Goborn Go

Apples beget apples, but can machines beget machines? Today it takes an elaborate manufacturing apparatus to build even a simple machine. Could we endow an artificial device with the ability to multiply on its own? Self-replication has long been considered one of the fundamental properties separating the living from the nonliving. Historically our limited understanding of how biological reproduction works has given it an aura of mystery and made it seem unlikely that it would ever be done by a man-made object. It is reported that when René Descartes averred to Queen Christina of Sweden that animals were just another form of mechanical automata, Her Majesty pointed to a clock and said, "See to it that it produces offspring."

The problem of machine self-replication moved from philosophy into the realm of science and engineering in the late 1940s with the work of eminent mathematician and physicist John von Neumann. Some researchers have actually constructed physical replicators. Forty years ago, for example, geneticist Lionel Penrose and his son, Roger (the famous physicist), built small assemblies of plywood that exhibited a simple form of self-replication [see "Self-Reproducing Machines," by Lionel Penrose; SCIENTIFIC AMERICAN, June 1959]. But self-replication has proved to be so difficult that most researchers study it with the conceptual tool that von Neumann developed: twodimensional cellular automata.

Implemented on a computer, cellular automata can simulate a huge variety of self-replicators in what amount to austere universes with different laws of physics from our own. Such models free researchers from having to worry about logistical issues such as energy and physical construction so that they can focus on the fundamental questions of information flow. How is a living being able to replicate unaided, whereas mechanical objects must be constructed by humans? How does replication at the level of an organism emerge from the numerous interactions in tissues, cells and molecules? How did Darwinian evolution give rise to self-replicating organisms?

The emerging answers have inspired the development of selfrepairing silicon chips [see box on page 32] and autocatalyzing molecules [see "Synthetic Self-Replicating Molecules," by Julius Rebek, Jr.; SCIENTIFIC AMERICAN, July 1994]. And this may be just the beginning. Researchers in the field of nanotechnology have long proposed that self-replication will be crucial to manu-

By Moshe Sipper and James A. Reggia Photoillustrations by David Emmite

facturing molecular-scale machines, and proponents of space exploration see a macroscopic version of the process as a way to colonize planets using in situ materials. Recent advances have given credence to these futuristic-sounding ideas. As with other scientific disciplines, including genetics, nuclear energy and chemistry, those of us who study self-replication face the twofold challenge of creating replicating machines and avoiding dystopian prescription could be used in two distinct ways: first, as the instructions whose interpretation leads to the construction of an identical copy of the device; next, as data to be copied, uninterpreted, and attached to the newly created child so that it too possesses the ability to self-replicate. With this two-step process, the self-description need not contain a description of itself. In the architectural analogy, the blueprint would include a plan for building a phothe cellular-automata world. All decisions and actions take place locally; cells do not know directly what is happening outside their immediate neighborhood.

The apparent simplicity of cellular automata is deceptive; it does not imply ease of design or poverty of behavior. The most famous automata, John Horton Conway's Game of Life, produces amazingly intricate patterns. Many questions about the dynamic behavior of cellular

Her Majesty pointed to a clock and said, "See to it that it produces offspring.

dictions of devices running amok. The knowledge we gain will help us separate good technologies from destructive ones.

Playing Life

SCIENCE-FICTION STORIES often depict cybernetic self-replication as a natural development of current technology, but they gloss over the profound problem it poses: how to avoid an infinite regress. A system might try to build a clone using a blueprint-that is, a self-description. Yet the self-description is part of the machine, is it not? If so, what describes the description? And what describes the description of the description? Self-replication in this case would be like asking an architect to make a perfect blueprint of his or her own studio. The blueprint would have to contain a miniature version of the blueprint, which would contain a miniature version of the blueprint and so on. Without this information, a construction crew would be unable to re-create the studio fully; there would be a blank space where the blueprint had been.

Von Neumann's great insight was an explanation of how to break out of the infinite regress. He realized that the self-detocopy machine. Once the new studio and the photocopier were built, the construction crew would simply run off a copy of the blueprint and put it into the new studio.

Living cells use their self-description, which biologists call the genotype, in exactly these two ways: transcription (DNA is copied mostly uninterpreted to form mRNA) and translation (mRNA is interpreted to build proteins). Von Neumann made this transcription-translation distinction several years before molecular biologists did, and his work has been crucial in understanding self-replication in nature.

