October 1987.

<sup>13</sup>Hans Moravec, 1988.

<sup>14</sup>Roy L. Walford, *The 120 Year Diet*, 1986

<sup>15</sup>Nanotechnological medicine is described in Drexler, 1986, and Drexler, Peterson, Pergamit, 1991.

<sup>16</sup>For an introduction to A-Life, see Simon! D. Levy's "Neurocomputing 5: Artificial Life," in *Extropy* #8 (vol.3, no.2), Winter 1991-92.

<sup>17</sup>His first book illustrating this often startling technique is *The Disowned Self*.

<sup>18</sup>I critique the deadly policies of the FDA in my talk, "Recreational Drugs and Smart Drugs: Paternalism and Responsibility," available on audio tape.

<sup>19</sup>An effective portrait of the developed, self-transformed person is presented in Wayne Dyer's books, *Your Erroneous Zones*, *Pulling Your Own Strings*, and especially *The Sky's the Limit*. Unfortunately I cannot recommend later books by Dyer, who appears to have abandoned useful insights for New Age vagueness and platitudes.

<sup>20</sup>A recent SF novel by Greg Egan, *Quarantine*, contains the best portrayal to date of these possibilities and their effects on our self-conception and sense of identity.

## BIBLIOGRAPHY

Robert Axelrod, *The Evolution of Cooperation*. (Basic Books, New York, 1984.)

D. David Bourland, Jr. and Paul Dennisthorpe Johnstone, eds., *To Be or Not: An E-Prime Anthology*. (International Society for General Semantics, 1991.)

Nathaniel Branden, *The Disowned Self*. (Nash Publishing Corporation 1972, Bantam Books, New York, 1973.)

Nathaniel Branden, *Honoring the Self: The Psychology of Confidence and Respect*. (J.P. Tarcher 1983, Bantam Books, New York, 1985.)

Paul M. Churchland, *A Neurocomputational Per-*

*spective: The Nature of Mind and the Structure of Science*. (Bradford Books, MIT Press, 1989.)

Eric Drexler, *Engines of Creation*. (Anchor Press/Doubleday, 1986.)

Eric Drexler and Chris Peterson with Gayle Pergamit, *Unbounding the Future: The Nanotechnology Revolution*. (William Morrow and Company, Inc., 1991.)

Greg Egan, *Quarantine*. (Legend Books, London, 1992.)

ETC. *A Review of General Semantics*, Summer 1992 (vol.49, no.2).

J.D. Foley, "Interfaces for Advanced Computing," *Scientific American*, October 1987: 127-135.

Daniel Kahneman, Paul Slovic, and Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*. (Cambridge University Press, 1982.)

E.D. Klemke, Robert Hollinger, A. David Kline, eds., *Introductory Readings in the Philosophy of Science*. (Prometheus Books, 1980.)

Alfred Korzybski, *Science and Sanity*. (International Non-Aristotelian Library, 1933.)

Bart Kosko, *Fuzzy Logic and Neural Networks...* (or forthcoming popular book).

Karel Lambert & Gordon G. Brittan, Jr., *An Introduction to the Philosophy of Science*. (Ridgeview Publishing Company, 2nd edition 1979.)

Simon! D. Levy, "Neurocomputing 5: Artificial Life." *Extropy* #8 (vol.3, no.2), Winter 1991-92.

Hans Moravec, *Mind Children: The Future of Robot and Human Intelligence*. (Harvard University Press, 1988.)

Friedrich Nietzsche, *Thus Spake Zarathustra*. Parts I and II published in 1883, Part III published in 1884, Part IV written in 1885, published in 1892.

Friedrich Nietzsche, *The Gay Science*. (1882, 2nd expanded edition 1887.)

Friedrich Nietzsche, *Assorted Opinions and Maxims*. (Published in 1879 as the First Supplement to *Human, All Too Human*; 2nd edition, 1886.)

Richard E. Nisbett & Lee Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*. (Prentice-Hall, 1980.)

David L. Norton, *Personal Destinies: A Philosophy of Ethical Individualism*. (Princeton University Press, 1976.)

Bruce Sterling, *Schismatrix*. (Ace Books, New York, 1985.)

Bruce Sterling, *Islands in the Net*. (Ace Books, New York, 1988.)

Roy L. Walford, M.D., *The 120 Year Diet: How to Double Your Vital Years*. (Pocket Books, Simon and Schuster, 1986.)

E.A. Wan, G.T.A. Kovacs, J.M. Rosen, and B. Widrow, "Development of neural network interface for direct control of neuroprostheses." *Proc. Second Int'l Conf Neural Networks* (Washington, D.C., Jan 15-19 1990) II-3-21.

Walter Jon Williams, *HardWired*. (Tom Doherty Associates, Inc, 1986.)

# A conversation with Mark S. Miller

by David Krieger

## Part one: Creole physics and the credit theory of identity

Mark S. Miller is one of the system architects of the Xanadu project, the electronic hypertext system conceived by Ted Nelson as the future of publishing. Miller is also author, with K. Eric Drexler, of the Agorics papers (published in *The Ecology of Computation*, B.A. Huberman, ed.; New York: Elsevier-North Holland, 1988) which first presented the idea of agoric open systems — computer operating systems in which system resources such as memory and processing cycles are traded by programs and processes on an internal open market. I spoke with Mark in Palo Alto, California in September 1992, and began by asking him about what he calls "creole physics":

In several different areas, what we find is that we seem to have, in the case of language, a particular grammar wired into our brains as kind of our initial grammar. When I say wired into our brains, it's not necessarily the case that you can look at the genotype and see the grammar directly encoded in there, but in the sense of Dawkins' extended phenotype<sup>1</sup>, that a particular grammar seems to be the phenotype of our genotype. The reason we have confidence in this, is that there is this process that there are at least three clear instances of, in which what linguists call a creolegrammar arose. This process is one in which the first generation of children who grow up speaking creole are in a linguistic environment which does not constrain their grammar and to a large degree doesn't teach them grammar.

You may want to go into the definitions of a pidgin and a creole. When a bunch of people speaking a bunch of different languages are somehow thrust together by weird historical circumstances — especially when they're trying to engage in trade with each other, which is a reason for different language speakers to constantly be in contact with each other — the language that they end up speaking is termed by linguists a pidgin language. The characteristics of a pidgin language are that the vocabulary is a compromise between the vocabularies of the different languages that fed into the pidgin, but the grammar is not. To a significant extent, to a good first approximation, a pidgin language has no grammar. It really doesn't do very much positional encoding, and you simply communicate by trying to put all of the content into the vocabulary words. However, if you listen to a given pidgin speaker, you can tell what his original language was, because he'll be using the grammatical constructs of his native language, he just will no longer be encoding meaning in that.

### Right.

The children of pidgin speakers are in the closest approximation to sort of the "ideal fantasy linguistic experiment."

### The "state of grace."

Yeah, the "state of grace fantasy linguistic experiment." The sort of thing that you would clearly love to be able to do as a linguist would be to somehow get a group of human infants who've never heard a human language and put them together on an island tended by robots or deaf-mutes or somebody — specifically, deaf-mutes who didn't know sign language, a completely non-linguistic environment — and let them just, in interacting with each other, spontaneously start speaking some kind of language, and just see what the structure of that is. Also, you'd want to do that experiment repeatedly with groups of children over and over again so you get some statistical distribution. Now, there's all sorts of reasons why, logically and ethically, and developmentally probably, this is not a possible experiment.

However, the children of the pidgin speakers are raised in what is probably the closest to this experiment you can ever get to, which is: They're raised in an environment in which their parents don't consider any grammatical construct to be correct or incorrect, so they're free to use whatever grammatical constructs they use, and it doesn't get corrected. Every time that there's been a clear case of this historical circumstance, the creole language that results — the language that the children of the pidgin-speakers speak — has the same grammar. And the most recent case, which is Hawaiian pidgin and Hawaiian creole<sup>2</sup> — they did this research while not only the original Hawaiian creole speakers were still alive but while the original Hawaiian pidgin speakers were still alive.

They were able to determine that the grammatical constructs of the Hawaiian creole speakers were not only the same

constructs as in other historical creole circumstances, but they were grammatical constructs that were found in *none* of the languages that were inputs to Hawaiian pidgin. It couldn't have been transmitted culturally, so you have definite non-transmission of a grammar that arises spontaneously every time this thing happens.

In addition, once they had this hypothesis in hand with this one verification, they said, "Okay, if this hypothesis is true, then what we should find when studying children growing up learning to speak English, is that where English and creole agree, the children should basically get the English constructs immediately, and where English and creole disagree, the children should start out speaking the creole construct, and have to have it corrected into the English construct. And they found exactly that.

The one particular case that I remember from the article — although the article has a lot of examples of this — is "double negative meaning emphasized negative." That's a creole construct. From a logical first-principle point of view, it's clearly more complicated than a "double negative meaning a negated negative," i.e. —

### A positive.

A positive, or at least the absence of a negative. So double negatives are something that children end up starting to use and they have to get it corrected into not using double negatives. As we all know in English-language culture at least it's only partially corrected.

In any case, that's one example of having this native inborn large structure that's an approximation of the culturally-evolved structures that we're used to, and it's the thing that starts the boot-strapping process off, where we come at life with this biologically-constructed —

*Set of pre-conceptions, actually.*

Yeah, and then we get it corrected into a more culturally-evolved and hopefully a more sophisticated one.

Other examples are "naive physics." With naive physics, what they've done is they've interviewed people, given them little tests, and said, "Intuitively, what do you think would happen if you took a ball that was going around a circular wall, and when it got to the end of the circular wall, which direction would it drop?" The correct Newtonian answer should be obvious —

*Tangent, basically. It continues going in a straight line.*

Right. But an amazing number of people will think that it either continues circular or will spiral outward. With "naive physics," what you find is that, even in physics students who have gone through first- and second-year Newtonian physics, naive physics is fairly unperturbed. If you ask the questions in the right way, where you still get at what their intuitions are, as opposed to making them feel like they're on a physics test and keying into their training, what you find is, their intuitions have been largely untouched by the training, and they still have the same intuitions as people who have not learned physics, and the intuitions actually resemble Aristotelian physics to a large degree. The significance is, the Greeks not being experimentalists, Aristotle cataloged not physics, but human psychology about physics, and I think it's interesting to think of that as a theory.

Popper is very clear on the analogy between biological evolution and theories, and in fact over here we have a very nice mixed case of a biological evolution of a theory of physics, which was "good enough". Biology stops, basically, when it gets to a "good enough" solution, especially when it's an adequate-enough solution that any creature is able to carry forward from there using cultural evolution.

*To an extent, biologically-derived theories are falsified through species or individual selection, in that, if your model of physics is "wrong enough" to expect the ball to do something completely—not even in line with "naive physics," for instance to move out of the plane of its motion, something that wrong—animals with that model will presumably be selected out, if it's wrong enough to make a survival difference.*

This is a complete aside, but I think it's interesting, and I also think it's original: A lot of people look at naive physics, and think just on the kind of stuff we've already said. "It's good enough, and there

are things about it that are wrong, but that's okay; culturally, we can correct." I think that it's actually better adapted as a physical theory in many ways, than Newtonian physics is, to the world that our ancestors were in. Here's specifically the thing that I mean by that: Everybody assumes that naive physics is primarily a physics of inanimate objects. However, I think that a lot of the ways in which it's wrong are actually by virtue of being more adapted to being a physics of animals.

*There's a certain animism in expecting the ball to continue on a curved path as if it has an intention.*

Exactly. There's also a certain animism in thinking that "It's been constrained, and now it's free, so now it's going to overcompensate." Either one of those are intuitions that are good at trying to predict the behavior of intentional creatures.

---

## **Popper is very clear on the analogy between biological evolution and theories, and in fact over here we have a very nice mixed case of a biological evolution of a theory of physics, which was "good enough."**

---

*I was going to say, you've "adopted the intentional stance" toward your particle or projectile.*

Exactly. Most of what we needed to interact with and predict successfully were animate objects, many of whom were trying to evade us, so it would be appropriate to devote more of our resources to trying to predict them than to predict these nice cooperative, or at least not hostile, inanimate objects.

*So to a certain extent, Aristotelian physics is still with us, and will continue to be with us until we start monkeying around with our genes.*

By the way, there's one very interesting example of this. This is the thing that really made me add the animism to this. The way I came by this animism thing is after reading a fascinating paper about some computer animation work.<sup>3</sup> An interesting peculiarity which has long been known to animators is that, when you draw an animation of a bouncing ball,

there are certain things that you can do that are violations of Newtonian physics — rather gross violations — that actually are known to make the animation sequence look more realistic. If you do it realistically, it actually looks less realistic than if you introduce these distortions. Now, what the distortions are, is that a ball, before it hits, elongates in the direction of motion, and then it hits and it flattens out. Also, there's another similar thing — let's just deal with that one, that's sufficient.

In this computer animation work, it was work in animating animals. In which what they did, well, they actually used a Luxo lamp as a "virtual animal", a Luxo lamp being this nice simple structure that's easy to animate, that you already have the database for —

*I've seen the cartoon; the one that I saw was called "Luxo, Jr."*

I'm not sure if "Luxo, Jr." actually used this technique; I think instead it's because the database from "Luxo, Jr." was already available, that they decided to do this work with the Luxo lamp.

What they did is, they wanted to make the Luxo lamp move around the way you would imagine an animal would move around. They did that by setting up an energy minimization problem. They sort of said, "Okay, the struts of the Luxo lamp are a skeleton, the joints are joints, and we will just, in our calculation, come up with a muscular model for bending the joint." Specifically, a model of the relationship of force and torque versus energy expenditure. Then what they did is they said, they basically had the following problem: the Luxo lamp is stationary on the ground over here, there is some kind of barrier between it and the goal, and it needs to get from where it is to the goal, and end up stationary at the goal, not just at the goal but stationary at the goal, and it needs to do that with a minimum expense of muscular energy. Fortunately, problems like this are sufficiently monotonic that you can solve them with straight hill-climbing.

What they ended up with was a motion in which the Luxo lamp first crouched down, then extended in a leap, then contracted back together, and moved its base forward. After it was over the barrier, it elongated itself —

*In preparation for landing.*

In preparation for landing, and then did a soft landing and came to a stationary position. So, they observed in their paper, that this has reproduced in a principled way, the trick that animators always knew about, and what I think they didn't go on

to explicitly say, although they may have, but I'll certainly go on to explicitly say, is that even at those very low levels of our visual apparatus, we have this very intentional physics, in terms of our expectations and perceptions of motion, and that's why motion of the inanimate object which is distorted towards animate motion, actually looks more realistic.

So, back from the digression: Kahneman, Slovic, and Tversky<sup>4</sup>, and also a book called *Human Inference* by Nisbett and Ross<sup>5</sup>, catalog a whole bunch of ways in which human statistical inference is very badly flawed, but distorted in systematic ways, and there are many many problems that result from the distortion, but, to some extent, society — we are all escaping from those problems by virtue of the modern understanding of statistics.

*Another topic I wanted to ask you about was your views on the question of identity, especially as it applies to cases like uploading and duplication.*

It seems to me that most everybody comes at the identity question with the presumption that there is some hard notion of identity that's real, and that all we need to do is find a definition that's good enough that corresponds to the notion of identity that we have this compelling sense we must be able to find. I think that the whole sense that there must be some real notion of identity that's preserved over time and over change of state, and change of information, et cetera, and it's really a discrete sense of identity, that you can still use to reason about a world with uploading, with backup copies, and with, especially, I think the most challenging one, having a communication medium between people such that the bandwidth between brains is the same as the bandwidth within a brain, in which case, the discreteness of the individual — I think that's probably the one most challenging to the notion of discrete individuals.

