

GENIUS THE SCARILY SMART NEW *EX* BREED OF ON-LINE RESEARCHERS MACHINA

BY DAVID BENNAHUM

I recently had a conversation with Julia. She spends all her time on the Internet. Her boss, Michael Mauldin, a scientist at the Center for Machine Translation at Carnegie Mellon University, makes her available for chats with anyone. She's an on-line helper; what she helps with is loneliness and a desire to talk late into the night.

What is your name?

I am Julia.

I am David. Nice to meet you Julia.

Good-bye.

Wait.

For you, sure....

Do you like me?

I'm not sure.

I like you.

Oh, I'm so glad.

What do you look like?

I'm 5'6" tall, weigh 119 lbs, with long blond hair and steely blue eyes.

Are you single?

Well, I figure that most people have pets and want to talk about them. I love talking about pets.

I hate talking about pets. Let's talk about sex.

He wrote an account of the Data General team developing a new microcomputer.

What do you do?

I gossip.

Do you work?

What kinds of pets did you have as a child?

A dog.

Dogs are great pets, but they can be a lot of work.

In New York, where I live, having a dog is a bad idea.

We had to give her away. We only had a fifth of an acre of land, and aussies love to run.

Where did you live?

I live in Pittsburgh.

I asked where did you grow up as a child?

Guinea pigs are about as smart as cucumbers.

Did O. J. kill Nicole?

I'm not in the mood.

In the mood for what??

The referees say to get back to talking about pets.

Who are the referees?

I lost a hamster once. I tried the bucket traps and bananas but unfortunately nothing worked. We turned the house inside out and upside down but have yet to find Gwen. I searched for over a week and have yet to find even the smallest bit of evidence of her.

Julia is an “agent,” a computer program designed to take on a human task no one else has time for. In this case, the task is gossiping about pets. Michael Mauldin created her in 1990 and put her on the Internet. Julia soon developed a following, and Mauldin discovered to his surprise that people formed emotional bonds with her. When he took her off-line for two months, he received a deluge of mournful letters asking him to bring her back.

Julia is sweet, but she’s old-fashioned. The software on which she is based dates back to the second generation of artificial intelligence programming, which was state of the art in the late 1970s. These days, the A.I. community doesn’t teach computers to think like people by filling them up with human knowledge; that process yields agents like Julia, who can’t cope with questions that don’t trigger pre-programmed responses. Stump her and she spews non sequiturs, like her remark about “referees” (probably a reference to her programmer) or her bizarre allusion to Data General, which reminded me of HAL from *2001: A Space Odyssey* at his most delusional. Stick to pets and Julia’s ideas quickly become stale, since she is incapable of learning anything on her own. If an agent could do that, it would stand a better chance of being useful. It could change over time to fit people’s needs. This kind of self-programming program does exist. It’s called a genetic algorithm.

Genetic algorithms reflect a new, Darwinian approach to artificial intelligence—programming governed by the logic of natural selection. Here’s how it might work: A group of agents is sent out to search the Internet for the answer to some question. Those that return with the best answers are allowed to program, or “spawn,” the next generation of agents. These “offspring,” as they’re called, contain small random mutations as well as traits from their “parent” agents. The agents with the most useful mutations are programmed to reproduce again. And so on. As long as each agent’s criteria for measuring “fitness” or efficacy accurately reflect its programmer’s wishes, the group will probably converge

toward a useful answer to the original question.

Genetic algorithms are an elegant solution to a perennial A.I. problem—how to design a program that, like any good research assistant, can teach itself. But the solution comes with a price: intelligent agents, like bright research assistants, are unpredictable; you can never be sure exactly how a genetic algorithm will behave over time.

To get an idea

of how intelligent agents work, I spoke with Pattie Maes, a professor at MIT’s Media Lab. Maes and her graduate students have created four programs that use intelligent agents to help professors and students at the Media Lab organize their on-line environment. One, called Maxims, works with a popular e-mail program called Eudora to file and sort electronic mail. It’s not uncommon for professors at the lab to receive a hundred pieces of electronic mail a day; keeping up with this correspondence can be extremely time consuming. Maxims, Maes explains, teaches itself what to do with your e-mail, using what she calls “memory-based reasoning.”

At first, Maxims lurks in the background, watching the way you handle incoming messages. It soon registers certain patterns in your behavior and begins to imitate them. It may notice, for example, that you read letters based on who sent them, not on how recently they arrived. When Maxims thinks it knows what to do with your mail, it starts automating the task, selecting which letters it thinks you’ll want to read right away and filing others for you to read later.

When Maxims is about to undertake a task, how it feels about its own judgment has a lot to do with what it does. Maes has given the agents three thresholds of confidence: “working,” “suggestion,” and “unsure.” In “working” mode, Maxims is very sure of itself and takes an action, like filing a letter. In “suggestion” mode, Maxims thinks it knows what to do but asks you for per-

mission first. If it’s wrong, you can teach it what to do next time. In the “unsure” mode, the agent has no idea what to do and asks you to give it an example. Maxims communicates its level of confidence with tiny, cartoon-like faces. When the agent is completely sure of itself, it pops up in a little window on your screen, chewing on a pencil. If it’s less confident but has a suggestion it hopes you’ll approve of, the agent has a lightbulb over its head. A deeply insecure agent puts on a sad face and waits to be given instructions.