To prove these ideas, von Neumann and mathematician Stanislaw M. Ulam came up with the idea of cellular automata. A cellular-automata simulation involves a chessboardlike grid of squares, or cells, each of which is either empty or occupied by one of several possible components. At discrete intervals of time, each cell looks at itself and its neighbors and decides whether to metamorphose into a different component. In making this decision, the cell follows relatively simple rules, which are the same for all cells. These rules constitute the basic physics of automata are formally unsolvable. To see how a pattern will unfold, you need to simulate it fully [see Mathematical Games, by Martin Gardner; SCIENTIFIC AMERICAN, October 1970 and February 1971; and "The Ultimate in Anty-Particles," by Ian Stewart, July 1994]. In its own way, a cellular-automata model can be just as complex as the real world.

Copy Machines

WITHIN CELLULAR AUTOMATA, selfreplication occurs when a group of components-a "machine"-goes through a sequence of steps to construct a nearby duplicate of itself. Von Neumann's machine was based on a universal constructor, a machine that, given the appropriate instructions, could create any pattern. The constructor consisted of numerous types of components spread over tens of thousands of cells and required a booklength manuscript to be specified. It has still not been simulated in its entirety, let alone actually built, on account of its complexity. A constructor would be even more complicated in the Game of Life because the functions performed by single cells in von Neumann's model-such as transmission of signals and generation of new components-have to be performed by composite structures in Life.

Going to the other extreme, it is easy to find trivial examples of self-replication. For example, suppose a cellular automata has only one type of component, labeled +, and that each cell follows only a single rule: if exactly one of the four neighboring

MOSHE SIPPER and JAMES A. REGGIA share a long-standing interest in how complex systems can self-organize. Sipper is a senior lecturer in the department of computer science at Ben-Gurion University in Israel and a visiting researcher at the Logic Systems Laboratory of the Swiss Federal Institute of Technology in Lausanne. He is interested mainly in bio-inspired computational paradigms such as evolutionary computation, self-replicating systems and cellular computing. Reggia is a professor of computer science and neurology, working in the Institute for Advanced Computer Studies at the University of Maryland. In addition to studying self-replication, he conducts research on computational models of the brain and its disorders, such as stroke.

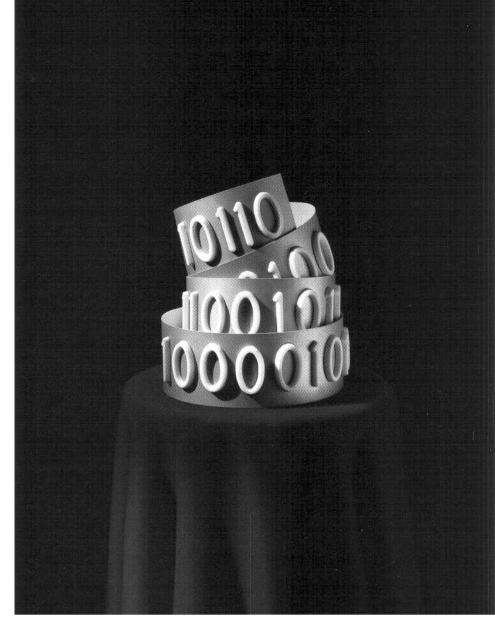
THE AUTHORS

cells contains a +, then the cell becomes a +; otherwise it becomes vacant. With this rule, a single + grows into four more +'s, each of which grows likewise, and so forth.

Such weedlike proliferation does not shed much light on the principles of replication, because there is no significant machine. Of course, that invites the question of how you would tell a "significant" machine from a trivially prolific automata. No one has yet devised a satisfactory answer. What is clear, however, is that the replicating structure must in some sense be complex. For example, it must consist of multiple, diverse components whose interactions collectively bring about replication-the proverbial "whole must be greater than the sum of the parts." The existence of multiple distinct components permits a self-description to be stored within the replicating structure.