The reason why we have this compelling sense that there must be a definition of identity is what I'll call "creole epistemology." I think that so far historically we haven't lived with uploading and with backup copies and with high-bandwidth communication between brains —

*So our intuition hasn't been selected for accurate portrayal of that sort of world, and what we intuitively guess about those is probably not accurate and probably not the most evolutionarily-stable strategy of epistemology.* So this set me on the following search. I think it's also the case that trying to dispense with the notion of identity altogether is not the right way out.

*Though it's very popular.*

There's all sorts of abstractions that we use to understand complex systems, and denying any reality to some kind of notion of identity out of a reductionist sense that identities aren't real, is to fall into the reductionist fallacy, to use a metaphor of Dean[Tribble]'s, of "denying that the window on my Macintosh is real because it's only built out of pixels." So what I'm looking for is some notion of identity which has the right kind of fluidity to it and still the right kind of sense of coherence and continuity over time and over change, such that, by using that notion of identity, it can be an aid to intuition about post-Singularity life, not a hindrance to it. What I came up with is the notion that we already have the right kind of intuition about the identity of civilizations, the identity of cultures, the identity of natural languages.

---

**So what I'm looking for is some notion of identity which has the right kind of fluidity to it and still the right kind of sense of coherence and continuity over time and over change, such that, by using that notion of identity, it can be an aid to intuition about post-Singularity life, not a hindrance to it.**

---

Let's concentrate on civilizations. One of the key insights that led me down this path was something that I got from my pediatrician, Dr. Einhorn. When I was a kid, I was reading some stuff about the Greeks, and I was really fascinated by the Greeks, and I was home sick and my pediatrician was seeing me, and I got into a conversation with him about Greek civilization; and he also thought the Greek civilization was real neat, and then I mentioned that, "It's really a shame the Greek civilization died." And he said, "They haven't died; we're it." And at first I just had no idea what he meant by that, and then he explained that so much of what our civilization is and where it comes from, is based on the ideas of the Greek civilization, and so many of the values that we have, evolved out of the values of

the Greek civilization, that, in a lot of ways, it really makes sense to think of the Greek civilization, for some descriptive purposes, as being who we are.

The thing that's very nice about what people understand about the identity of civilizations is that people are already used to the fact that civilizations aren't really discrete, they blend and flow into each other, there's multiple influences from multiple paths, and drawing a boundary in space-time around a bunch of people and cultural practices and saying "This is Civilization X" is clearly something that we do as observers, and that we have the choice of several different boundaries. We can't argue that any one set of boundaries is correct. So, in that sense, no set of boundaries is objective, but neither is it the case that all boundaries are equally good; it's not the case that "Anything goes," and "It's all just subjective" in the sense of "Whatever the hell, anything goes, it's all just in the eye of the beholder anyway" and "Why should my boundaries be any less good than your boundaries?"

Another thing about this is that, when we really think about whether or not we should say that "The Greek civilization isn't dead, we're it," whether there's that kind of continuity, largely we're not thinking so much in terms of "pattern identity," to use the phrase from *Engines of Creation*<sup>6</sup>, because our civilization currently is very, very different from Greek civilization, and we're also not using the notion that's very common of "continuity of consciousness," because people generally don't speak of civilizations as being conscious, and besides that, the Greek civilization was suspended in cryonic suspension until its literature was rediscovered.

The question is, what's the criterion we use when we say, "We are the successor to the Greek in such a way that this vague descriptive notion of identity is preserved." That was resolved for me by hearing about Fukuyama's book *The End of History and the Last Man*<sup>7</sup>. So, leveraging off the description that I've heard of that book, I realized that the key issue is credit — credit, and respect, and admiration — Which previous state do we feel good about giving the credit for having been our younger selves?

I think this lets us get past what would otherwise be the terrible paradox of seeking long life in a post-Singularity world, which is: Should you seek to not grow? If you're coming at things from a pattern identity point of view, then to the degree that you expand and learn and grow, you also die — your previous self has died. So, from the point of view of

someone who constructs his goal structure by starting with "I choose life, I want to live," the pattern identity notion would lead you to seek to constrain your future choices to prevent your growth into something that is so far beyond your current self that you're no longer like your current self.

I think that the goal, if you reconstruct the goal of wanting to live in terms of this notion of continuity of identity through credit, a particular kind of credit that's really neither of the ones that Fukuyama mentioned, that's the one that I know I most crave in my intellectual pursuits, is the respect of people I admire. The greater my admiration for those people, the greater my satisfaction in getting any respect from them. They might not even respect me as a peer, but if they respect me at all, the higher my estimate of them, the more that respect for me makes a difference to me.

What I seek to become is a creature that is vastly, vastly greater than my current self, something that is magnificently greater than my current self; and I translate my desire to live into my desire to have that future magnificent self look back on my current puny self with some credit—

*And recognize the debt that it owes you.*

Right. Yeah, "I'm thankful to that guy for having had the seeds to become me." Just like I look back on my four-year-old self and realize, "Oh, yeah, that guy, I can just kind of give him credit for having had what it took to grow into my current self." Let me make a meta-point, by the way. Both myself and the Xanadu project as a whole are very, very careful about credit and references which is in some sense why we're in fact doing this hypertext system where you can make all these links to source material —

*Yes, that's been noted several times before about the whole XanAMIX crowd, and one of the reasons that attracted me to come up here and join in, is the meticulous pinpointing of giving credit where credit is due, and the willingness to spread around credit among sources. That ties in closely with what you said about credit identity —*

To the extent that we're citing our formal selves, our more primitive selves, for the germs of what we become.

About these boundaries being between subjective and objective — Don Lavoie referred me to a book<sup>8</sup>, the basic point of which is that philosophy has been stuck in this false dichotomy: that things are either objectively true, or they're in an anything-goes state. In fact, most everything is in this third state.

Don comes at it from a hermeneutical

point of view, and you can think of these issues as very literary ones, where certain choices of descriptive boundaries "make for a better story" about history than others. You can't objectively say that one novel is better than another, but we all know that novels are not all equal; that novels really are better than each other, but there's no objective way to determine in a disagreement which is which. In some sense it comes down to a matter of taste, but in some sense it doesn't really, because there really is a content to what we're arguing here.

*A lot of matters of taste rely on being consistent with a set of cultural assumptions, in that, although there's no direct reason for saying that one novel is better than another, there are cultural qualities, based on which, valuing what we consider to be the inferior novel would be inconsistent with certain other assumptions in our culture.*

But I believe that even cross-culturally, and even, hypothetically, across species, you'll find that there's enough agreement about some of the extreme cases, especially the extremely bad cases, that there's clearly something going on here that's not just arbitrary and anything goes, but it's also not objectifiable. I don't come at this from a hermeneutical point of view, but from an evolutionary-epistemological point of view. That's simply because I have not gotten into the hermeneutical literature yet. I believe, from what I understand of hermeneutics through Don Lavoie, that there's actually a lot more agreement between the two philosophies than the practitioners of either are inclined to recognize or acknowledge, which is unfortunate.

From this evolutionary-epistemological point of view, all of our knowledge arises out of a process of variation and selection. You can think of the variation component of the evolution of knowledge as being the subjective component, and the selection component (where the selection is by external criteria — there's also subjective selection, but we'll leave that aside), the selection is by virtue of getting mugged by reality. It's through the selection that we get the objective component, but therefore all of our actual knowledge embodies a mixture of the two where it's impossible to separate the mixture. We can understand the process, but we can't tell, for an individual piece of knowledge, how much of its content came from the variation process and how much came from the selection process.

*You said a little bit earlier about continuity of consciousness as a criterion for identity, and you said that Greek civilization was*

*"cryonically suspended" in the form of literature until it was revived. Does consciousness exist? Is consciousness necessary as an assumption, as a hypothesis, or can your personal epistemology get along fine without consciousness?*

I think that, similar to identity, there's a lot of strong intuitions we have about consciousness that are a result of creole epistemology. As well as a result of things that culturally evolved between Freud and computers. There was a lot of evolution of the notion of what consciousness is between those two events, and essentially our folk psychology — which is a wonderful phrase of Dennett's — is still largely a result of that pre-computer evolution.

---

**What I seek to become is a creature that is vastly, vastly greater than my current self, something that is magnificently greater than my current self; and I desire to translate my desire to live into my desire to have that magnificent future self look back on my current puny self with some credit.**

---

There's many parts of what we mean by consciousness. The idea of there being a "real you" at a point inside you that experiences things in a serial order and knows what it believes — understand, I have not read the book — but from what I understand, Dennett does a very good job of trashing that, and I think it well deserves to be trashed: homunculus theory?

*That's what Dennett calls it, the homunculus, the "little man in the control room" —* By the way, let me strongly recommend to everyone the segment of Woody Allen's movie "Everything You Always Wanted To Know About Sex, But Were Afraid To Ask" with Tony Randall in the control room of the guy who's having sex, and Woody Allen as a sperm who's agonizing about "What if he's not even having sex? Maybe I'll end up on the ceiling!" And the thing that's wonderful about that is it's a very strange and wonderful mixture of homunculus theories, and society-of-mind theories; it's kind of got a lot of

mixture of both. It's actually very sophisticated. I really recommend going back and watching that after having read *Society of Mind*.<sup>9</sup>

Okay.

So, our notion of consciousness that has much truth to it, is the notion of consciousness being reflective—that it really is quite a trick to be able to have been doing something without articulating to yourself what it was you were doing, and then to "hop back" and say, "Oh, so I was doing that" and then you're articulating, you now have a thought you can think about. And this whole ability, not just to introspect, but one of the things we notice about articulate knowledge versus inarticulate knowledge is that inarticulate knowledge is learned gradually with training. Articulate knowledge, through, I'll say, through the mechanism of consciousness—and what I mean there is purposefully vague—is able to learn many things on one hearing, because it's manipulating a very different kind of learning process. A lot of the trick of the brain is the gluing together of those two very different kinds of learning, and, in particular, the way in which they each bootstrap the other.

Another aspect of consciousness that is valid is that part of the trick of consciousness is that your articulate construction of the world has as part of its articulate construction a symbol in there if you will—a designator, a node—that represents this creature itself. So this notion of consciousness is intimately tied in with the notion that we have of identity, which is: our creole notion of identity is the thing that's there in that slot—

#### *Experiencing the consciousness.*

In the conscious articulation of how the world works and what's going on there, there's this idea structure of the identity of this personal self that's a creole self—a creole identity of the self as well as a creole identity for consciousness—but that's the thing that the rest of the conscious structure uses in thinking about its place in reality.

So I think that all of those are valid pieces and part of the pattern, but I think that not only is a lot of our other stuff about consciousness, like the homunculus stuff, wrong; but a lot of the stuff that's right is much less important than people verbally think it is—that, in some sense, we're much less conscious than we think we are, and the reason is that the thoughts that we're aware of, are the ones that we're aware of.

Can you give an example of an instance in which people are less conscious than they

think they are?

Okay. People generally think that they have a lot of introspective access to how they think. However, for relatively simple computational tasks—the kinds of things that people are conscious of what they do, of how they do it when they do it—are generally much simpler computational tasks than the things that people completely take for granted and do "without thinking." The fact that so many of those are of such incredibly greater sophistication than the things that we do "with thinking" means that the fact that we think that it's in our "thinking" process that's where the sophistication lies, is probably wrong.

One particular example that I really love is "Hansel and Gretel." In the book *Distributed Artificial Intelligence*, edited by Bond and Gasser<sup>10</sup>, one of the papers has this diagramming notation which they're exploring so they can diagram not only "Actor X believes A," but they can diagram what people believe about other people's beliefs. "At this point in the story, this guy believed that that guy believed that this person was going to do that, whereas this person himself believed that this guy believed that he himself was going to do something else."

Right.

So they've got a good diagramming notation for laying out these kinds of nested belief structures. They explain this by diagramming the beginning of the Hansel and Gretel story—the part of the story that is *before* it gets interesting—the part that you just read to a kid and assume that he gets it "without really thinking about it." It's not considered to be a tough comprehension for very small children, but you end up with this massive structure to explain it. I think I'm good enough at evaluating notations and formalisms to look at it and say, the massive structure is not the result of a bad formalism. The formalism is good; the formalism does not create any combinatorial explosions here. The massive structure is the result of the fact that human beings are wired for this incredible ability to do nested modeling of belief structures—very nested modeling, three and four levels deep without straining.

*It could also be considered empathic in the sense of putting yourself into someone else's position and modeling their beliefs by adopting their viewpoint.*

I think that one of the reasons that people always seem conscious to us, and no AI system has seemed conscious to us yet, is that when I form a model of you "without thinking about it," part of my model is my

model of your model of me, and my model of your model of my model of you. It's certainly the case that no AI system has enough knowledge of our social conventions and of other human beings that, when we're interacting with an AI system, the models that we construct that have explanatory power to us about how it's going to behave, don't include that kind of deep nesting because it doesn't have enough ability to accurately understand my model of it, or its modeling of my model of it, to have influence in what my model of it is. There's a lack of the kind of deep nesting that we normally do "without thinking about it" with human beings, and I think it's what that kind of deep nesting provokes, that is what we have attached the label *consciousness* to. That's why, no matter how reflective a computer system is, it never seems conscious to us.

*In Part Two of this interview, Mark and I discuss the five kinds of libertarianism, the Reverse Polish Moon Treaty, "nanarchy," and the role of gossip in a free society.*

## REFERENCES

- <sup>1</sup>Dawkins, Richard. *The Extended Phenotype*. Oxford: Freeman, 1982.
- <sup>2</sup>"Creole Languages" by Derek Bickerton, *Scientific American*, July 1983.
- <sup>3</sup>"Spacetime Constraints" by Andrew Witkin and Michael Kass, *ACM SIGGRAPH*, Vol. 22, No. 4; August, 1988.
- <sup>4</sup>Daniel Kahneman, Paul Slovic, and Amos Tversky, eds. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1982.
- <sup>5</sup>Nisbett, Richard E., and Lee Ross. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- <sup>6</sup>Drexler, K. Eric. *Engines of Creation*. Garden City, N.Y.: Anchor Press/Doubleday, 1986.
- <sup>7</sup>Fukuyama, Francis. *The End of History and the Last Man*. New York: Free Press, 1992.
- <sup>8</sup>Bernstein, Richard J. *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis*. Philadelphia: University of Philadelphia Press, 1983.
- <sup>9</sup>Minsky, Marvin. *The Society of Mind*. New York: Simon and Schuster, 1986.
- <sup>10</sup>Alan H. Bond and Les Gasser, eds. *Readings in Distributed Artificial Intelligence*. San Mateo, Calif.: M. Kaufmann, 1988.

# Now! Freely Available To Americans... EUROPEAN SMART DRUGS!

You don't have to wait for the FDA to "approve" of new memory and intelligence enhancing therapies. For your personal use, you can obtain therapies that can:

- Enhance mental function *four times* better than Hydergine!
- Increase mental energy, concentration, and alertness!
- Maximize the ability to memorize facts, figures, and scientific findings
- Increase learning, mental agility, and I.Q.
- Improve test scores leading to higher grades and work productivity!
- Prevent and treat degenerative brain diseases including Alzheimer's Disease, Parkinson's Disease, etc.

The Life Extension Foundation has compiled the sources, the doctors and the information to enable Americans to obtain life extension therapies from around the world. These drugs have been safely used in France, Germany and other countries for years, but until now, have been denied to Americans.

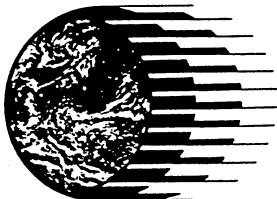
Europeans are using these drugs to improve their mental condition to unprecedented levels. Now Americans can gain access to advanced life extension therapies such as **Piracetam**, **Centrophenoxine**, **human growth hormone**, and a new memory-enhancing product that is *four times* more effective than **Hydergine**.