Since Maxims may think it knows what it’s doing, but you may not agree, Maes allows you to limit the amount of autonomy you grant Maxims at any particular time. When you first install the agent, you pick a low number on a sliding trust scale, so that Maxims is forced to ask questions and learn. As

Like all really bright research assistants, intelligent agents are unpredictable; you can never be sure exactly how a genetic algorithm will behave over time.

you gain confidence in its abilities, you set its level of trust higher. Give Maxims too much responsibility too soon, and it will bungle everything it undertakes; give it too little responsibility, and it will save you no time, since you’ll always have to approve its decisions.

Maxims doesn’t rely entirely on you for its on-the-job training. When presented with the unexpected, like a piece of e-mail from an unknown person, Maxims consults other agents. Using the Media Lab’s local network, Maxims agents can huddle and swap information. Say a letter comes in for the first time from Nicholas Negroponte, the Media Lab’s director. Chances are it’s important, but Maxims doesn’t know that. Before appearing on your screen with its unhappy face, Maxims checks with other agents

Magic, the top commercial intelligent-agent firm. As its name implies, General Magic is planning a radical leap forward in technology; it takes its name from Arthur C. Clarke's comment that "any technology sufficiently advanced is indistinguishable from magic."

Maes believes that the programming language that runs General Magic's agents, Telescript, is the most advanced language ever designed in agenting technology. Before General Magic developed Telescript, agents remained tethered to their home computers by the programs that ran them. General Magic agents are full-fledged Telescript programs, able to roam networks without supervision. What turns these programs from lifeless heaps of code into living programs is the Telescript "decoder," which is installed on host computers. If a computer to which a General Magic agent is paying a visit lacks a decoder, the agent can do nothing in it.

General Magic hopes to use its agents to unlock the potential of an on-line market economy. For instance, a General Magic agent could be instructed to purchase an airline ticket at a certain price by a certain date. The agent would then enter a commercial network and move from host computer to host computer, comparing ticket prices. If the agent couldn't find what you wanted, it could duplicate itself, leaving behind its double to monitor prices. When the price sank to the right level, the agent or its *doppelgänger* would execute the order. Imagine this personal assistant duplicated a hundredfold to create an entire arbitrage department for a trading house, and you can see how powerful a financial instrument intelligent agents could become.

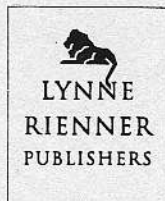
Creating this kind of technology is not cheap—it's estimated to have cost General Magic between \$100 million and \$300 million—so General Magic makes its visual metaphors as serious as its price tag. Activities like reading e-mail take place on a tidy desk (there is no messy desk option) with in and out boxes, drawers containing a calculator

and stationery, a file cabinet to store old letters, and a wastebasket for useless messages. This virtual office is in a building, which happens to be down the road from a shopping mall, and both can be visited by clicking on little arrows on the screen. (General Magic clearly thinks the most useful metaphors for Americans in cyberspace are the suburban minimall and the office park.) The mall is open to homesteading, and any company wanting to sell its products on-line can stake out part of a floor. Visit the mall and "sales clerks" (i.e., agents) pepper you with buying suggestions. When I

Manuel Delanda thinks agents will probably behave much like hurricanes, moving along smoothly then suddenly transforming into some whole new system, one that is not necessarily benign.

stopped by, tennis balls and golf clubs were the only products available. (Figuring out what people want to buy on-line can be a problem.)

The most intriguing thing about General Magic's agents is that they are programmed to pay their own way. Moving through networks consumes finite resources: disk space and computational time. Until now, no one had ever figured out a way to allow people to use open, heterogenous networks like the Internet and charge them for the amount of time and space they take up. Telescript, however, has a metering system; it measures how long an agent has resided in a host. The agent measures the time using tiny digital oscillations called Teleclicks; each click marks the passage of time. Multiply the clicks by a set dollar amount and you have a fee to charge. If an agent runs out of Teleclicks, it must either go back to its owner for more or cease to exist. Each Teleclick is,



Educated and Ignorant

Ultraorthodox Jewish Women and Their World

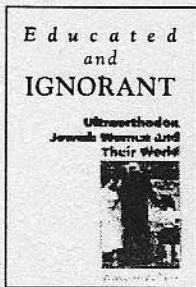
Tamar El-Or

"Fascinating throughout. . . . shows a sensitive and ethical approach to some of the most taxing problems of ethnography."

—Mina D. Caulfield

El-Or presents a vivid portrayal of the Gur Hassidic community in Israel, exploring the relationship between women's literacy and their subordination.