In the years since von Neumann's seminal work, many researchers have probed the domain between the complex and the trivial, developing replicators that require fewer components, less space or simpler rules. A major step forward was taken in 1984 when Christopher G. Langton, then at the University of Michigan, observed that looplike storage devices-which had formed modules of earlier self-replicating machines-could be programmed to replicate on their own. These devices typically consist of two pieces: the loop itself, which is a string of components that circulate around a rectangle, and a construction arm, which protrudes from a corner of the rectangle into the surrounding space. The circulating components constitute a recipe for the loop-for example, "go three squares ahead, then turn left." When this recipe reaches the construction arm, the automata rules make a copy of it. One copy continues around the loop; the other goes down the arm, where it is interpreted as instructions.

By giving up the requirement of universal construction, which was central to von Neumann's approach, Langton showed that a replicator could be constructed from just seven unique components occupying only 86 cells. Even smaller and simpler self-replicating loops have been devised by one of us (Reggia) and our colleagues [see box on next page]. Be-



cause they have multiple interacting components and include a self-description, they are not trivial. Intriguingly, asymmetry plays an unexpected role: the rules governing replication are often simpler when the components are not rotationally symmetric than when they are.

Emergent Replication

ALL THESE SELF-REPLICATING structures have been designed through ingenuity and much trial and error. This process is arduous and often frustrating; a small change to one of the rules results in an entirely different global behavior, most likely the disintegration of the structure in question. But recent work has gone beyond the direct-design approach. Instead of tailoring the rules to suit a particular type of structure, researchers have experimented with various sets of rules, filled the cellular-automata grid with a "primordial soup" of randomly selected components and checked whether selfreplicators emerged spontaneously.

In 1997 Hui-Hsien Chou, now at Iowa State University, and Reggia noticed that as long as the initial density of the free-floating components was above a certain threshold, small self-replicating loops reliably appeared. Loops that collided underwent annihilation, so there was an ongoing process of death as well as birth. Over time, loops proliferated, grew in size and evolved through mutations triggered by debris from past collisions. Although the automata rules were deterministic, these mutations were effectively random, because the system was complex and the components started in random locations.

Such loops are intended as abstract machines and not as simulacra of anything biological, but it is interesting to compare them with biomolecular structures. A loop loosely resembles circular DNA in bacteria, and the construction arm acts as the enzyme that catalyzes DNA replication. More important, replicating loops illustrate how complex global behaviors can arise from simple local interactions. For example, components move around a loop even though the rules say nothing about movement; what is actually happening is that individual cells are coming alive, dying or metamorphosing in such a way that a pattern is eliminated from one position and reconstructed elsewhere—a process that we perceive as motion. In short, cellular automata act locally but appear to think globally. Much the same is true of molecular biology.

In a recent computational experiment,

Jason Lohn, now at the NASA Ames Research Center, and Reggia experimented not with different structures but with different sets of rules. Starting with an arbitrary block of four components, they found they could determine a set of rules that made the block self-replicate. They discovered these rules via a genetic algorithm, an automated process that simulates Darwinian evolution.

The most challenging aspect of this work was the definition of the so-called

BUILD YOUR OWN REPLICATOR

SIMULATING A SMALL self-replicating loop using an ordinary chess set is a good way to get an intuitive sense of how these systems work. This particular cellular-automata model has four different types of components: pawns, knights, bishops and rooks. The machine initially comprises four pawns, a knight and a bishop. It has two parts: the loop itself, which consists of a two-by-two square, and a construction arm, which sticks out to the right.

The knight and bishop represent the self-description: the knight, whose orientation is significant, determines which direction to grow, while the bishop tags along and determines how long the side of the loop should be. The pawns are fillers that define the rest of the shape of the loop, and the rook is a transient signal to guide the growth of a new construction arm.

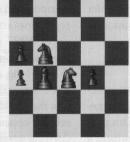
As time progresses, the knight and bishop circulate counterclockwise around the loop. Whenever they encounter the arm, one copy goes out the arm while the original continues around the loop. HOW TO PLAY: You will need two chessboards: one to represent the current configuration, the other to show the next configuration. For each round, look at each square of the current configuration, consult the rules and place the appropriate piece in the corresponding square on the other board. Each piece metamorphoses depending on its identity and that of the four squares immediately to the left, to the right, above and below. When you have reviewed each square and set up the next configuration, the round is over. Clear the first board and repeat. Because the rules are complicated, it takes a bit of patience at first. You can also view the simulation at Islwww.epfl.ch/chess

The direction in which a knight faces is significant. In the drawings here, we use standard chess conventions to indicate the orientation of the knight: the horse's muzzle points forward. If no rule explicitly applies, the contents of the square stay the same. Squares on the edge should be treated as if they have adjacent empty squares off the board. —*M.S. and J.A.R.*

STAGES OF REPLICATION



INITIALLY, the selfdescription, or "genome"—a knight followed by a bishop—is poised at the start of the construction arm.