As a member of **The Life Extension Foundation**, you will learn about the scientific basis for using these life-enhancing therapies, their side effects, and how you can get them!

For just \$50.00, you will receive *Life Extension Report* and *Life Extension Update* each month plus:

1. ***The Directory Of Life Extension Doctors***. A nationwide directory of doctors who are knowledgeable about these advanced therapies and may be willing to prescribe them for you.
2. ***The Physician's Guide To Life Extension Drugs***. The first book ever published to provide American doctors with information about safe and effective "unapproved drugs". This book is referenced to enable the lay person to understand and find therapies for specific purposes.
3. ***The Directory Of Innovative Medical Clinics***. If you were told you had an incurable disease would you believe your doctor? A disease your doctor says is "untreatable" may already have a cure that the FDA has not yet "approved" of. There are scientists with impeccable credentials who are effectively treating so-called "terminal" victims of cancer, Alzheimer's Disease, etc. You can now access these advanced research centers with ***The Directory Of Innovative Medical Clinics***.
4. ***Discounts of 25% to 50%*** on your vitamin purchases. Members buy name brand nutrient supplements and advanced life extension formulas at super discount prices.
5. ***Discounts of 20%*** on all your prescription drug purchases including popular life extension drugs such as **Hydergine** and **Eldepryl**. THE MAIL-ORDER PHARMACY saves members hundreds of dollars a year on their prescription drug purchases.

The Life Extension Foundation is the *only* organization in the world that tells you how to obtain the most advanced life extension therapies in the world...long *before* they are "approved" by the FDA. You will be the *first* to find out about products that will enhance your life.



Mail to:  
Life Extension Foundation  
P.O. Box 229120  
Hollywood, Florida 33022

To join, use the coupon or call: **1-800-841-5433**

Enclosed is my \$50.00 membership. Please enroll me in your life extension program which includes two newsletters each month and the three directories of life extension doctors, drugs, and clinics and the super discounts on my vitamin and prescription drug purchases.

EXTR

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_

Credit Card # \_\_\_\_\_ Exp Date \_\_\_\_\_

# Nanocomputers

## 21st Century hypercomputing

J. Storrs Hall

If the price and performance figures for transportation technology had followed the same curves as those for computers for the past 50 years, you'd be able to buy a top-of-the-line luxury car for \$10. What's more, its mileage would be such as to allow you to drive around the world on one gallon of gas. That would only take about half an hour since top speed would be in the neighborhood of 50,000 mph (twice Earth's escape velocity).

Oh, yes, and it would seat 5000 people.

Comparisons like this serve to point out just how radically computers have improved in cost, power consumption, speed, and memory capacity over the past half century. Is it possible that we could see as much improvement again, and in less than another half century? The answer appears to be yes.

At one time, people measured the technological sophistication of computers in "generations". There were vacuum tubes, discrete transistors, IC's, and finally large scale integration. However, since the mid-70's, the entire processing unit has increasingly come to be put on a single chip and called a "microprocessor." After these future advances have happened, today's microprocessors will look the way ENIAC does to us now. The then-extant computers need a different name; we'll refer to them as nanocomputers. "Micro" does exemplify at least the device size (on the order of a micron) and instruction speed (on the order of a microsecond) of the microprocessor. In at least one design, which we'll examine below, the nanometer and nanosecond are the appropriate measures instead.

Do we really need nanocomputers? After all, you have to be able to see the screen and press the keys even if the processor is microscopic. The answer to this question lies in realizing just how closely economics and technological constraints determine what computers are used for. In the mid-sixties, IBM sold a small computer for what was then the average price

of a house. Today, single-chip micros of roughly the same computational power cost less than \$5 and are used as controllers in toaster-ovens. Similarly, we can imagine putting a nanocomputer in each particle of pigment to implement "intelligent paint", or at each pixel location in an artificial retina to implement image understanding algorithms.

This last is a point worth emphasizing. With today's processing technology, robots operating outside a rigid, tightly controlled environment are extremely expensive and running at the ragged edge of the state of the art. Even though current systems can, for example, drive in traffic at highway speeds, no one is going to replace truck drivers with them until their cost comes down by some orders of magnitude. Effective robotics depends on enough computational power to perform sensory perception; nanocomputers should make this cost-effective the way microcomputers did text processing.

Beyond providing robots with the processing power humans already have, there is the opportunity of extending those powers themselves. Nanocomputers represent enough power in little enough space that it would make sense to implant them in your head to extend your sensorium, augment your memory, sharpen your reasoning. As is slowly being understood in the world of existing computers, the interface to the human is easily the most computationally intensive task of all.

Last but not least – in some sense, the most important application for nanocomputers – is as the controllers for nanomechanical devices. In particular, molecular assemblers will need nanocomputers to control them; and we will need assemblers to build nanocomputers. (In the jargon of nanotechnology, an "assembler" is a robot or other mechanical manipulator small enough to build objects using tools and building blocks that are individual molecules.)

### What is a nanocomputer?

Currently, the feature sizes in state-of-the-art VLSI fabrication are on the order of half a micron, i.e. 500 nanometers. In fifteen years, using nothing more than a curve-fitting, trend-line prediction, this number will be somewhere in the neighborhood of 10 nanometers; would it be appropriate to refer to such a chip as a nanocomputer?

For the purposes of this article, no. We want to talk about much more specific notions of future computing technology. First, we're expecting the thing to be built with atomic precision. This does *not* mean that there will be some robot arm that places one atom, then another, and so forth, until the whole computer is done. It means that the "design" of the computer specifies where each atom is to be. We expect the working parts (whether electrical or mechanical) to be essentially single molecules.

We can reasonably expect the switches, gates, or other embodiment of the logical elements to be on the order of a nanometer in size. (They may have to be further apart than that if electrical, due to electron tunneling.) In any case, it is quite reasonable to expect the entire computer to be smaller than a cubic micron, which contains a billion cubic nanometers.

### Nanotechnology

The nanotechnology-assembler-nanocomputer dependence sounds like a self-referential loop, and it is. But many technologies are that way; machine tools make precision parts used in machine tools. Bootstrapping into a self-supporting technology is not a trivial problem, but it's not an impossible one either.

Another self-referential loop in nanotechnology is slightly more complicated. We would like assemblers to be self-reproducing. This would allow for nanocomputers and other nanotechnological products to be inexpensive, be-

cause each fixed initial investment would lead to an exponentially increasing number of nanomechanisms, rather than a linearly increasing number.

But how can a machine make a copy of itself? The problem is that while we can imagine, for example, a robot arm that can screw, bolt, solder, and weld enough to assemble a robot arm from parts, it needs a sequence of instructions to obey in this process. And there is more than one instruction per part. But the instructions must be embodied in some physical form, so to finish the process, we need

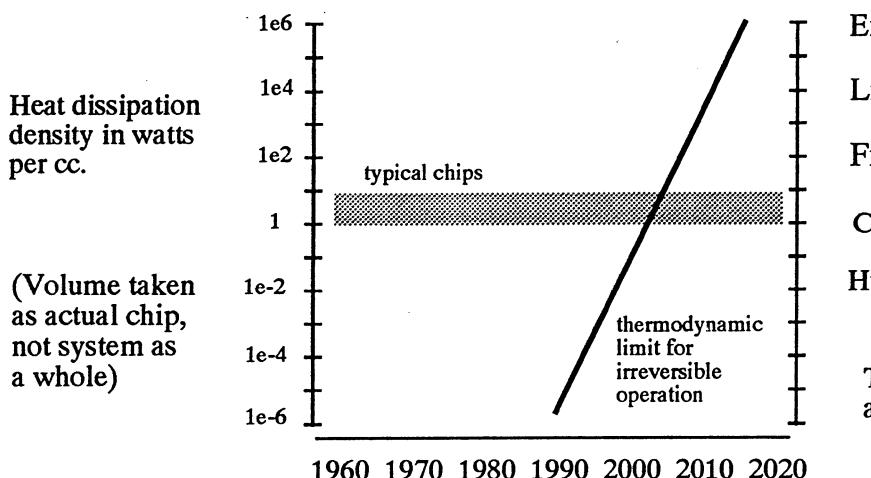


Figure 1.

Heat dissipation vs its theoretical limit

instructions to build the instructions, and so on, in an infinite regress. The answer to this seeming conundrum was given mathematically by John von Neumann, and at roughly the same time (the '50's) was teased out of the naturally occurring self-reproducing machines we find all around us, living cells. It turns out to be the same answer in both cases.

First, design a machine that can build machines, like the robot arm above. (In a cell, there is such a thing, called a ribosome.) Next, we need another machine which is special purpose, and does nothing but copy instructions. (In the cell it's called a replisome.) Finally, we need a set of instructions that includes directions for making both of the machines, plus whatever ancillary devices and general operating procedures may be needed. (In the cell, this is the DNA.) Now we read the instructions through the first machine, which makes all the new machinery necessary. Then we read it through the second machine, which makes a new set of instructions. And there's our whole new self-reproducing system, with no infinite regress.

## General computer principles

When you get right down to it, a computer is just a device for changing information. You put in your input, and it gives you your output. If it's being used as the controller for anything from a toaster to a robot, it gets its input from its sensors and gives its output to its motors, switches, solenoids, speakers, and what have you.

Internally, the computer has a memory, which is used to store the information it's working on. Many functions, even so simple as a push-on, push-off

computer, then, we need some way to remember bits; some way to perform boolean logic; and some way to clock the sequence of remember, perform, remember, perform, etc.

In computers from historical times to date, information has been encoded either as the position of a mechanical part, the voltage on a wire, or a combination (as in a relay-based computer). The major reason is that these are the easiest encodings to use in the logic part of the particular technology in question. It seems reasonable to expect to see these encodings at the nano level, for the same reasons.

## Additional constraints for nano-computers

About a year ago I had the occasion to design a nano-computer using Eric Drexler's mechanical rod logic, which will be examined in detail later in this article. As someone who was used to the size and speed constraints of electronics, I was in my glory with this new medium. I went

wild, adding functionality, pipelining, multicore computing, the works. I could build a super-computer beyond the wildest dreams of Cray, the size of a bacterium! What I *didn't* do was pay any attention to "this crazy reversible computing stuff."

— Until I did the heat dissipation calculations. The problem is that there really is a fundamental physical limit involved in computation, but it represents an amount of energy so small (it's comparable to the thermal energy of one atom) that it is totally negligible in existing physical devices. But in a nanocomputer, it far outweighs all the other heat-producing mechanisms; in fact, my nanocomputer design had the same heat dissipation per unit volume as a low-grade explosive. Back to the drawing board...

Since the earliest electronic computers in the 1940's, energy dissipation per device has been declining exponentially with time. Device size has undergone a similar decline, with the result that overall dissipation per unit volume has been relatively constant (see the horizontal bar in Fig. 1). Historically, the portion represented by the thermodynamic limit for irreversible operations was completely

lightswitch, need memory by definition. But the computer also uses memory to help break the job down to size, so as to be able to change the data it receives in little pieces, one at a time. The more memory you allow, and the smaller the pieces, the simpler the actual hardware can be.

There is a "folk theorem" in the computer world that the NAND gate is computationally universal. This is true in the sense that one can design any logic circuit using only NAND gates. However, something much more surprising is also true. Less than ten years ago, Miles Murdocca, working at Bell Labs, showed (in an unpublished paper) how to build a universal computer using nothing but delay elements and *one single OR gate*. Just one. Not circuits using arbitrarily many of just one kind of gate.

Murdocca's computer works by driving the notion of a computer to its very barest of essentials. A computer is a memory and a device to change the information in the memory. Generally we add two more specifics: The information is encoded and changed under the rules of Boolean logic; and the changes happen in synchronized, discrete steps. To build a

insignificant; it is still in the parts per million range. However, with densities and speeds in the expected range for nanocomputers, it is extremely significant. Thus nanocomputer designers will be forced to pay attention to reversibility.

Efficient computers, like efficient heat engines, must be as nearly reversible as possible. Rolf Landauer showed in a landmark paper (in 1961) that the characteristic irreversible operation in computation is the erasure of a bit of information; the other operations can be carried out in principle reversibly and without the dissipation of energy. And as in heat engines, the reason reversibility matters is the Second Law of Thermodynamics, the law of entropy.

## Entropy

The subject of entropy seems to give rise to more misconceptions and disagreements than any other scientific principle. Relativity and quantum mechanics have been similarly abused, but much of the abuse is in the form of extensions of the concepts that are frankly metaphorical. Relativity and quantum mechanics do not, in general, apply to the everyday world, but entropy does. When people use metaphorical extensions of entropy on phenomena that are governed by actual entropy, confusion occurs.

It's possible, on the other hand, to give a metaphorical explanation of entropy that is carefully rigged to give all the same answers as actual entropy. Here it is; but if you'd prefer to take my word for it that nanocomputers *must* be reversible, you can skip to the next section.

Let us suppose that we are going to have a computer simulation of some closed physical system. We can have as high an accuracy as we like, but the total amount of information, i.e. the number of bits in the computer's memory, is in the end some fixed finite number. Now since the physical system we're simulating is closed, there will be no input to the simulation once it is started.

Since there is a fixed number of bits, say  $K$ , there is a fixed number of possible descriptions of the system the simulation can ever possibly express, namely  $2^K$  of them. Now by the first law of thermodynamics, conservation of energy, total energy in a closed system is constant. Thus we can pick all of the states with a given energy, and call them "allowable", and the rest are forbidden. The first law constrains the system to remain within the allowable subset of states but says no more about which states within that set the system will occupy.

There is another constraint, however,

in the sense that the laws of physics are deterministic; given a state, there is a single successor state the system can occupy in the next instant of time. (In the real world, this is more complicated in two ways: Time and the state space are continuous, and quantum mechanics provides multiple successor (and predecessor) states. However, the mathematical form of quantum mechanics (i.e. Hamiltonian transformations) gives it properties analogous to the model, so for perspicuity, we will stick with the discrete, deterministic model.) What is more, the laws are such that each state has not only a unique successor, but a unique predecessor.

Let's try to make this notion a little more intuitive. Each "state" in our computer simulation corresponds to some description of all the individual atoms in the physical system. For each atom, we know exactly where it is, exactly how fast it is going, exactly in what direction, etc. As we move forward in time, we can calculate all the electrical, gravitational, and if we care, nuclear, forces on that atom due to all the other atoms, and compute just where it will be some tiny increment of time in the future. Clearly, to just the same degree of precision, we can calculate exactly where it must have been the same tiny amount of time in the past. The math of the physical laws allow you to simulate going backwards just as deterministically as you can simulate going forwards.

So, suppose we have a simulation of a box which has a partition dividing it in half. There is some gas in one half, i.e. atoms bouncing around, and none in the other. Now suppose the partition disappears: the atoms that would have bounced off it will continue on into the empty half, which pretty soon won't be empty any more. The atoms will be distributed more or less evenly throughout the box.

What happens if we suddenly stop the simulation and run it backwards? In fact, each atom will retrace the exact path it took since the partition disappeared, and by the time the partition should reappear, the atoms will all be in the original half.

In reality, we don't see this happen. Remember that in our model there is a distinct causal chain of states from the state where the atoms are all spread out but about to move into half the box, to the state where they are actually in half the box. This means that the number of states from which the atoms are about to compress spontaneously (in some specific number of timesteps) is the same as the number of states in which they are all in one half of the box.

The important thing to remember is

that the total energy (which is proportional to the sum of the squares of the velocities of the atoms) must be the same. If we used a simulated piston to push the atoms back into the original half, we would find a 1-to-1 mapping between spread-out states and compressed ones; but the compressed ones would be higher-energy states.