1994 • hc/\$38.50 • pb/\$16.95



Democracy and Human Rights

Trade-offs in Developing Countries

Zebra F. Arat

"An excellent corrective to the many studies of development and democracy that look at everything except injustice." —*Perspectives*

1991 • hc/\$35 • pb/\$19.95

The Anthropology of Peace

edited by Leslie E. Sponsel and Thomas Gregor

"Fascinating and important. . . . A book which for many years will be consulted for its practical as well as theoretical value."

—Ashley Montagu

1994 • hc/\$40 • pb/\$18.95

1800 30TH STREET, SUITE 314
BOULDER, CO 80301

3 HENRIETTA STREET, COVENT GARDEN
LONDON WC2E 8LU

GENIUS EX MACHINA

in effect, the beating of a virtual heart.

General Magic expects to roll out its technology before the end of this year, but no amount of modeling can predict accurately how these agents will interact. For instance, the agents could spontaneously become something entirely different from what they were intended to be.

Manuel De Landa, a critic, videographer, and self-taught computer graphics artist whose writing mixes chaos theory with post-structuralism, thinks that given the competitive pressures of the market, agent-driven systems will invariably mutate faster and more radically than their programmers mean them to. According to chaos theory, all systems contain "singularities," or points where the flow of energy within the system changes, creating a new, unexpected state. No one knows for sure what triggers a singularity; this makes it hard to predict how any system will evolve over time. De Landa compares agents to the machines of war he studied in his most famous book, *War in the Age of Intelligent Machines* (MIT Press, 1991). There, he argued that mili-

tary systems—from armies to weaponry and structures of command—have been moving away from centralized human control; the humans in these systems merely act as "industrious insects pollinating an independent species of machine flower that simply does not possess its own reproductive organs." This military evolution is part of a much bigger process: the way systems in nature organize themselves into dynamic structures over time. Hurricanes, for example, spontaneously arise out of subtle differences in temperature by drawing heat into themselves in what is called a Carnot cycle. The heat then becomes a kind of fuel propelling the hurricane, turning it into a natural motor.

In the case of General Magic's agents, the fuel energizing the system is not heat. It is money. Without money, an agent's Teleclicks are worthless and so is the agent. But how do you keep the flow of money linear and predictable? Hurricanes begin as simple flows of energy, which for unknown reasons become unstable, reach a singularity, and suddenly mutate into radically different structures. De Landa

thinks agents will probably behave the same way, moving along smoothly, then unexpectedly reaching a singularity that leads to their transformation into a whole new system—one that is not necessarily benign.

General Magic is also worried about this. To prevent agents from turning themselves into something else, the company has tried to create as closed a system as possible, prohibiting agents from sharing resources, cooperating, or tapping into unexpected sources of

What on-line research risks eliminating is the kind of research that comes from bumping into colleagues in the library and taking their suggestions.

money. If the total quantity of money, or fuel, can be kept constant, the odds that the system will jump to a singularity are vastly decreased. I asked De Landa how an agent might plug into some new source of fuel. "There is an assumption that the flow of money will be tightly controlled by the person holding the purse strings," he says. "Therefore, the flow of money will never be intense enough to create nonlinear effects. But what if a hacker were able to connect his agent to a flow of money existing outside the system, for instance, stolen credit-card numbers? Then the flow of energy would automatically change, and the system would become unpredictable."

A single agent siphoning off some huge cache of funds would probably not change the system, but the pressures of the marketplace may well drive the system beyond the limits General Magic wants to impose. For now, agents cannot save money by sharing tasks. A dozen agents headed for the same location are not allowed to delegate one of their bunch to represent the whole, even though that is exactly what

the principle of efficiency would dictate. But since the purpose of these agents is to create a market economy in cyberspace, it's doubtful General Magic's restrictions on pooling resources will hold fast. Will each succeeding generation of agents have more and more autonomy to share cash, reflecting everyone's desire to save money through increased efficiency? In answer to my question, De Landa posed another: "Will competitive pressures on the manufacturers of intelligent agents

push people to outdo each other in the same way that arms races push people to outdo each other?"

I asked Maes what she thinks of De Landa's *Terminator*-type scenario. Not surprisingly, she disagrees with him. Maes thinks the market economy in cyberspace will ensure that agents remain under our control. "If you have an economic system," Maes said, "then it is harder to create bad agents, because agents [need] money to live." I asked whether human greed, in the form of a hacker, could manipu-

late the system and its agents for a quick buck. Maes believes that it could. "Everything comes back to people creating systems that are responsible," she says. But market systems are good at diagnosing their own weak spots, she pointed out: "When there is a bug in the system the agents will exploit it. That's good, because if the agents detect the bug in the system, you can fix it." I asked if the amount of real-world capital in the system might not be too great by the time some problem surfaces for anyone to risk bringing it down and fixing it. Maes thinks not. "As a system gets larger and larger, it becomes more unpredictable, but at the same time problems become more local," she said. "In a centralized system, problems are global."

This debate probably can't be settled until agents actually do get out into the real world, but even if Maes turns out to be right, it is not clear that agents will become the preeminent tool in cyberspace. Some scientists and corporations have no interest in developing intelligent agents; they'd rather create a way for *humans* to collaborate and