1 The knight and bishop move counterclockwise around the loop. A clone of the knight heads out the arm.



2 The original knightbishop pair continues to circulate. The bishop is cloned and follows the new knight out the arm.



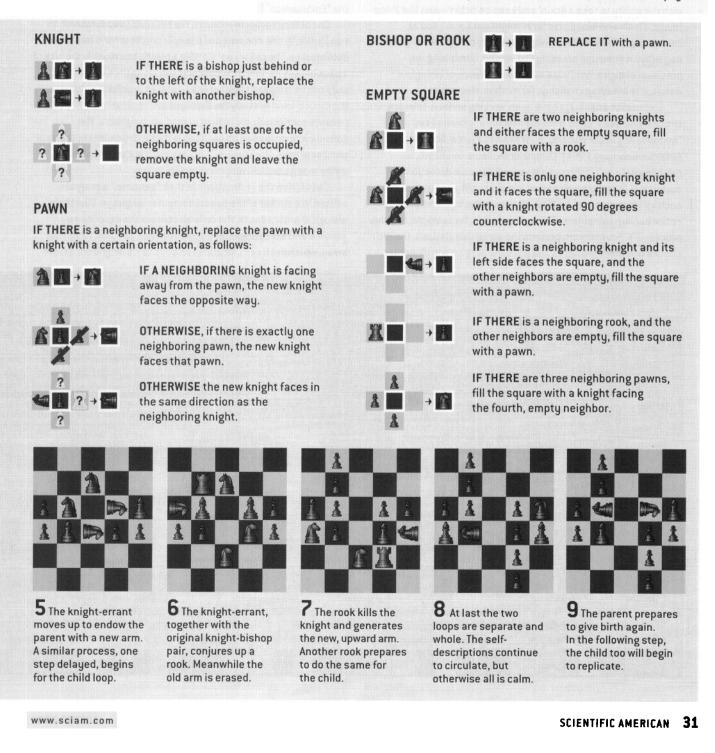
3 The knight triggers the formation of two corners of the child loop. The bishop tags along, completing the gene transfer.

4 The knight forges the remaining corner of the child loop. The loops are connected by the construction arm and a knight-errant.

fitness function—the criteria by which sets of rules were judged, thus separating good solutions from bad ones and driving the evolutionary process toward rule sets that facilitated replication. You cannot simply assign high fitness to those sets of rules that cause a structure to replicate, because none of the initial rule sets is likely to allow for replication. The solution was to devise a fitness function composed of a weighted sum of three measures: a growth measure (the extent to which each component type generates an increasing supply of that component), a relative position measure (the extent to which neighboring components stay together) and a replicant measure (a function of the number of actual replicators present). With the right fitness function, evolution can turn rule sets that are sterile into ones that are fecund; the process usually takes 150 or so generations.

Self-replicating structures discovered in this fashion work in a fundamentally

different way than self-replicating loops do. For example, they move and deposit copies along the way—unlike replicating loops, which are essentially static. And although these newly discovered replicators consist of multiple, locally interacting components, they do not have an identifiable self-description—there is no obvious genome. The ability to replicate without a self-description may be relevant to questions about how the earliest biological *Continued on page 35*



ROBOT, HEAL THYSELF

Computers that fix themselves are the first application of artificial self-replication

LAUSANNE, SWITZERLAND—Not many researchers encourage the wanton destruction of equipment in their labs. Daniel Mange, however, likes it when visitors walk up to one of his inventions and press the button marked KILL. The lights on the panel go out; a small box full of circuitry is toast. Early in May his team unveiled its latest contraption at a science festival here—a wall-size digital clock whose components you can zap at will—and told the public: Give it your best shot. See if you can crash the system.