How many states are we talking about here? Well, suppose that all we know about any specific atom is which side of the box it is in, which we can represent with a single bit. If the box has just 100 atoms in it, there will be more than 1,267,000,000,000,000,000,000,000,000 states in which the atoms are spread around evenly, and one state where they are all on one side. A similar ratio holds between the number of states (with full descriptions) where the atoms are spread-out, and the subset of those states where they are about to pile over into one side.

---

**It's clear that quantum mechanics allows for mechanisms that capture a single electron and hold it reliably in one place. Individual electrons doing specific, well-defined things under the laws of quantum mechanics is what happens in typical chemical reactions... there is no basic physical law that prevents us from building nanocomputers that handle electrons as individual objects.**

---

We are now going to talk about entropy. In order to relate the simulation model of a physical system to the way physical scientists view physical systems, we'll use the term "microstate" to represent what we have been calling a state in the simulation, i.e. one specific configuration of the system where all the bits are known. We'll use "macrostate" to refer to what a physical scientist thinks about the system. This means knowing the temperature, pressure, mass, volume, chemical composition, physical shape, etc, but not knowing everything about every

atom.

Clearly, there are lots of microstates in a macrostate. The log of the number of microstates in a given macrostate is the entropy of the macrostate. (Physical entropies are generally given as natural logs, but we will talk in terms of base 2 logs with the understanding that a scaling factor may be needed to translate back.) To put it more simply, the entropy of a given macrostate is the number of bits needed to specify which of its microstates the system is actually in.

With these definitions, the second law of thermodynamics is quite straightforward. If a proposed physical transformation maps a macrostate with a higher entropy into one with a lower entropy, we know it is impossible. Remember the causal chains of (micro)states: they can neither branch nor coalesce. Now suppose at the beginning of some physical process, the system was in a macrostate with a trillion microstates; we have no idea which microstate, it could be any one of them. Therefore at the end of the process, it can still be in a trillion microstates, each at the end of a causal chain reaching back to the corresponding original microstate. Obviously, the system cannot be in any macrostate with fewer than a trillion microstates, i.e. with a lower entropy than that of the original macrostate.

Now suppose I have a beaker of water at a specific temperature and pressure. It has, according to a physicist or chemist, a specific entropy. But suppose I happen to know a little more about it, e.g. I have a map of the currents and vortices still flowing around in it from when it was stirred. There are a lot of the microstates that would be allowed from the simpler description, that I know the water is not really in. Isn't its entropy "really" lower? Who gets to say which is the "real" macrostate, whose size determines the "true" entropy of the system?

The answer is, that entropy isn't a property of the physical system at all, but a property of the description. After all, the real system is only in one single microstate! (Ignoring quantum mechanics.) This does sound a bit strange: Surely the "true" entropy of any system is then 0. And we should be able to induce a transformation from this system into any macrostate we like, even one with much lower entropy than that of the original macrostate of the system as conventionally measured.

Let's consider the little box with the atoms of gas in it. The gas is evenly spread through the box, a partition is placed, and there is a Maxwell's Demon with a door to let the atoms through selectively. But the

demon isn't going to try anything fancy. We're going to assume that we know the exact position and velocity of each atom in advance, so we will be able to provide the demon with a control tape that tells him when to open and close the door *without observing the atoms at all*. In fact, this would work; the demon can herd all the atoms into one side without expending any energy.

Why doesn't this violate the second law? Well, let's count up the causal chains. The entropy problem in the first place is that there are many many fewer microstates in the final macrostate, namely the one with all the atoms on one side, than in the original, so that many original microstates must somehow map into a single final one. But with the demon at work, we can run the simulation backwards by running the demon backwards too; the sequence of door-opening and closing that got us to our particular final microstate is clearly enough information to determine which original microstate we started from. Thus the final state, *including the tape*, is still in a one-to-one mapping with the original state, and the second law is not violated.

The curious thing to note about this gedanken experiment is that the demon *can* compress the gas without expending energy; what he *cannot* do is erase the tape! This would leave the system with too few final states.

What happens if the demon starts with a blank tape, instead of one where the microstate of the system is already recorded? Can he measure the system on the fly? Again yes, but only if he writes his measurements on the tape. Again the critical point is that the data on the tape serves to make the number of possible final microstates as large as the number of possible original microstates.

In practice, of course, the way one would obtain the same result, i.e. moving the atoms into half the box, would be to use a piston to compress the gas and then bring it in contact with a heat sink and let it cool back to the original temperature. Energy, in the form of work, is put into the system in the first phase and leaves the system, in the form of heat, in the second phase. At the end of the process the system is the same as the demon left it but there is no tape full of information. Clearly there is some sense in which the dissipation of heat is equivalent to erasing the tape.

In terms of the simulation model, the demon directly removes one bit from the position description of each atom (storing it on the tape). The piston compression moves a bit from the position to the velocity description, and the cooling process

removes that bit (storing it in the heat sink). The entropy of the gas decreases, and that of the heat sink increases.

Of course, dissipating heat is not the only way to erase a bit. Any process that "moves" entropy, i.e. decreasing it in one part of a system at the expense of another part, will do. For example, instead of increasing the temperature of a heat sink, we could have expanded its volume. Or disordered a set of initially aligned regions of magnetization (in other words, written the bits on a tape). Or any other physical process which would increase the amount of information necessary to identify the system's microstate. However, heat dissipation is probably the easiest of these mechanisms to maintain as a continuous process over long periods of time, and it is well understood and widely practiced.

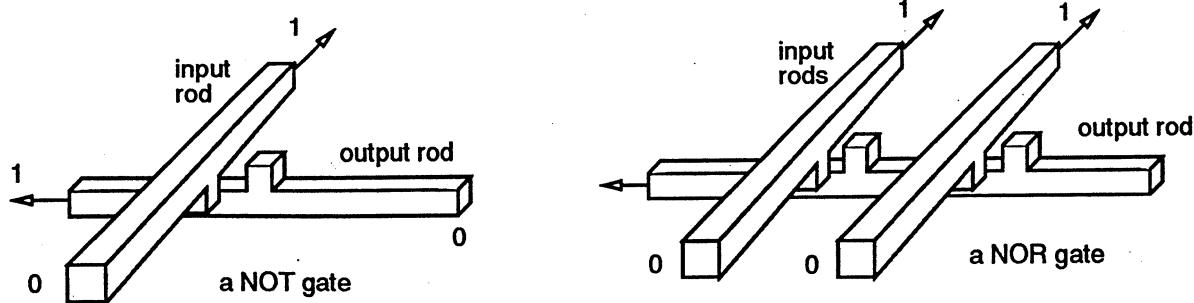
A state-of-the-art processor, with 100,000 gates erasing a bit per gate per cycle, at 100 MHz, dissipates about 28 nanowatts due to entropy. (At room temperature. Each bit costs you the natural log of 2, times Boltzmann's constant, times the absolute temperature in Kelvins, joules of energy dissipation, which comes to about 2.87 mJ (milli-attoJoule,  $10^{-21}$  Joules)). Since it actually dissipates 100 million times this much, or more, nobody cares. But with a trillion-fold decrease in volume and thousand-fold increase in speed, the nanocomputer is "a whole 'nother ball game."

Thus there are two new design rules that the nanocomputer designer must adopt:

- (1) Erase as few bits as possible.
- (2) Eliminate entropy loss in operations that do not erase bits.

We eliminate entropy loss in logical operations by what is known as "logical reversibility". Suppose we have in our computer registers A and B, and an instruction ADD A,B that adds A to B. Now in ordinary computers that would be done by forming the sum A+B, erasing the previous contents of register B, and then storing the sum there. However, it isn't logically necessary to do this; since we can recreate the old value of B by subtracting A from the new value, no information has been lost, and thus it is possible to design a circuit that can perform ADD A,B *without* erasing any bits.

Addition has the property that its inputs and results are related in such a way that the result can replace one of the inputs with no loss of information. However, many useful, even necessary, functions don't have this property. We can still use those functions reversibly; the



**Figure 2. Rod Logic.**

only trick needed is not to erase the inputs! Ultimately, of course, you have to get rid of the input in order to process the next one; but you can *always* erase the *output* without entropic cost if you've saved the input.

This leads to structures in reversible computation called "retractile cascades." Each of a series of (logical) circuits computes a function of the output of its predecessor. If the final output is erased first, and then the next-to-last, and so forth, the entire operation is reversible, and can be done (in theory!) without any energy dissipation.

If we adopt these rules throughout our computer design, we can reduce the number of bits erased per cycle from around 100,000 to around 10.

### Drexler's mechanical logic

K. Eric Drexler, "the father of nanotechnology", has designed and subjected to a thoroughgoing analysis, a *mechanical* logic for nanocomputers. A mechanical logic has the disadvantage of being slower than an electronic one, but has the major advantage that at the molecular level it is possible to design, and analyse the operation of, a mechanical logic with current molecular simulation software, and be reasonably certain that the design, if built, would work.

By the time we get around to actually building molecular computers, our analytical tools will be better than they are now, and what's more we'll be able to augment the simulations with physical experiments. So real nanocomputer designs won't have to be nearly so conservative as this one. In particular, they'll probably be electronic, and thus probably some orders of magnitude faster. But don't worry: this mechanical design is already plenty fast.

This formulation is sometimes called "rod logic" because instead of wires, it uses molecular-sized rods. (E.g. a nanometer in diameter and from ten to a hundred nanometers long.) Each rod represents a logic 0 or 1 by its position, sliding slightly along its length to make the transition.

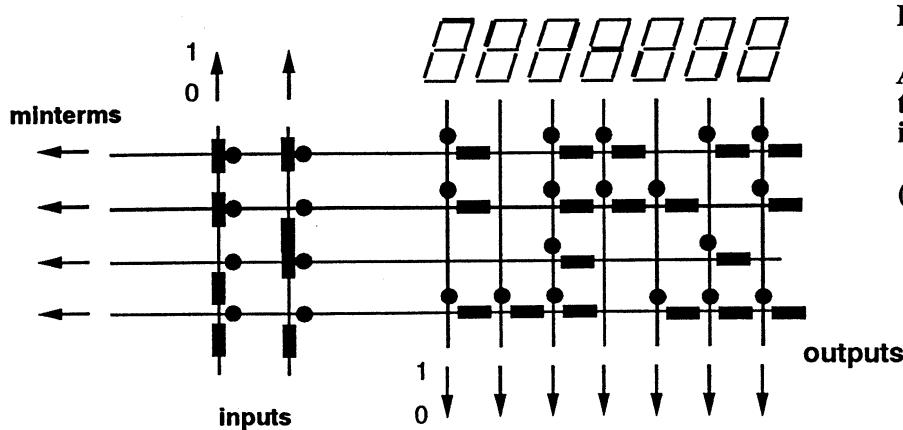
To do logic, the rods have knobs on them which may or may not be blocked by something, preventing the rod from changing state. The "something" is simply other knobs on other rods, which block or don't block the first rod, depending on *their* state. (see Fig. 2) The logic is clocked by pulling on rods through the equivalent of a spring, so that it moves unless blocked. (We can draw a workable parallel to transistors, which block or don't block a clock from changing the voltage on a wire, depending on the voltage of another wire.)

The rods move in a fixed, rigid housing structure which might be thought of as a hunk of diamond with appropriate channels cutout of it (although it wouldn't be built that way). The rods are supported along their entire length so the blocking does not place any bending stress on them.

Any logic function is now simply constructed: for an AND gate, for example, take two input rods and place knobs so that they block the output rod when they are in the "0" position. The output rod will only be able to move to the "1" when both inputs are "1."

(Now that we can build a single gate, aren't we just about finished, by virtue of Murdocca's design? Well, we'd still need clocking and some mechanism to handle a delay-line memory; but more to the point the design produces a computer that is about a billion times slower than you could build with conventional logic designs!)

The motion of the rods is limited by the speed of sound (in diamond); but they are so short (e.g. one-tenth of a micron) that the switching times are still a tenth of a nanosecond. The speed of an entire nanocomputer of this kind of design will be limited by thermal noise and energy dissipation, which can produce enough variation in the shapes of the molecular parts to keep them from working right. Drexler gives a detailed analysis of the

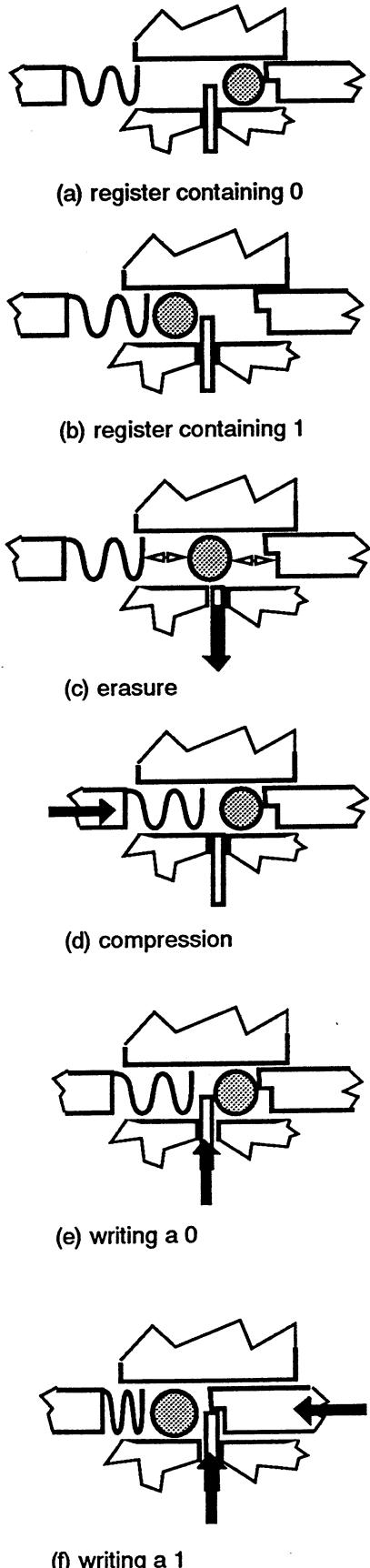


**Figure 3**

A PLA which translates a two-bit number (0, 1, 2, 3) into a 7-segment display

( $0, 1, 2, 3$ ).

sources of such error in *Nanosystems* (chapter 12). The energy dissipated per switch-



**Figure 4 Register operation**

ing operation is conservatively estimated at about 0.013 mJ.

This is less than 5% of the fundamental limit for bit destruction; as long, of course, as the gate in question doesn't destroy bits! Most of the logic design in a rod logic nanocomputer *must* be either conservative logic or retractile cascades. To demonstrate the difference, consider a NOT gate. One can implement a conservative NOT because it has an exact inverse (which happens to be itself). One could implement a conservative NOT in rod logic with a single gear meshing with racks on input and output rods. A retractile NOT on the other hand would be a single rod crossing with knobs preventing the output from moving to "1" if the input was "1". The "retractile" part is that the output must be let back down easy, in such a way that the energy stored in the spring is retrieved, before the input is reset for the next operation.

If this is not done, e.g. if the input is released first, the output rod will return under force of its spring and the energy stored in the spring will be dissipated as heat. In order for this not to happen, the output rod must be returned first, and then the input may be.

One very powerful and widely used technique in logic design is called the PLA (programmed logic array). A PLA is readily designed in retractile cascade style; it also has a remarkably good match to the geometric constraints of the rod logic, which requires the input and output rods from any interaction to form a right angle. The PLA consists of three sets of rods: the inputs, the minterms, and the outputs (see Fig. 3). First the input rods are moved, i.e. set to the input values. Then the minterm rods are pushed; some of them move and some don't depending on which inputs blocked them. In an electronic PLA each input is both fed directly into these interactions, and its negation is; this need not be done in the rod logic since its effect can be had by altering the position of the knobs. Sometimes the number of minterms can be reduced for the same reason.

After the minterm rods are pushed, the output rods are pushed, and the appropriate value is encoded in which ones actually move. Now, the important thing in preserving reversibility (what makes this a retractile cascade) is that after this operation, first, the output rods must be let back down gently; then the minterms let back down gently; and finally the inputs can be released.

Notice that in the figure, the input rods are at rest in the 0 position, while the output rods are at rest in the 1 position. (And in any given operation, exactly one

of the minterm rods will slide to the left.)

PLA's can implement any logic function necessary in a computer, although there are more efficient circuits for some of them that are commonly used instead. More crucial, however, to a full grasp of the mechanisms of a nanocomputer, is memory.

## Registers and memory

Memory is a problem; if we follow the rules for conservative or retractile reversible logic, memory is impossible to implement. This is because any memory with a "write" function erases bits *by definition*.

In Drexler's rod-logic design, all the bit-erasing functionality is concentrated in the registers. The register design is fairly complex, to keep the energy dissipated in this process near its theoretical minimum.

The main problem is that for a physical system to retain one of two states reliably, which is what you want in a memory, there must be a potential barrier between the states that is significantly higher than  $kT$ , or thermal fluctuations will be sufficient to flip the bit at random. But in simple implementations, the height of the barrier determines how much energy is lost when the system changes state.

Consider an ordinary light switch. When you flip it, there's a spring that resists your finger until the halfway point, and then it snaps into place, dissipating all the energy you put into it as heat, vibration, and sound. (A "silent" switch is worse, since it dissipates by friction and you have to push all the way across.) The weaker the spring, and the more likely that some vibration will flip the switch when you didn't intend it. The trick is to have some way to change the strength of the spring (or to have the effect of doing so).

In Figure 4, there is a simplified version of Drexler's register. The bit it contains is reflected in the position of the shaded ball (In the real design it's more complicated so that the value can be read!). (a) and (b) show the register when it contains 0 and 1 respectively. In (c), the barrier has been lowered and the ball is free to wander freely between both positions; this stage increases entropy. In (d), the register is reset to 0. The similarity to compressing a gas-filled cylinder is apparent; this is where  $\ln(2) kT$  joules of work are converted into heat. Now to write the next bit, the input rod (on the right) is either extended (a 1, see (f)) or not (a 0, see (e)) and then the barrier raised. Finally, the spring rod (on the left) is retracted to get back to (a) or (b). If a 1 was written, the input rod did work to com-

press the ball into the spring, but that energy can be retrieved when the spring rod is retracted. The mechanisms to do this are just the same as in the logic portions, e.g. having the rods mechanically coupled to a flywheel.

Registers like this which are going to be used to erase bits will tend to be located near heat sinks or coolant ducts; bit erasure is the largest component of power dissipation in the rod logic design. Memory can be implemented as lots of registers; registers occupy about 40 cubic nanometers per bit. Thus about 3 megabytes worth of registers fill a cubic micron. One would probably use register memory for cache, and use a mechanical tape system for main storage, however. The "tape" would be a long carbon chain with side groups that differed enough to be 1's and 0's. Since the whole computer is mechanical, the difference in speeds is not as bad as macroscopic tapes on electronic computers. Such a tape system might have a density in the neighborhood of a gigabyte per cubic micron. Access times for using a tape as a random access memory consist almost entirely of latency; if the length of individual tapes is kept to under 100 kbytes, this is in the 10's of microseconds.

## Motors

In order to drive all this mechanical logic we need a motor of some kind; Drexler has designed an electric motor which is nothing short of amazing. (Clearly this is of import well beyond computers.) The reason is that the scaling laws for power density are in our favor as we go down toward the nanometer realm. At macroscopic sizes, almost all electric motors are electromagnetic; at nano scales, they will be electrostatic. The motor is essentially a van de Graff generator run in reverse (but it works just fine as a generator, as do some macroscopic electric motors). The power density of the motor is over  $10^{15}$  W/m<sup>3</sup>; this corresponds to packing the power of a fanjet from a 747 into a cubic centimeter. (It's not clear what you'd do with it if you did, though!)

Ultimately, the ability to make small, powerful motors is going to be more important for nanorobots than nanocomputers per se. The speed advantage of electronics over mechanical logic is almost certain to drive the descent into nanocomputer design.

## Other logics for nanocomputers

Before going into other extensions of conventional digital logic, there is another

form of nanocomputer that may appear earlier for technological reasons. That's the molecular biocomputer.

Imagine that a DNA molecule is a tape, upon which is written 2 bits of information per base pair (the DNA molecule is a long string of adenine-thymine and guanine-cytosine pairs). Imagine, in particular, this to be the tape of a Turing machine, which is represented by some humongous clump of special-purpose enzymes that reads the "tape," changes state, replaces a base pair with a new one, and slides up and down the "tape." If one could design the enzyme clump using conventional molecular biology techniques (and each of the individual functions it needs to do are done somewhere, somehow, by some natural enzyme) you'd have a molecular computer.

## Other mechanical logics

Now, back to mechanical logic. Most macroscopic mechanical logic in the past has typically been based on rods that turned instead of sliding. It's reasonable to assume that similar designs could be implemented at the nano scale.

## Electronic logic

It's clear that quantum mechanics allows for mechanisms that capture a single electron and hold it reliably in one place. After all, that's what an atom is. Individual electrons doing specific, well-defined things under the laws of quantum mechanics, is what happens in typical chemical reactions. Clearly there is no basic physical law that prevents us from building nanocomputers that handle electrons as individual objects.

What is not so clear is how, specifically, they will work. Quantum mechanics is computationally very expensive to simulate, and intuitively harder to understand, than the essentially "physical object" models used in mechanical nanotechnology designs thus far. Indeed, the designs are typically larger and slower than they would have to be in reality, simply to avoid having to confront the analysis of quantum effects.

Ultimately, however, nanotechnologists will be "quantum mechanics." Computers based on quantum effects will be even smaller, more efficient, and much faster than mechanical ones of the type presented above. They will use much the same logical structure: it's quite possible to design retractile cascades even in conventional transistors (where it's an extension of techniques called "dry switching" in power electronics and "hot clocks" in VLSI design).

There are schemes, with some mathematical plausibility, to harness quantum state superposition for implicit parallel processing. In my humble opinion, these will require some conceptual breakthrough (or at the very least, significant experimental clarification) about the phenomenon of the collapse of the Schroedinger wavefunction before they can be harnessed by a buildable device. Keep your fingers crossed!

## Conclusion

Beyond certain rapidly approaching limits of size and speed, *any* computer must use logical reversibility to limit bit destruction. This is particularly true of nanocomputers with molecular-scale components, which if designed according to standard current-day irreversible techniques, explode.

We can design nanocomputers today which we are virtually certain would work if constructed. They use mechanical parts that are more than one atom but less than ten atoms across in a typical short dimension. The parts move at rates of up to ten billion times per second; processors built that way could be expected to run at rates of 1000 MIPS. Such a processor, and a megabyte of very fast memory, would fit in a cubic micron (the size of a bacterium). A gigabyte of somewhat slower memory would fit in another cubic micron. A pile of ten thousand such computers would be just large enough to see with the naked eye.

## FURTHER READING

Drexler, K. Eric: *Nanosystems: Molecular Machinery, Manufacturing, and Computation*, Wiley Interscience, New York, 1992

*Proceedings of the Physics of Computation Workshop*, Dallas, October 1993: IEEE Press, (in press). (particularly papers by Merkle, Hall, and Koller)

Hennessy, J.L. & Patterson, D.A.: *Computer Architecture: A Quantitative Approach*, Morgan Kaufmann, San Mateo, CA, 1990

Watson, Hopkins, Roberts, Steitz, & Weiner: *Molecular Biology of the Gene*, Benjamin/Cummings, Menlo Park, CA, 1987 (4th ed.)

Leff, Harvey S. and Andrew F. Rex: *Maxwell's Demon: Entropy, Information, Computing*, Princeton University Press, Princeton, NJ, 1990 (Particularly papers by Landauer and Bennett)

# EXTROPY INSTITUTE

## Why an Extropy Institute?

Fundamentally, two related reasons led us to form Extropy Institute (ExI) out of the persons and forces attracted by the intellectual gravitation of *Extropy* magazine. The first objective was to draw together people of shared values and goals, to act as focal point – a nexus – to facilitate their interaction, mutual aid, and exchange of information. I saw ExI as the nucleating agent which would crystallize a fresh, dynamic culture to sustain those who find themselves alienated from the surrounding society's bizarre religious, political, and intellectual practices, institutions, and beliefs.

I defined the Extropian Principles of Boundless Expansion, Self-Transformation, Dynamic Optimism, Intelligent Technology, and Spontaneous Order to give shape to the vision of life I shared with Tom Morrow, co-founder of *Extropy*. The Extropian culture grows rapidly as many others encounter a life-stance which explicitly interrelates their own values and goals, including political individualism and voluntarism, personal responsibility, rationality, skepticism, enthusiasm for technology, a desire for perpetual self-improvement, and an urge to overcome historical, cultural, biological, genetic, neurological, and spatial limits, including many that most humans rarely think to question, such as the process of aging and dying.

The second, closely related purpose of ExI, is educational. By gathering and focusing the energies of transhumanists all over this planet, we seek to shift cultural attitudes in an extropic direction, both by persuasion and by personal example. We will encourage the more traditional, uncritical, and timid members of our species to doubt both the desirability and inevitability of death and taxes, to outgrow irrationalist and mystical religions, and to challenge every limit to our potential. ExI promulgates a philosophy of life appropriate for rational persons rushing headlong into a future of boundless possibility, where many of the rules of existence will be changed.

## ExI Membership

Following incorporation in May 1992, ExI began to offer memberships in June 1992, and has been growing strongly since. Members receive *Extropy: The Journal of Transhumanist Thought* – the central source of extensive intellectual

exploration. In addition, members receive the bi-monthly newsletter, *Exponent* (edited by Simon! D. Levy). *Exponent* provides regular news updates, shorter writings, book, movie, and software reviews of Extropian interest, reports on crucial advances in science, and provides information on meetings and electronic gathering places.

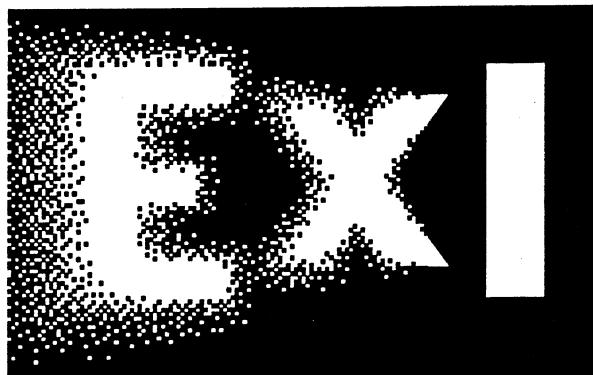
ExI members receive discounts, currently on T-shirts and audio tapes, and later on books, software, and events.

We are in the early information gathering stages of organizing the first International Extropy Institute Conference, likely to be held in 1994. The conference will draw together top-flight brains in many fields for a unique and intense round of lectures, panel discussions, and social events.

The Extropian network grows as individuals organize local meetings, both social and purposive. The first gatherings in Los Angeles in 1991, have been followed by meetings in New York, San Francisco, and Boston.

Members and non-members alike can dive into prolific discussions across the Internet (see back cover for details). The main Extropians list, managed by Harry Shapiro, is a torrent of fascinating discussion, its overwhelming volume now tamed by the addition of a digest version, thanks to Ray Cromwell. There are several email lists for local organization, and an essay list for extensive original essays. Thanks to the ease of communication granted by a modem and the Internet, a flourishing Extropian virtual community has arisen over the past 17 months since ExI member Perry Metzger originally created the Extropians list.

Further information on our other proposed projects can be found in the last issue of *Extropy*. The ambitious roster of



## Extropy Institute

projects found there should be understood as a prospectus for our future, not a description of our current status. Extropy Institute is yet a newborn, without wealthy parents. Yet ExI grows and develops vigorously. Join us, and fly upward and outward with us as we take charge and create the future we seek to inhabit.

Max More, Executive Director, ExI

## Ten Books

The dauntingly long list of books following the Extropian Principles 2.0 (see *Extropy* #9) may deter some from even starting to read these mind-expanding works. So, acting on a suggestion by Mark Plus, here are the ten books that I believe serve jointly to explain much of the Extropian conception of self and universe, in the past, present, and probable future.

Paul M. Churchland, *Matter and Consciousness*.

Richard Dawkins, *The Selfish Gene*.

K. Eric Drexler, *Engines of Creation*.

David Friedman, *The Machinery of Freedom*.

Hans Moravec, *Mind Children: The Future of Robot and Human Intelligence*.

Ed Regis, *Great Mambo Chicken and the Transhuman Condition*.

Julian L. Simon, *The Ultimate Resource*.

Robert Anton Wilson, *Prometheus Rising*.

Ayn Rand, *Atlas Shrugged*.

Marc Stiegler, *The Gentle Seduction*.

## Extropy Institute Directors:

Max More, Executive Director and CEO, Editor *Extropy*. [more@usc.edu](mailto:more@usc.edu)

Tom Morrow, Associate Executive Director. [twb3@midway.uchicago.edu](mailto:twb3@midway.uchicago.edu)

Simon! D. Levy, Editor, *Exponent*. [levy@yalehask.bitnet](mailto:levy@yalehask.bitnet)

David Krieger. [dkrieger@netcom.com](mailto:dkrieger@netcom.com)

Russell E. Whitaker. [whitaker@eternity.demon.co.uk](mailto:whitaker@eternity.demon.co.uk)

Ralph Whelan. [alcor@cup.portal.com](mailto:alcor@cup.portal.com) [not a private address]

# BACK ISSUES

## BACK ISSUE PRICES:

#1, 2, 4, 5, 6: \$4 each; #7, 8, 9: \$4.50 each. Available from:

**EXTROPY INSTITUTE, P.O. Box 57306, Los Angeles, CA 90057-0306. Tel/fax: 213-484-6383.**

### **#9, Vol.4 No.1 (Summer 1992):**

The Extropian Principles, 2.0, by Max More; Extropy Institute Launches, by Max More; Persons, Programs, and Uploading Consciousness, by David Ross; Nanotechnology and Faith, by J. Storrs Hall; The Making of a Small World (fiction), by R. Michael Perry; Genetic Algorithms, by Simon! D. Levy; Time Travel and Computing, by Hans Moravec; Futique Neologisms 3; Exercise and Longevity, by Fran Finney; The Transhuman Taste (Reviews): *The Anthropic Cosmological Principle*, *The Blind Watchmaker*, *The Ultimate Resource*, *Population Matters*, *The Resourceful Earth*, *Bionomics*

### **#8 Vol.3 No.2 (Winter 1991-92):**

Idea Futures: Encouraging an Honest Consensus, by Robin Hanson; Dynamic Optimism, by Max More; Neurocomputing 5: Artificial Life, by Simon! D. Levy; Futique Neologisms (futurist lexicon); Extropia: A Home for Our Hopes, by Tom Morrow; Human-Transhuman-Posthuman, by Max More; Reviews of: Stiegler's *David's Sling*, Drexler's *Unbounding the Future*, Platt's *The Silicon Man*; News of scientific advances and movement news; Reviews of zines.

### **#7 Vol.3 No.1 (Spring 1991):**

A Memetic Approach to 'Selling' Cryonics, H. Keith Henson & Arel Lucas; Privately Produced Law, Tom Morrow; Order Without Orderers, Max More; Futique Neologisms; Neurocomputing 4: Self-Organiza-

zation in Artificial Neural Networks, by Simon! D. Levy; Forum on Transhumanism; Reviews of *Smart Pills*, *Surely You're Joking Mr Feynman*, *Great Mambo Chicken and the Transhuman Condition*; and more...