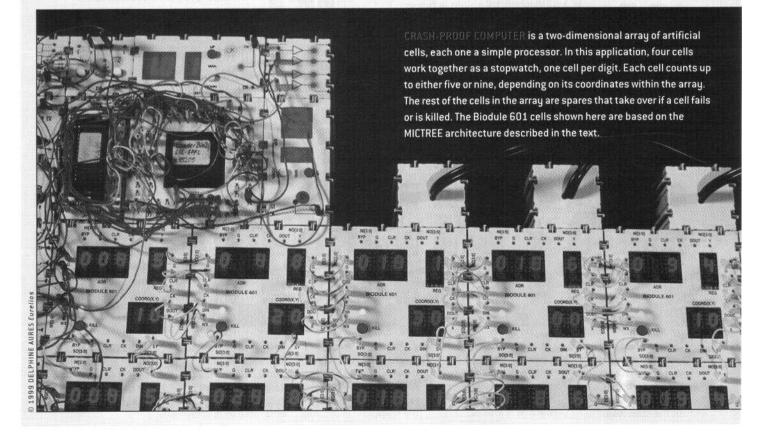
The goal of Mange and his team is to instill electronic circuits with the ability to take a lickin' and keep on tickin'—just like living things. Flesh-and-blood creatures might not be so good at calculating π to the millionth digit, but they can get through the day without someone pressing Ctrl-Alt-Del. Combining the precision of digital hardware with the resilience of biological wetware is a leading challenge for modern electronics.

Electronics engineers have been working on fault-tolerant circuits ever since there were electronics engineers [see "Redundancy in Computers," by William H. Pierce; SCIENTIFIC AMERICAN, February 1964]. Computer modems would still be dribbling data at 1200 baud if it weren't for error detection and correction. In many applications, simple quality-control checks, such as extra data bits, suffice. More complex systems provide entire backup computers. The space shuttle, for example, has five processors. Four of them perform the same calculations; the fifth checks whether they agree and pulls the plug on any dissenter. The problem with these systems, though, is that they rely on centralized control. What if that control unit goes bad?

Nature has solved that problem through radical decentralization. Cells in the body are all basically identical; each takes on a specialized task, performs it autonomously and, in the event of infection or failure, commits hara-kiri so that its tasks can be taken up by new cells. These are the attributes that Mange, a professor at the Swiss Federal Institute of Technology here, and others have sought since 1993 to emulate in circuitry, as part of the "Embryonics" (embryonic electronics) project.

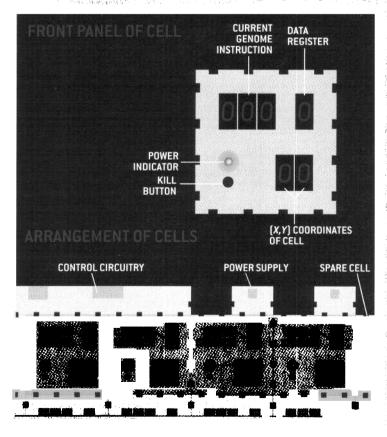
One of their earlier inventions, the MICTREE (microinstruction tree) artificial cell, consisted of a simple processor and four bits of data storage. The cell is contained in a plastic box roughly the size of a pack of Post-its. Electrical contacts run along the sides so that cells can be snapped together like Legos. As in cellular automata, the models used to study the theory of self-replication, the MICTREE cells are connected only to their immediate neighbors. The communication burden on each cell is thus independent of the total number of cells. The system, in other words, is easily scalable unlike many parallel-computing architectures.

Cells follow the instructions in their "genome," a program written in a subset of the Pascal computer language. Like their biological antecedents, the cells all contain the exact same genome and execute part of it based on their position within the array, which each cell calculates relative to its neighbors. Waste-



fol though it may soom, this redundancy allows the array to withstand the loss of any cell. Whenever someone presses the KUL button on a cell, that cell shuts down, and its left and right neighbore because directly connected. The right neighbor recalculates its position and starts executing the deceased's program. Its tanks, in turn, are taken up by the next cell to the right, and so on, with a coll designated as a spare is pressed into service. Writing programs for any parallel processor is tricky, but the 2.2 MICTREE array requires an especially unconventional approach. instead of giving explicit instructions, the programmer must devise simple rules out of which the desired function will emerge. Beine Swies, Mange demonstrates by building a superreliable stopwatch. Displaying minutes and seconds requires four cells in a row, one for each digit. The genome allows for two cell tupes: a counter from zero to nine and a counter from zero to five. An oscillator feeds one pulse per second into the rightmost cell. After 10 pulses, this cell cycles back to zero and sends a pulse to the cell on its left, and so on down the line. The watch takes up part of an array of 12 cells: when you kill one, the clock transplants itself one cell over and carries on. Obviously, though, there is a limit to its resilience: the whole thing will fail after, at most, eight kills.