### **#6 (Summer 1990):**

Transhumanism: Towards a Futurist Philosophy, by Max More; The Thermodynamics of Death, Michael C. Price; The Opening of the Transhuman Mind, by Mark Plus; The Extropian Principles, by Max More; Neurocomputing Part 3, by Simon! D. Levy; Forum on Arch-Anarchy and Deep Anarchy; Reviews: *Order Out of Chaos*, *The Emperor's New Mind*, *A Neurocomputational Perspective*, *Loompanics Greatest Hits*, *The Machinery of Freedom*; Extropian Resources, and more.

### **#5 (Winter 1990):**

Forum: Art and Communication; Leaping the Abyss, by Gregory Benford; Arch-Anarchy, by A; Deep Anarchy, by Max O'Connor; I am a Child, by Fred Chamberlain; Perceptrons (Neurocomputing 2), by Simon D. Levy; On Competition and Species Loss, by Max O'Connor; A Review of *Intoxication*, by Rob Michels; Intelligence at Work, by Max O'Connor and Simon D. Levy; Extropian Resources, by Max O'Connor and Tom W. Bell; The Extropian Declaration, by Tom W. Bell and Max O'Connor; Our Enemy, 'The State,' by Max O'Connor and Tom W. Bell

### **#4 (Summer 1989):**

Forum; In Praise of the Devil, by Max O'Connor; Neurocomputing, by Simon D. Levy; Why Monogamy? by Tom. W. Bell; What's Wrong With Death? by Max O'Connor; Reviews: *Are You a Transhuman?* Postscript to "Morality or Reality" by Max O'Connor; Efficient Aesthetics, by Tom. W. Bell; Intelligence at Work: Advances in Science by Max O'Connor

### **#2 (Winter 1989):**

Review of *Mind Children*, by Max O'Connor; Darwin's Difficulty, by H. Keith Henson and Arel Lucas; A Truly Instant Breakfast, by Steven B. Harris M.D.; Wisdomism, by Tom W. Bell; Nanotechnology News, by Max O'Connor; Weirdness Watch, by Mark E. Potts

### **#1 (Fall 1988):**

A brief overview of extropian philosophy and an introduction to some of the topics we plan to address: AI, Intelligence Increase Technologies, Immortalism, Nanotechnology, Spontaneous Orders, Psychochemicals, Extropic Psychology, Morality, Mindfucking, Space Colonization, Libertarian Economics and Politics, Memetics, and Aesthetics; "Morality or Reality," by Max O'Connor.

#3 (Spring 1989) is out of print.

## PICHARDO

Adolpho S. Pichardo Associates, Ltd. A Limited Partnership Chartered for Fiduciary Services

**SERVING AS TRUSTEES FOR REANIMATION TRUSTS SPECIFICALLY DESIGNED FOR SAFEGUARDING YOUR PROPERTY DURING THE PERIOD OF YOUR CRYONIC SUSPENSION**

**3850 Galloway Road, Suite 306, Miami, FL 33165 (305) 551-0991 FAX: (305) 551-3051**

# The Transhuman Taste

## Reviews of Extropian interest

### Beyond the poor man's Extropianism: A review of two books about Ayn Rand and Objectivism.

by Mark Plus

(A) *Atheism, Ayn Rand, and Other Heresies*, by George H. Smith. (Buffalo, New York: Prometheus Books, 1991). 324 pages. ISBN 0-87975-577-6.

(B) *The Ideas of Ayn Rand*, by Ronald E. Merrill. (La Salle, Illinois: Open Court, 1991). 191 pages. ISBN 0-8126-9157-1.

#### Introduction

The Extropian eupraxophy, although currently associated with relatively few individuals, is an example of a spontaneous order. Rather than being presented as a completed intellectual system in the manner of, say, Thomas Aquinas, it is instead the open-ended confluence of ideas from many individuals and disciplines. A "fountainhead" of much Extropian thinking has been the late novelist-philosopher Ayn Rand, who, although succumbing to entropy in 1982, sketched out a philosophy, called Objectivism, which is compatible with the Extropian principles of Boundless Expansion, Self-Transformation, Intelligent Technology, Spontaneous Order, and Dynamic Optimism. (Indeed, her novel *Atlas Shrugged* is on the recommended reading list for "The Extropian Principles v. 2.0.")

Unfortunately Rand never developed the full extropic implications of her system. She has left her followers to make do with a bug-ridden "Objectivism 1.0," a kind of poor man's Extropianism. Hence Objectivism, although still a powerful eupraxophy, lacks the vigor and appeal it would have if certain errors were corrected and many Extropian features, e.g. immortalism, were explicitly expressed.

Fortunately in recent years a number of independent Randian scholars have been working on versions of "Objectivism 2.0," which, although falling short of full Extropianism, demonstrate attempts to grapple with the realities of tomorrow. These efforts have helped to ground Objectivism in the thought of the past, while preparing for its growth and improvement in the future. As this review will make clear, Objectivism must engage in Boundless Expansion to maintain its intellectual power. The two books under review describe these revisions.

#### (A) Objectivism's Heritage: Thomism without "God" (the G-word)

George H. Smith is an independent scholar who moves comfortably in both freethought and libertarian circles. His earlier book, *Atheism: The Case Against God*, has been a steady seller at Prometheus Books for over a decade, while during the same time he has been writing and lecturing in advocacy of libertarian ideas.

In his new book, *Atheism, Ayn Rand, and Other Heresies*, Smith combines his interests by grouping both freethought and libertarian essays together around the theme of "heresy." As Smith explains in his Introduction, heresy (from Greek *hairesis*, "choice") was originally a morally neutral disagreement with someone else's "right belief" or orthodoxy. When one ideological group attains political power, however, its orthodoxy defines heresy as evil and corrupting, giving the orthodox both the means and the incentive to persecute heretics. It is only in the open society, where everyone can freely associate and delimit his or her respective orthodoxy, that the heretic can live without fear of physical sanctions, although there may be moral and social sanctions instead.

Though most of Smith's essays are both readable and thought-provoking, of relevance to this review are the three essays in the middle section of the book dealing with Ayn Rand. Rand was certainly a heretic, especially in her uncompromising secularism. In his essay "Atheism and Objectivism," Smith points out that Rand's atheism offends conservatives who might otherwise be open to her philosophy. According to the traditional religious critique of atheism, rejection of belief in the g-word leads to all sorts of theoretical and practical disasters.

Ironically, as Smith demonstrates, there is a grain of truth in this criticism, though the religionists are right for the wrong reasons. Many philosophies promote atheism, but at unacceptable costs.

After analyzing different approaches to atheism, Smith discusses the various kinds of epistemological atheism, one of which is Objectivism. Most of these – Humean skepticism, logical positivism, linguistic analysis – reject theism as nonsensical, but for the same reason also jettison objective ethics, certainty, causality, and metaphysical speculation.

Objectivism, like the other epistemological critiques of theism, requires that the g-word be defined in an intelligible way, and then places the onus of proof upon the theist. Since no intelligible, non-trivial definition of the g-word has ever been produced, and no argument for the existence of the g-word has withstood analysis, Objectivism accepts atheism as the only rational possibility.

Where Rand differs from other atheistic philosophers, however, is that she is an Aristotelian, with Aristotle's conception of philosophy as a necessity for human life. Thus her philosophy is equipped for worldly success, unlike the other "stripped down" secular philosophies of modernity. This seems rather odd, for the majority of Aristotelians on the scene today, the Thomists — the Catholic intellectual heirs of Thomas Aquinas — are opposed to secularism. As Smith writes, "With Objectivism and Thomism we have two philosophical movements claiming Aristotle as their intellectual ancestor, but that are on opposite sides of the religious spectrum." In other words, although Smith does not phrase it in this way, Objectivism is a kind of Thomism without the g-word.

Smith maintains that Rand's "rejection of God does not stem from the limitation or distrust of reason, rather, it is in the name of reason that she rejects faith, mysticism, and belief in the supernatural." He then concludes in this essay:

In short, the atheism of Ayn Rand is not destructive in the least. In rejecting God, Rand does not reject metaphysics, ethics, certainty, or the possibility of happiness. On the contrary, it is because Rand has so much positive value to offer that she considers atheism to be a comparatively minor issue.

Of course, from our perspective Rand would have offered a lot more if she had advocated—and practiced!—immortalism. This is a major defect of Objectivism 1.0. Nevertheless she deserves credit for promoting a secularism which is a radical

improvement over the dreary atheistic philosophies of Marxism, existentialism, and secular humanism.

In his second Randian essay, "Ayn Rand: Philosophy and Controversy," Smith develops the Rand / Aquinas parallel further. He commits his own heresy by arguing that, although Rand was not well read in philosophy, many of her epistemological and ethical arguments are similar to those made by modern Aristotelians such as the Thomists, while her political philosophy is clearly derived from classical liberalism. The components of much of Rand's thought may be borrowed from others, or else independently reinvented, but the way she puts them together into one system gives Objectivism its freshness and vitality. By identifying the ancestors and antecedents of Objectivism, Smith can place it firmly in the context of Western philosophy, while showing that it is not that far out of the mainstream. Rand may be a heretic, but she is one with a familiar genealogy.

After documenting these similarities, Smith speculates about the future of Objectivism. He considers the "official" school, led by Leonard Peikoff (the "Randian Grand Inquisitor") to be brain dead, and argues that the future lies in the work of independent, "neo-Randian" philosophers such as David Kelly, Tibor Machan, Douglas Rasmussen, and others. (I had considered reviewing Leonard Peikoff's book, *Objectivism: The Philosophy of Ayn Rand*, but it is not comparable to the other two in quality. Under Peikoff's care, dogmatic Objectivism, following the example of its cousin Thomism, is hardening into a dead scholasticism.) However, these neo-Randians may not get much closer to Extropianism than Rand if they continue to accept the limitations of the human envelope. The book reviewed in section B, written by a neo-Randian not mentioned by Smith, implies a way beyond the human condition.

Finally, in the third essay, "Objectivism as a Religion," Smith deplores the tendency of many Objectivists to turn Randian ethics into a system of rules one has to follow, regardless of the consequences to one's happiness. As he argues, a rule-based ethics is essentially religious, whereas Rand's ethics is based on the idea of ethical standards, which one chooses not out of fear or guilt, but rather out of intellectual conviction. One can obey or disobey a rule, but it makes no sense to obey or disobey a standard. In the Objectivist ethics one follows a standard as a guide for attaining the goal of personal happiness. As Smith concludes in this final Randian essay:

Whatever her errors, Ayn Rand struggled mightily against a religious view of morality, and she sought to place ethics on a rational foundation, free of any appeal to faith or force. Rand was a humanist in the best sense; for her, the happiness of the human being is the *summum bonum* of ethics.

Of course Rand, like her philosophical opponents, promoted a dechristianized concept of happiness. Beyond dechristianized eudaimonism lies the challenge of immortalist Dynamic Optimism.

## (B) Objectivism's Future: The Trans-Randian condition

If George H. Smith can show Objectivism's place in the thought of the past, Ronald E. Merrill suggests and implies what Objectivism has to do if it wishes to reach for the future, even to the threshold of transhumanism.

I had not heard of Merrill until I read his book *The Ideas of Ayn Rand*, but I can now state without reservation that it is the most fascinating study of Rand's thinking I have yet encountered. Indeed, there is so much material in the book that I have trouble selecting what I want to review.

Like Smith, Merrill wishes to locate Rand and Objectivism in their historical context. Unlike the recent biographies of Rand, however, Merrill avoids the prurient details, and concentrates instead on the *ideas* of Ayn Rand, as the book's title says. He writes that he became a "Randian" at fifteen after reading *Atlas Shrugged*, and was associated with the Nathaniel Branden Institute (NBI) for awhile back in the 1960's. Although he met some eccentric people, he never encountered the cape-wearing Rand cultists other critics of the NBI period have described.

The break between Rand and Branden in 1968 destroyed what could have otherwise become a powerful force for reason and freedom, but Rand nevertheless continues to exert a covert influence on American thinking. Her novels still sell in the hundreds of thousands annually (in what is otherwise an increasingly illiterate society), while many of her ideas, advocated without proper acknowledgement through "conservative" ideologues, are becoming more widely accepted. (For example, the "conservative" broadcaster Rush Limbaugh sounds Randian much of the time, except when he starts mouthing off about the g-word. As Smith explains in his book, Rand's atheism is an impediment to a general recognition of her legitimacy.) Her on-going popularity suggests

that Rand's vision is timeless, a vision which Merrill tries to explain.

In Chapters 2 through 5, where Merrill discusses Rand's development as a writer, he makes it clear that he considers her to be a deep and complex literary artist. As she struggled to master English prose while trying to survive in her adopted country, Rand also struggled with and eventually overcame her devotion to the philosophy of Friedrich Nietzsche.

Although according to the "official" history Rand flirted with Nietzsche briefly as a young woman before inventing Objectivism, Merrill documents how Rand was still a Nietzschean in the 1930's when she wrote the first version of *We the Living*. Comparing the original novel with the version Rand "revised" in the 1950's shows that she felt embarrassed by the Nietzschean dialogue she had written in the 1930's. It was only in *The Fountainhead*, where Rand contrasted the Objectivist Howard Roark with the Nietzscheans Dominique Francon and Gail Wynand, that Rand was able to show her break with Nietzsche. By this time Rand had decided that true mastery over one's circumstances came not from wielding the whip over the common people, as Nietzsche taught and Wynand practiced, but rather from self-discipline exercised through intellectual and economic productivity, as shown in the life of Roark.

Finally in *Atlas Shrugged* Rand dramatizes the outlines of her mature philosophy. Merrill argues that Rand's critics have misrepresented or misinterpreted many of that novel's controversial passages (the train tunnel explosion, for example), and that a full exegesis of all of *Atlas Shrugged*'s plots, ideas, psychological insights, and symbolism has still to be done. (He describes one theory that Galt's strike is based on a Talmudic exegesis of the biblical story of the destruction of Sodom, whose inhabitants were guilty not of perversion, but of collectivism. Rand did come from a Jewish family, but I have seen no hard evidence that she ever used her heritage as a source of literary allusions.) Clearly there is more information encoded in *Atlas Shrugged* than either its admirers or its detractors have yet extricated.

After analyzing Rand's literary development, in Chapter 6 Merrill comes the meat of his subject, Rand's formal philosophical ideas. As Merrill argues, Rand carried on the tradition of the Greek sophists, the market intellectuals who sold their wisdom to all comers. This placed her in direct opposition to the Socratic tradition in philosophy, for Rand took an anti-skeptical and anti-relativistic approach to the fundamental issues. Philoso-

phy is too important to be left as a game for the intellectual elite.

Starting with the essentials of Aristotle's metaphysics, Rand assumes that "existence is identity," thus relieving her of the burden of explaining ultimate origins. Her metaphysics form the "boundary conditions" of the rest of her philosophy, to borrow a metaphor from mathematics, which determine everything else that follows. From there Merrill describes how Rand's theory of concepts, building blocks of her epistemology, emphasizes understanding over proof. This has the unfortunate consequence of leaving us without epistemological guidance in a lot of areas – including, as it will be shown, in Rand's theory of ethics. Merrill laments that Rand never directed her mind towards the epistemological paradoxes raised by general relativity and quantum mechanics. (Since Merrill seems to be well read in both science and philosophy, why doesn't he try to solve these puzzles?) And as an example of the epistemological wilderness in which Rand has left us, he writes:

It not infrequently occurs that we must choose between two theories. One theory, call it A, is internally self-consistent, but there are some experimental facts that contradict it. The other theory, call it B, is consistent with all the facts but contains internal contradictions. No better theory is currently available, but one must make decisions, and right now – to design a spacecraft, to plan a campaign against a deadly epidemic, to prevent an explosion in a refinery. One must act, and on the basis of one theory or the other. Which should be chosen, A or B? This is an epistemological problem. The solution is left as an exercise for the reader.

From an analysis of Rand's epistemology Merrill moves on to his attempt to state and then reformulate Rand's derivation of ethics ("ought" statements) from the facts of human nature ("is" statements). Ironically the above thought-experiment bears directly on one weakness in Merrill's reformulation, though he does not seem to be aware of it.

Merrill paraphrases the "Randian Argument" as follows:

- i. Living beings, and only living beings, have values (goals).
- ii. Man, being volitional, must choose his values.
- iii. Values – goals – may be means to an end, but must lead to some ultimate end. An infinite chain of means leading to no final end would be meaningless and

impossible.

iv. Life is an ultimate end, and furthermore it is the only possible ultimate end, the only "end in itself."

v. Therefore, the only meaningful or justifiable values a man can choose are those which serve to sustain his life.

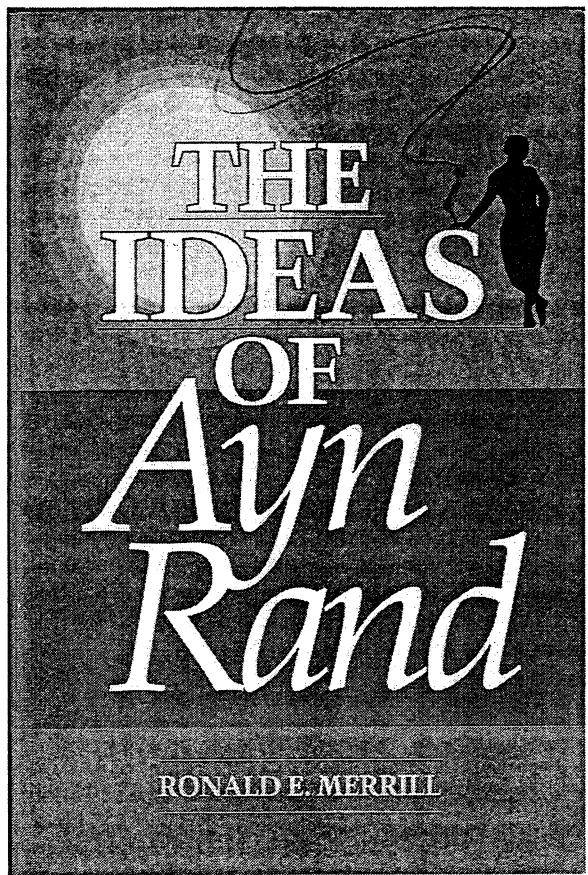
The sticking point in this argument is the fourth premise, involving the concept of "ends in themselves." Merrill better defines this concept by quoting from other Objectivist writings to the effect that life is an end in itself in that it "is an ordered collection of activities, which are means to achieving an end, which is — simply those activities." He also disposes of the problem of competing ends in themselves by arguing that life is the necessary prerequisite to all other ends, so that it is the most important end in itself, regardless of whether ends in themselves exist.

Next Merrill shows how to derive ethical statements from the facts of reality by equating normative ought statements – e.g., "You ought always to tell the truth." – with operational ought statements – e.g., "You ought to format a new disk before attempting to write a file to it." – in a manner analogous to Einstein's equation of gravity with acceleration in General Relativity. If one accepts this approach, then "ethics reduces to a matter of engineering" and "Objectivist ethics ... can provide a prescription for any specified moral dilemma."

Unfortunately Merrill's reformulation of the Randian Argument suffers from its own problems. First of all, if one accepts the proffered definition of life as an end in itself, the fact remains that the activities of life lose their efficiency, and eventually break down, due to aging. It cannot be argued that aging is one of the activities of life one engages in today in order to repeat those activities tomorrow. Aging eventually destroys one's ability to engage in the activities of life. Hence one's life is a depreciating ethical standard unless the aging process could be prevented. But neither Rand nor Merrill derives immortalism from this argument.

Second, if Rand, as interpreted by Merrill, really has turned ethics into a "matter of engineering," then the burden in decision-making is shifted from ethics onto epistemology – which by Merrill's own admission is the weak link in Objectivism! The thought-experiment I quoted above is perhaps an ethical problem which Objectivism cannot obviously solve.

Fortunately Merrill provides himself an escape from the first dilemma by presenting life not as one state of a binary condition (the other state being death), but rather as a continuum of value-seek-



ing. In a passage that may express ideas familiar to many cryonicists, Merrill writes:

If life is defined in terms of an organism's exhibition of goal-directed behavior then we must visualize the possibility that it can exist on a multitude of levels. On what we might call the "hamburger" level, cellular behavior is very primitive, consisting of a small set of rather simple tropisms. The "warm body" level, in which all or at least some organs are functioning, represents a higher level of complexity. When the organism – a human being, in the case we are concerned with – is able to perceive and to act, then the range of goals accessible to him expands enormously, and it is only at this level that we would regard him as "completely alive." And yet, need we stop here? If our subject is more alert, more intelligent, more healthy, more strong, is he not more able to select goals and pursue them, and is he therefore not more alive yet?

Merrill does not delimit the upper range of this continuum, but the implication is that it is open-ended. Nor does he draw the obvious inference, so I will: The Objectivist ethics pushes us in the direction of increasing the value of our lives without limit, which implies transform-

ing ourselves into immortal superhumans. Merrill himself acknowledges that the Objectivist virtue of pride requires the Objectivist to "make a commitment to self-education and self-improvement, to the constant expansion of his competence to deal with reality" (thereby approaching the Extropian principles of Self-Transformation and Boundless Expansion). The irony of Merrill's analysis is that he fails to see that an unaging transhuman would meet the Objectivist standard of ethical excellence better than an ordinary human.

Beyond the Objectivist ethics lies its practical application in politics, described in Chapter 7. Merrill analyzes Rand's philosophical differences with both conservatism and libertarianism. Her critique of conservatism is on the mark, for conservatives are only half-way committed to the cause of freedom; as soon as they assume

"power," they begin to compromise their principles for the sake of popularity. Despite the fact that the Republican Party has had its way with the Executive Branch of the U.S. Government for twelve years, it has not stopped the progress towards socialism in this country.

On the other hand, Rand's critique of libertarianism is problematic, for as Merrill complains, she never explained her view of the Objectivist Just State. Part of the problem is that "libertarianism" is a catch-all term for a range of philosophies united around the assumption that the State is inherently evil. As a consequence, the only way the adherents of these otherwise disparate philosophies can work together is under the banner of moral tolerance.

As Merrill sees it, libertarian moral tolerance contributes to the dispute about the nature of the social contract. The anarcho-capitalist school of libertarianism, which "defends the indefensible," argues that the Lockean view of the social contract — where the individual gives his "free" consent to be governed by the State — is really a sham, for the State controlling any geographic area is a monopoly and is able to dictate its terms by force. Hence true freedom can be found only in a competitive market for the services that the State traditionally offers.

Rand's view, on the other hand, is that (a) political freedom needs to be based

on an objective standard of morality, and (b) "the Objectivist Just State is to merit its authority because it implements objective, knowable moral principles." Because of her assumption that there can be no conflict of interests among rational beings, Rand therefore seems to argue that one's entry into an ethical social contract is "pure profit." As Merrill writes, "In giving up the chance to live as a predator, the individual is losing nothing."

There are serious difficulties with the Randian view, however, because humans make decisions based mostly on emotions and naive or neurologically hardwired heuristic procedures, and rarely based on explicit epistemological principles — and we have just seen a major weakness in Rand's epistemology. For evolutionary reasons humans currently cannot sustain lifelong, consistent rationality. It will take technological self-transformation of the human brain, along with progress in epistemology, before any hypothetical society of rational individuals becomes possible. (In which case one might speak of a "society of rational transhumans.") Under the limitations of the human condition, a society based on anarcho-capitalism might be an improvement over ones based on statism, but such a society would still have problems caused by the discoordination of the human brain. Rand's ignorance of the modular-mind perspective leaves her philosophy vulnerable to attack as a form of rationalism rather than as a systematization of rationality.

Where Merrill and Rand are both right, however, is in their emphasis on objective morality as an antidote to the growth of the State. Merrill argues that "libertarianism has failed by its success" in eroding social controls over shiftless behavior. Its doctrines of ethical subjectivism and "doing one's own thing" have produced more anomie than anarcho-capitalism, with the result that the law-abiding people are calling on the State to protect their families from an ever-worsening breakdown in social order. "Moral tolerance" has reached bankruptcy when self-proclaimed "community leaders" demand amnesty towards thugs arrested for arson, looting, and beating up innocent bystanders on live television.

Only a people who can maintain a sense of "doing the right thing," as Albert Jay Nock phrased it, can live in a sustainably free society. Because of the collapse of morality, "the defenders of the indefensible face the political backlash," as suggested by the recent Republican National Convention. This reaction would not be called for if people, like the inhabitants of Galt's Gulch, "did the right thing" in the absence of the State's coercive

supervision.

Fortunately Merrill ends his book on a positive note. He calls for the re-creation of the Objectivist social network that existed before the abolition of NBI, and recommends that Objectivists study new areas of knowledge, such as sociobiology and chaos theory, to add "smart weapons" to their intellectual ammunition: "Technology has changed, the economy has changed, society has changed — and our vision of the organization of the free society must change too." And: "Half the genius of Ayn Rand was that she saw what everybody saw, and said what nobody dared to say. We need once again to become intellectual leaders, to have the courage to approach the cutting edge of new thought." (Merrill seems to be looking once again for Self-Transformation and Boundless Expansion.)

Along with networking, education, and setting good examples, the Objectivists' most powerful tool is Rand's concept of "the sanction of the victim." Evil survives by drawing its strength from the good, so withholding the sanction is one of the keys to victory in the long term. (Merrill's example of a sanction of the victim: "When businesses give money to universities and foundations that attack capitalism, ... the moral force of their example far outweighs the financial assistance. Whenever and however we may choose to fight, this kind of issue is the where.") However, Objectivist leaders ought not to "excommunicate" freethinking neo-Randians who disagree on peripheral issues, for Rand's, and now Peikoff's, habit of condemning dissenting Objectivists has caused more harm than good. Merrill then expresses his hope that Rand will be remembered by history as one of the greatest thinkers of our century, with our views of the just society and the function of philosophy changed in her favor.

Clearly the writings of George H. Smith and Ronald E. Merrill represent progress in upgrading the Objectivist worldview into something I call "Objectivism 2.0." Merrill, especially, is a "trans-Randian" groping towards something like Extropianism. His kind of boundlessly expansive Objectivist thinking needs to be nurtured and encouraged to counter the dogmatism of the Peikoff school. Both of the books I have reviewed are worth adding to one's Extropian library, but Merrill's is the better value of the two.

## Nanosystems: Molecular Machinery, Manufacturing, and Computation

By K. Eric Drexler

New York: Wiley-Interscience, 1992; 556 pp. Hardcover \$42.95, ISBN 0-471-57547-X Paperback \$22.95, ISBN 0-471-57518-6

Reviewed by J. Storrs Hall

In 1610, Galileo published a small book entitled *Siderius Nuncius* which set the world on its ear. It was, in many respects, the keynote publication of the 17th century. In it he described what happened when he took an existing tool, the telescope, and turned it on the realm of the heavens. He had discovered mountains on the moon, and moons around Jupiter. The old conceptions of the universe were, at that point, doomed, and the following century would see a complete restructuring of our basic understandings of nature, and kick the Renaissance into high gear.

In *Nanosystems*, Drexler takes the tool of mechanical engineering and turns it on the realm of the molecular. *Nanosystems*, in my humble opinion, is destined to be the keynote publication of the 21st century. There will not be a person alive in 2100 who is completely unaffected by the consequences of the ideas presented here.

Suppose you had a mechanical manipulator so small, and so precise, that it could handle individual molecules, maybe even atoms? You could build (small) objects with a precision that would allow you to specify in your design where each individual atom was to be, exactly which atoms would be covalently bonded to which other atoms, and so forth. If indeed you could build objects with this degree of precision, you could design and build some rather incredible things — like the mechanical manipulator we assumed in the first place.

If we could bootstrap ourselves into such a technology, we could gain essentially atomic-level control over the structure of matter for a multitude of purposes. Such a technology is called nanotechnology or, to distinguish it from more conventional approaches to nanometer-scale operations, molecular nanotechnology.

Drexler's previous books, *Engines of Creation* and *Unbounding the Future*, gave a glimpse of what molecular nanotechnology might mean. Indeed "*Engines*" was one of the major factors behind the flourishing of the cryonics movement in the mid-80's. However,

these books were written for the popular audience. *Nanosystems* is written for the technical audience.

Even so, much of *Nanosystems* is introductory or explanatory compared to the average technical book. The reason for this is clear: most technical books are written for the specialist in a particular field, and there are virtually no specialists in this one.

The burden of *Nanosystems* is simple in overall concept: it is a proof that molecular nanotechnology can work. As such it has two major parts.

First Drexler develops and justifies the theoretical tools he needs to analyze nanomechanical devices. This involves a unification of principles from physics, chemistry, and engineering. He explores the various levels of abstraction at which physical laws are used, from quantum mechanics to the continuum models of engineering which ignore the molecular nature of matter entirely. Of particular interest are the empirical molecular mechanics models, which seem to be the appropriate tradeoff between accuracy and efficiency for analyzing nanomechanical designs.

Second, he exhibits actual designs, with analyses, sufficient to make a strong case that a full-fledged, self-reproducing, molecular technology could be built from the devices presented and others like them. We have axles and bearings, gears and cams, pumps, motors, generators, and computers. There are chapters on the internal processes of, and the overall structure of, molecular factories. There is a detailed design for a robot arm 100 nanometers long.

There is a point of some importance that is generally ignored in higher-level discussions of nanotechnology. It forms the bridge between the first and second sections. If you had the manipulator arm with the ability to put each atom where you wanted it, would that really give you the ability to build objects with them? Surely the reactive atom you were trying to add to your workpiece would bond more quickly to the manipulator itself?

Of course, adding one atom at a time, like laying bricks, is not what is actually envisioned in the synthesis operations. What must actually be done is quite a bit more complex and looks a lot more like chemistry. The explanation of this "mechanosynthesis" is the "missing link" in the common understanding of nanotechnology; but it's not missing here. It's chapter 8.

Anyone with experience in the construction of macroscopic objects is aware that the process typically involves a plethora of tools and clamps and jigs and molds and scaffolding. So too in molecular construction: we might have a complex molecule that will undergo a chemical reaction if held to a certain kind of spot on a surface, adding one or two atoms in a certain configuration. We might build up the structure we want by holding a succession of the "tool" molecules to the workpiece in a careful pattern, recycling the depleted byproduct molecules to be recharged by conventional chemical means.

One example of molecular scaffolding is the relatively recent ability to grow diamond. This depends on the presence of hydrogen, without which the surface of the diamond is unstable and graphite is formed instead. Clearly nanomechanical synthesis of diamondoid structures would use the same kind of scaffolding!

The final chapter is not part of the logical development of the "proof of nanotechnology" but is just as important: the question of how we get there from here. There is a host of pathways. Drexler sketches a few of them here; but it is the very multiplicity of options that makes the ultimate goal relatively certain, and thus the book as a whole so important.

*Nanosystems* is intended to be accessible to the interested technical reader from any field, providing the person is willing to do enough brainwork to assimilate the concepts. If you've read and largely understood Asimov's *Intelligent Man's Guide to Science*, you'll find *Nanosystems* comprehensible; however, the more work you put into it, the more you'll be rewarded. I spent half a year working out, in my own specialty, some of the implications from just a few pages in the nanocomputers chapter (from a pre-publication draft).