The prototype MICTREE cells are hardwired, so their processing power cannot be tailored to a specific application. In a finished product, cells would instead be implemented on a fieldprogrammable gate array, a grid of electronic components that can be reconfigured on the fly [see "Configurable Computing," by John Villesenor and William H. Mangione-Smith; Scientific American, June 1997]: Mange's team is now custom-designing a gate array,



known as MUXTREE (multiplexentree), that is optimized for artificial cells. In the biological metaphor, the components of this array are the "molecules" that constitute a cell. Each consists of a logic gate, a data bit and a string of configuration bits that determines the function of this gate.

Building a cell out of such molecules offers not only fieldbillity but also extre endurance. Each molecule containe two capies of the gate and three of the storage bit. If the two gates ever glue different results, the molecule kills itself for the greater goad of the cell. As a last gasp, the molecule sends its data bit (preserved by the triplicate storage) and configuration to its right neighbor, which does the same, and the process continues until the rightmost molecule transfers its date to a spare. This second level of fault tolerance prevents a single error from wiping out an entire cell.

A total of 2,000 molecules, divided into four 20-by 25 cells, make up the BioWall—the giant digital clock that Mange's team has just put on display. Each molecule is enclosed in a small box and includes a KILL button and an LED display. Some mulecules are configured to perform computations; others serve as pixels in the clock display. Making liberal use of the KILL buttons, I did my utmost to crash the system, something I'm usually quite good at. But the plucky clock just wouldn't submit. The clock display did start to look funny—numerals bent over as their pixels shifted to the right—but at least it was still legible, unlike most faulty electronic stana.

That said, the system did suffer from display glitches, which Mange attributed mainly to timing problems. Although the processing power is decentralized, the cells still rely on a central oscillator to coordinate their communications; sometimes they fall out of sync. Another Embryonics team, led by Andy Tyrrell of the University of York in England, has been studying making the cells asynchronous, like their biological counterparts. Cells would generate handshaking signals to orchestrate data transfers. The present system is also unable to catch certain types of error, including damaged configuration strings. Tyrrell's team has proposed adding watchdog molecules—an immune system—that would monitor the configurations [and one another] for defects.

Although these systems demand an awful lot of overhead, so do other fault-tolerance technologies. "While Embryonics appears to be heavy on redundancy, it actually is not that bad when compared to other systems," Tyrrell argues. Moreover, MUCIREE should be easier to scale down to the nano level; the "molecules" are simple enough to really be molecules. Says Mange, "We are preparing for the situation where electronics will be at the same scale as biology."

On a philosophical level, Embryonics comes very close to the dream of building a self-replicating machine. It may not be quite as dramatic as a robot that can go down to Radio Shack, pull parts off the racks, and take them home to resolder a connection or build a loving mate. But the effect is much the same, Letting machines determine their own destiny—whether reconfiguring themselves on a silicon chip or reprogramming themselves using a neural network or genetic algorithm—sounds scary, but perhaps we should be gratified that machines are becoming more like us: imperfect, fallible but stubbornly resourceful. —Seorge Musser, imperfect but resourceful stoff editor and writer

Continued from page 31

replicators originated. In a sense, researchers are seeing a continuum between nonliving and living structures.

Many researchers have tried other computational models besides the traditional cellular automata. In asynchronous cellular automata, cells are not updated in concert; in nonuniform cellular automata, the rules can vary from cell to cell. Another approach altogether is Core War [see Computer Recreations, by A. K. Dewdney; SCIENTIFIC AMERICAN, May 1984] and its successors, such as ecologist Thomas S. Ray's Tierra system. In these nents, one for the program and the other for data. The loops can execute an arbitrary program in addition to self-replicating. In a sense, they are as complex as the computer that simulates them. Their main limitation is that the program is copied unchanged from parent to child, so that all loops carry out the same set of instructions.