In the years following 1610, *Siderius Nuncius* was highly sought after. People adopted Galileo's methods, and new discoveries came thick and fast: Spots on the sun; phases of Venus; rings around Saturn.

K. Eric Drexler

# Nanosystems

Molecular  
Machinery,  
Manufacturing,  
and Computation



With the publication of *Nanosystems*, our old bulk technologies are doomed. The following century should see a complete restructuring, not only of manufacturing, but of our basic relationship to nature. A new Renaissance is the least we should expect.

## Genius - The Life and Science of Richard Feynman

by James Gleick

Published by Pantheon Books, 1992, a division of Random House, Inc., New York  
ISBN: 0-679-40836-3

Reviewed by Harry Shapiro

No scientist has ever captured my imagination more than Richard Feynman; Certainly I am not alone. James Gleick, the author of *Chaos*, has tried to provide a more detailed, accurate and scholarly view of Richard Feynman. Having previously read, *Surely you're joking Mr. Feynman*, and *What do you care what other people think?* [1] I thought I knew him already. Gleick has clearly tried to find and present the Richard Feynman that isn't revealed by the humorous volumes mentioned above. "He penetrates beyond the gleeful showman depicted in Feynman's own memoirs and reveals a darker Feynman:

his ambition, his periods of despair and uncertainty..." [2]

I have made a practice in reviewing books for *Extropy* to focus not only on the book but how it directly relates to Extropian philosophy. I will continue this practice here. I doubt that any Extropian minded individual has any doubt about how the science of Richard Feynman relates to the Extropian view. Feynman is regarded as the "intellectual father" of nanotechnology because of his famous speech arguing that "there's plenty of room at the bottom." [3] I was delighted to learn about other areas in which

he seems to me to directly reveal an Extropian outlook. I will discuss this below. Overall I was both satisfied and greatly disappointed with this book.

Gleick clearly was trying to present the whole man; he succeeds in many ways. He reveals Richard Feynman as an explorer of self, and nature. Many details are provided about the major and key efforts that Feynman made when working on the Manhattan Project—details that Feynman left out of his own books. Neither are personal details left out, including facts about his sexual relationships, and use of psychoactive chemicals (marijuana and LSD [4]). At the same time, other details seem left completely out. In Feynman's books, for example, he spends much time talking about his own paintings. This is barely mentioned by Gleick. The lack of details about his painting seems surprising because Gleick does expend much effort writing about Feynman's ability to visualize. I would have enjoyed a critical look at the paintings as much as the looks Gleick provides of Feynman's other "art work," "the Feynman Diagrams."

Some of the disappointment I found in the book relates directly to how Gleick

presented Feynman's various scientific discoveries. It wasn't enough to talk about each topic (from particles and waves to DNA). Some detail of previous discoveries was clearly called for. Gleick not only provided these details but what amounted to an oral history of modern physics, and, almost, a primer in modern physics. I am not a physicist, though I have taken courses in it. I found Gleick's long, often several page descriptions, background details, and philosophical overviews to either tell me what I already knew or didn't really tell me enough. I suspect, that except for a select group of readers these parts of the book will be seen as rather boring, though Gleick was true to the book's sub-title – "The ... Science of Richard Feynman."

The most interesting parts of the book were for me, as indicated above, those details that seem to scream out: "Richard Feynman was an Extropian!" This included a rejection of religion, and a seemingly libertarian view of economics (page 397). I will present a few of these below.

One of the famous, or perhaps infamous, topics on the Extropian Mailing list [5] is the uploading of human consciousness into computer-like devices, and related topics such as the nature of consciousness. On page 124 Gleick details a cross-examination of philosopher Adolph Gruenbaum by Feynman who was identified as "Mr. X."

Gruenbaum: I want to say that there is a difference between a conscious thing and an unconscious thing.

Mr. X: What is the difference?

Gruenbaum: Well ... I would not be worried if a computer is unemployed,... I would worry about the sorrows which that human being experiences in virtue of conceptualized self-awareness.

Mr. X: Are dogs conscious?

Gruenbaum: Well, yes. It is going to be a question of degree. But I wonder whether they have conceptualized self-awareness.

Mr. X: Are cockroaches conscious?

Gruenbaum: Well, I don't know about the nervous system of the cockroach.

Mr. X: Well, they don't suffer from unemployment.

Gleick writes, "It seemed to Feynman that a robust conception of "now" ought not to depend on murky notions of mentalism. The minds of humans are manifestations of physical law, too, he pointed

out. Whatever hidden brain machinery created Gruenbaum's coming into being must have to do with a correlation between events in two regions of space—the one inside the cranium and the other elsewhere 'on the space-time diagram.' In theory one should be able to create a feeling of nowness in a sufficiently elaborate machine," said Mr. X."

Another topic that appeared on the Extropian list is the issue of how small computers could actually be made. Someone wondered if they could be reduced below the atomic level. On page 435 Gleick writes, "He joined two Caltech authorities on computation, John Hopfield and Carver Mead, in constructing a course on issues from brain analogues and pattern recognition to error correction and uncomputability. For several summers he worked with the founders of Thinking Machines Corporation, near MIT, creating a radical approach to parallel processing; he served as a high-class technician, applying differential equations to circuit diagrams... And he began to produce maverick research at the intersection of computing and physics: on how small computers could be; on entropy and the uncertainty principle in computing; on simulating quantum physics and probabilistic behavior; and on the possibility of building a quantum-mechanical computer, with packets of spin waves roaming ballistically back and forth through the logic gates."

Of course computers were not new to Feynman. He had worked with the primitive card-driven mechanical calculators at Los Alamos. On page 201 Gleick writes, "His computing team had put everything aside to concentrate on one final problem: the likely energy of the device to be exploded a few weeks hence at Alamogordo in the first and only trial of the atomic bomb.... He had invented a system for sending three problems through the machine simultaneously. In the annals of computing this was an ancestor to what would later be called parallel processing or pipelining. He made sure that the component operations of an ongoing computation were standardized, so that they could be used with only slight variations in different computations... He also invented an efficient technique for correcting errors without halting a run."

There are many other examples to choose from. I have chosen these because they seem to stick out in my mind. Examples of pushing technology to the edge of understanding, examples of innovation and the optimism to forge ahead in the face of adversity. Adversity in the nature of our universe—Feynman writes, "If it turns out there is a simple ultimate

law which explains everything, so be it—that would be very nice to discover. If it turns out it's like an onion with millions of layers ... then that the way it is." [6] Also adversity on more a more personal level—Feynman's first wife Arline was dying throughout his time at Los Alamos—he had the strength and courage to be with her emotionally and physically almost every weekend, while having the concentration to continue his work on "the bomb." He was dying of cancer during his work on the independent commission investigating the Challenger explosion. Again adversity was met and overcome, while Feynman was once again able to capture the attention of the nation.

If there is anything non-Extropian about him it is that he accepted his own death. However, he didn't retreat into the mythos of religion: "I can live with doubt and uncertainty and not knowing... I don't feel frightened by not knowing things, by being lost in a universe without any purpose, which is the way it really is as far as I can tell. It doesn't frighten me." [7]

In conclusion, I will say that for anyone who wants to have a deeper insight into Richard Feynman, *Genius*, is required reading. However the book does not stand alone. For readers who have a detailed knowledge of physics, or for those who don't want one, the book may very well prove disappointing. If you are looking to introduce someone to both Feynman and the science he espoused, then the book comes highly recommended. I found the book hard to put down, but did on occasion, yet I am glad I have read it. I would urge Mr. Gleick to write more about science, using people as backdrops rather than using a biography to teach science.

[1] Written by Ralph Leighton as told by Richard Feynman.

[2] From the book jacket to *Genius*

[3] A speech presented at the 1959 Annual meeting of the American Physical Society and later reprinted in Caltech's *Engineering and Science* magazine and *Popular Science Monthly*. (See *Genius* pages 17, 354-355)

[4] His use of LSD came as a revelation to me. I thought I remembered reading that he didn't use it for fear of disturbing his thought processes.

[5] You can join the Extropians list by sending e-mail to extropians-request@gnu.ai.mit.edu. Prepare to be besieged by many messages per day.

[6] Page 432

[7] Page 438

## CLASSIFIED ADS

Spread your Extropian memes with "Neural Tweaker" greeting card themes!

For FREE catalog send SASE to: Neural Tweakers, c/o Kevin Brown, P.O. Box 781, Netcong, NJ 07857, USA.

Cross verbal swords in *The Libertarian Connection*, open-forum magazine since 1968. Subscribers may insert one page/issue free, unedited. Lots of stimulating conversation. Eight issues (one year) \$20. Strauss, Box 3343X, Fairfax, VA 22038.

### Forward! Upward! Outward! Into Your T-Shirt!

ExI proudly announces, for the first time ever, an opportunity to boldly show what no one has shown before. While others are wearing their hearts on their sleeves, you'll be wearing your ideas on your torso. That's right, official ExI T-shirts are now for sale. These shirts, in blue, feature the five-spiral ExI logo in black and gold up front, and the motto "Forward! Upward! Outward! Into the Future!" on the back. The cost is \$16 per shirt (\$14 for ExI members). Sizes are men's large and extra-large. Send check or money order, in US dollars drawn on a US bank, payable to Extropy Institute, Dept. S, PO Box 57306, Los Angeles, CA 90057-0306.

## CONTRIBUTORS

**Hal Finney:** Hal received a BS in Engineering from Caltech. He developed several commercial video games for Mattel, then went on to head the Operating Systems Group at Ametek, a major manufacturer of parallel computers. He co-developed Neuralyst, a commercial neural-network simulation package. Hal played a major part in the development of version 2.0 of Phil Zimmermann's encryption program, PGP.

hal@alumni.caltech.edu

**J. Storrs Hall:** Moderator of the USENET sci.nanotech newsgroup and *Extropy*'s Nanotechnology Editor.  
josh@cs.rutgers.edu

**David Krieger:** David is Marketplace Administrator at the American Information Exchange (AMiX), *Extropy*'s Science Editor, an Extropy Institute director, and a former Technical Consultant to *Star Trek: The Next Generation*.  
dkrieger@netcom.com

**Hans Moravec:** Director of the Mobile Robot Laboratory at Carnegie Mellon University, Hans is author of *Mind Children: The Future of Robot and Human Intelligence*.  
moravec@cs.cmu.edu

**Max More:** Editor of *Extropy* and Executive Director of Extropy Institute, Max received his BA in Philosophy, Politics, and Economics from St. Anne's College, Oxford University in 1987, and now is writing his Ph.D dissertation on *The Diachronic Self: Identity, Continuity, and Transformation* at the Philosophy Department, University of Southern California.  
more@usc.edu

**Mark Plus:** Editor of *Venturist Monthly News* and Vice President of the Society for Venturism, Mark is manager of a small business. He hopes to have an e-mail address by next issue. PO Box 458, Wrightwood, CA 92397.

**Harry Shapiro:** Manager of the Extropians e-mail list, Harry is Manager of Computer Services for Warwick, Baker, Fiori, has consulted at Bell Laboratories, and worked in technical positions at E.F. Hutton and Shearson Lehman Brothers – the brokerage division of American Express.  
habbs@panix.com

*Extropy* #10 was produced on a Gateway 486-50DX2 PC with 8Mb of RAM and a 200Mb hard disk, running Windows 3.1, and using Aldus Pagemaker 4.0, WinWord 2.0, and Aldus Freehand 3.1.

Print run this issue: 2500

Printed by Hollywood Stationery & Printing, Los Angeles, California.

## ExI Audio Tapes For Sale

### Everything is Getting Better and Better – I'll Bet On It! by Julian Simon.

Economist Julian Simon uses hard data to counter prevailing gloomy beliefs about the current state of the world and its direction. Practically all measures of human well-being substantiate the Extropian's dynamically optimistic views: Life does tend to improve, though only through the efforts and applied intelligence of free persons. This tape makes an effective introduction to Simon's ideas, and, lent out to your pessimistic friends, will serve as a valuable intellectual catalyst.

\$10.95 (Members \$9.95)      EC1 (1-hour audio)

### Bionomics On Trial: A Discussion With Michael Rothschild

Rothschild outlines the main contentions of his book *Bionomics: Economy As Ecosystem*, and responds to audience questions. Topics discussed include electronic ecosystems; how Bionomics effectively draws 'liberals' into support for free markets; the relation between Austrian/process economics and bionomics; the role of government; how far the economy as ecosystem analogy can be pushed.

\$12.95 (Members: \$11.50)      EC2 (80 minutes audio)

### Spontaneous Orders and Future Evolution

by Max More

In this talk, originally delivered to the Albert Jay Nock Forum, Extropy Institute CEO Max More poses and answers the question: "How is humankind to evolve, transform, progress, to change its nature, without the oppression of central control and social engineering?" More distinguishes the spontaneous order embodied in free markets from the constructed orders that are more visible to most people. More shows why spontaneous orders are underappreciated, analyzes the problems endemic in central control, and shows the diverse areas in which spontaneous ordering principles reveal themselves, including evolutionary theory, memetics, ecology, agoristic computing, neural networks, genetic algorithms, and computational ecosystems. Finally, methods of transforming human nature while respecting the constraints of spontaneous order are proposed.

\$10.95 (Members \$9.95)      PH1 (1-hour audio)

### Recreational Drugs and Smart Drugs: Paternalism and Responsibility

by Max More

An analysis of the drug policies of the DEA and FDA, beginning with arguments for decriminalization of recreational drugs. The FDA comes under fire, with a history of that agency's increasing powers and growing safety fascism, the FDA's effects on death rates and pharmaceutical innovation, and the motivations facing FDA bureaucrats. More goes on to explain how smart drugs differ from other drugs, and how the FDA has mounted a war on the freedom to use intelligence-enhancing substances. The policies of the FDA and DEA are accelerating the infantilization of the American people, undercutting their self-determination and denying their personal judgment. This tape is crammed with the facts, statistics, and arguments needed to blast the FDA and DEA out of the water.

\$10.95 (Members \$9.95)      PL2 (1-hour audio)

**Postage: \$1 per tape. Overseas orders: Surface mail—\$1.50 first tape, \$1.25 each additional tape. Contact ExI for airmail rates. Check or money order in US dollars drawn on a US bank, payable to "Extropy Institute." Mail your order to: Extropy Institute, Dept. S, P.O. Box 57306, Los Angeles, CA 90057-0306.**

### How to join the ExI Virtual Community

Extropy Institute now sponsors, through the generous efforts of ExI member and List Manager Harry Shapiro, a number of electronic fora for sharing libertarian, free-market, life-extensionist, and other Extropian ideas with bright, like-minded individuals around the globe. All of these services reside at gnu.ai.mit.edu, so the string @gnu.ai.mit.edu should be added to the boldface addresses below when sending requests.

The most popular service is the Extropian e-mail list, which contains lively discussion and debate on numerous topics. To join, send a request to **extropians-request**. This list also has a recently-added digest feature which, instead of sending each message to you as it comes in, batches a group of messages into a segment of about 50K. That means getting one to three messages a day. To join, send a message to **exi-daily-request**.

A more recent service is the ExI Essay list, for posting longer, more carefully prepared electronic manuscripts. To get on this list, send a message to **exi-essay-request**.

There is also an Extropian conference on the Well, one of the longest-running professionally run BBS systems. On the Well, send mail to **habs**.

Those ExI-Essay papers containing explicit notices granting permission for redistribution are available by anonymous FTP at lynx.cs.wisc.edu (IP address 128.105.2.216). A list of available essays along with their file names is in the file pub/ExI-Essay/INDEX. Any questions should go to Derek Zahn at [derek@cs.wisc.edu](mailto:derek@cs.wisc.edu).

There are also two "local" lists for announcements and discussions in the San Francisco Bay Area and the Boston area. To join these lists, send a message to **exi-bay-request** (Bay Area) or **exi-bos-request** (Boston area).