In 1998 Chou and Reggia swept away this limitation. They showed how selfreplicating loops carrying distinct information, rather than a cloned program, can be used to solve a problem known as satisfiability. The loops can be used to determine whether the variables in a logical exsigning a parallel computer from either transistors or chemicals [see "Computing with DNA," by Leonard M. Adleman; SCIENTIFIC AMERICAN, August 1998].

In 1980 a NASA team led by Robert Freitas, Jr., proposed planting a factory on the moon that would replicate itself, using local lunar materials, to populate a large area exponentially. Indeed, a similar probe could colonize the entire galaxy, as physicist Frank J. Tipler of Tulane University has argued. In the nearer term, computer scientists and engineers have experimented with the automated design of robots [see "Dawn of a New Species?" by George

In a sense, researchers are seeing a continuum between nonliving and living structures.

simulations the "organisms" are computer programs that vie for processor time and memory. Ray has observed the emergence of "parasites" that co-opt the selfreplication code of other organisms.

Getting Real

SO WHAT GOOD are these machines? Von Neumann's universal constructor can compute in addition to replicating. but it is an impractical beast. A major advance has been the development of simple yet useful replicators. In 1995 Gianluca Tempesti of the Swiss Federal Institute of Technology in Lausanne simplified the loop self-description so it could be interlaced with a small program-in this case, one that would spell the acronym of his lab, "LSL." His insight was to create automata rules that allow loops to replicate in two stages. First the loop, like Langton's loop, makes a copy of itself. Once finished, the daughter loop sends a signal back to its parent, at which point the parent sends the instructions for writing out the letters.

Drawing letters was just a demonstration. The following year Jean-Yves Perrier, Jacques Zahnd and one of us (Sipper) designed a self-replicating loop with universal computational capabilities—that is, with the computational power of a universal Turing machine, a highly simplified but fully capable computer. This loop has two "tapes," or long strings of compopression can be assigned values such that the entire expression evaluates to "true." This problem is NP-complete—in other words, it belongs to the family of nasty puzzles, including the famous travelingsalesman problem, for which there is no known efficient solution. In Chou and Reggia's cellular-automata universe, each replicator received a different partial solution. During replication, the solutions mutated, and replicators with promising solutions were allowed to proliferate while those with failed solutions died out.

Although various teams have created cellular automata in electronic hardware, such systems are probably too wasteful for practical applications; automata were never really intended to be implemented directly. Their purpose is to illuminate the underlying principles of replication and, by doing so, inspire more concrete efforts. The loops provide a new paradigm for deMusser; SCIENTIFIC AMERICAN, November 2000]. Although these systems are not truly self-replicating—the offspring are much simpler than the parent—they are a first step toward fulfilling the queen of Sweden's request.

Should physical self-replicating machines become practical, they and related technologies will raise difficult issues, including the Terminator film scenario in which artificial creatures outcompete natural ones. We prefer the more optimistic, and more probable, scenario that replicators will be harnessed to the benefit of humanity [see "Will Robots Inherit the Earth?" by Marvin Minsky; SCIENTIFIC AMERICAN, October 1994]. The key will be taking the advice of 14th-century English philosopher William of Ockham: entia non sunt multiplicanda praeter necessitatem-entities are not to be multiplied beyond necessity.

MORE TO EXPLORE

Simple Systems That Exhibit Self-Directed Replication. J. Reggia, S. Armentrout, H. Chou and Y. Peng in Science, Vol. 259, No. 5099, pages 1282–1287; February 26, 1993. Emergence of Self-Replicating Structures in a Cellular Automata Space. H. Chou and J. Reggia in Physica D, Vol. 110, Nos. 3–4, pages 252–272; December 15, 1997. Special Issue: Von Neumann's Legacy: On Self-Replication. Edited by M. Sipper, G. Tempesti, D. Mange and E. Sanchez in Artificial Life, Vol. 4, No. 3; Summer 1998. Towards Robust Integrated Circuits: The Embryonics Approach. D. Mange, M. Sipper, A. Stauffer and G. Tempesti in Proceedings of the IEEE, Vol. 88, No. 4, pages 516–541; April 2000. Moshe Sipper's Web page on artificial self-replication is at Islwww.epfl.ch/~moshes/selfrep/ Animations of self-replicating loops can be found at necsi.org/postdocs/sayama/sdsr/java/ For John von Neumann's universal constructor, see alife.santafe.edu/alife/topics/jvn/jvn.html